Type-adapted representations of semidirect product groups

P. Kasperkovitz
Institut für theoretische Physik, Technische Universität, A-1040 Wien, Karlsplatz 13, Austria

(Received 7 August 1981; accepted for publication 16 September 1981)

Irreducible representations of compact groups can be partitioned into three classes (character test $+,0,-$). This classification is the same for real, complex, and quaternionic representations and in all three cases a peculiar, type-adapted form of the representation matrices may be chosen (t reps). In this paper it is shown how to construct t reps of semidirect products $G \circledS g$ starting with t reps of G and t reps of some covering groups of subgroups of g. The advantage of using t reps shows up in that the factor system of the little cogroups is real in two of three cases and that real, complex, and quaternionic representations are obtained simultaneously. The method is specialized to direct products and generalized to induction from normal subgroups.

PACS numbers: 02.20. + b

1. INTRODUCTION

Matrix representations of compact groups irreducible over the field of complex numbers can be divided into three disjoint classes: the representation either can be brought into real form, or is inequivalent to its complex conjugate, or is equivalent to its complex conjugate, but cannot be brought into real form. This partition does not only hold for complex representations but is also valid for real and quaternionic representations.[1,2] No matter what representation field is chosen, belonging to a certain class means always that the representation is *essentially* a real or a complex or a quaternionic matrix representation. This fact is often obscured by similarity transformations but it is always possible to make the intrinsic algebraic structure transparent: If the field characterizing the type of the representation is a subfield of the representation field (e.g. $\mathbb{R} \subset \mathbb{C}$) this is achieved by finding a matrix representation with elements in the subfield. If on the other hand the representation is characterized by an extension of the representation field (e.g. $\mathbb{Q} \supset \mathbb{C}$) then the matrix can be composed of small square matrices each having a peculiar structure which is typical for the extension field.

To choose type-adapted representations is to eliminate redundant information from the beginning. In fact, as can be seen from the definitions given in Sec. 2, this choice reduces the number of real parameters needed to fix a matrix representation by $\frac{1}{3}$ for complex representations (averaged over all three types) and by $\frac{5}{12}$ for noncomplex ones. To consider real and quaternionic representations of a group along with the familiar complex ones is both of mathematical and physical interest. For these representations have been shown to be equivalent to the combination of a complex representation and a commuting antiunitary operator and, if chosen in a type-adapted form, to simplify the matrix representation of invariant operators.[1,3] It therefore seems worthwhile to discuss the construction and properties of type-adapted representations over the reals, the complex numbers, and the quaternions. The topic considered in this paper is the construction of type-adapted representations of semidirect product groups. In this problem the advantage of the representations considered here shows up in that the factor system of the little cogroups is real in two of three cases and that real, complex, and quaternionic representations are obtained in one run.

2. TYPE-ADAPTED REPRESENTATIONS (t REPS)

We consider matrix representarions of compact groups over a (skew) field $\mathbb{F}'$ of characteristics zero (reals, complex numbers, quaternions). A matrix representation with elements from $\mathbb{F}'$ is called real if $\mathbb{F}' = \mathbb{R}$, complex if $\mathbb{F}' = \mathbb{C}$, and quaternionic if $\mathbb{F}' = \mathbb{Q}$. It is said to be of $\mathbb{F}$ type and adapted to its type if one of the following conditions is satisfied:

(i) If $\mathbb{F} = \mathbb{F}'$ the matrices $D_{\mathbb{F}'}(x)$, $x \in G$, are norm preserving and absolutely irreducible. A norm-preserving matrix is composed of orthonormalized row (or column) vectors, the components of which are elements of $\mathbb{F}'$. Therefore norm preserving means orthogonal if $\mathbb{F}' = \mathbb{R}$ and unitary if $\mathbb{F}' = \mathbb{C}$. For $\mathbb{F}' = \mathbb{Q}$, where the term hyperunitary has been introduced, care must be taken of the noncommutativity of the multiplication. The term "absolutely irreducible" means that it is impossible to find a norm-preserving matrix T such that the matrices $TD_{\mathbb{F}'}(x)T^{-1}$, $x \in G$, all decompose into a direct sum of smaller matrices, impossible even if the elements of T are taken from an extrension field $\mathbb{F}''$ ($\subseteq \mathbb{F}'$).

(ii) If $\mathbb{F} \subset \mathbb{F}'$ the matrices $D_{\mathbb{F}'}(x)$ coincide with the matrices considered in (i). This implies, for instance, that a comlex representation ($\mathbb{F}' = \mathbb{C}$) of $\mathbb{R}$ type ($\mathbb{F} = \mathbb{R}$) has to be real.

(iii) If $\mathbb{F} \supset \mathbb{F}'$ the matrices $D_{\mathbb{F}'}(x)$ are obtained from the matrices $D_{\mathbb{F}}(x)$ considered in (i) by replacing the matrix elements (belonging to $\mathbb{F}$) by small square matrices (with elements from $\mathbb{F}'$). This is done according to one of the following conventions (cf. Ref. 2):

$$a + ib \longleftrightarrow R^{\mathbb{C}}[a + ib] = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \tag{2.1}$$

$$a + ib + jc + kd \longleftrightarrow R^{\mathbb{Q}}[a + ib + jc + kd]$$

$$= \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}, \tag{2.2}$$

$$a + ib + jc + kd \longleftrightarrow C^{\mathbb{Q}}[a + ib + jc + kd]$$

$$= \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}, \tag{2.3}$$

 0022-2488/83/010001-07$02.50

Here $a,b,c,d\in\mathbb{R}$, and $i\in\mathbb{C}$ and $i,j,k,\in\mathbb{Q}$ are the familiar units of these fields. The matrices obtained by one of the substitutions (2.1–2.3) are also norm preserving and irreducible over $\mathbb{F}'$ (but not over $\mathbb{F}$).

Summarizing the above definitions one sees that t reps of $\mathbb{R}$ type are always real; those of $\mathbb{C}$ type are complex ($\mathbb{F}=\mathbb{C}$ or $\mathbb{Q}$) or real matrices composed of submatrices of the form (2.1) ($\mathbb{F}'=\mathbb{R}$); and the t reps of $\mathbb{Q}$ type are quaternionic ($\mathbb{F}'=\mathbb{Q}$) or composed of the complex 2×2 matrices (2.3) ($\mathbb{F}'=\mathbb{C}$) or of the real 4×4 matrices (2.2) ($\mathbb{F}'=\mathbb{R}$). To make the notation more concise we write $R(x)$ instead of $D_{\mathbb{R}}(x)$, etc. The upper index Λ labeling the equivalence class of the t rep $D^{\Lambda}(x)$ is chosen to be A (or α) if $D(x)$ is of $\mathbb{R}$ type, B (or β) if $D(x)$ is of $\mathbb{C}$ type, and Γ (or γ) if it is of $\mathbb{Q}$ type. The same labels can be used for all three representation fields; only if $\mathbb{F}'=\mathbb{C}$ must the label B be replaced by the pair B +, B −. In this case the t reps are assumed to satisfy

$$[C^{\mathsf{B}+}(x)]^{*}=C^{\mathsf{B}-}(x)\quad\text{for all } x\in G. \tag{2.4}$$

Note that these representations are equivalent over $\mathbb{Q}$. We choose

$$Q^{\mathsf{B}}(x)=C^{\mathsf{B}+}(x). \tag{2.5}$$

To each representation $D(x)$ (not necessarily irreducible) a character $\chi(x)$ can be assigned. This is done by the definitions

$$\chi_{\mathbb{R}}(x)=\text{trace } R(x), \tag{2.6}$$

$$\chi_{\mathbb{C}}(x)=\text{trace } C(x), \tag{2.7}$$

$$\chi_{\mathbb{Q}}(x)=\text{real part of trace } Q(x). \tag{2.8}$$

(The real part of $a+ib+jc+kd$ is a). Taking into account the definition of t reps this implies

$$\chi^{\mathsf{A}}_{\mathbb{R}}=\chi^{\mathsf{A}}_{\mathbb{C}}=\chi^{\mathsf{A}}_{\mathbb{Q}}, \tag{2.9}$$

$$\chi^{\mathsf{B}}_{\mathbb{R}}=\chi^{\mathsf{B}+}_{\mathbb{C}}+\chi^{\mathsf{B}-}_{\mathbb{C}}=2\chi^{\mathsf{B}}_{\mathbb{Q}}, \tag{2.10}$$

$$\chi^{\Gamma}_{\mathbb{R}}=2\chi^{\Gamma}_{\mathbb{C}}=4\chi^{\Gamma}_{\mathbb{Q}}. \tag{2.11}$$

The characters form a basis for the set of square-integrable class functions defined on G. However to be complete also for $\mathbb{F}'\neq\mathbb{C}$ this basis has to be supplemented by functions ψ^{B} defined by

$$\psi^{\mathsf{B}}_{\mathbb{R}}(x)=2\sum_{J}R^{\mathsf{B}}_{J1,J0}(x)=-i\chi^{\mathsf{B}+}_{\mathbb{C}}(x)+i\chi^{\mathsf{B}-}_{\mathbb{C}}$$

$$=\sum_{J}[-iQ^{\mathsf{B}}_{JJ}(x)+Q^{\mathsf{B}}_{JJ}(x)^{*}i]=2\psi^{\mathsf{B}}_{\mathbb{Q}}(x). \tag{2.12}$$

This set is then closed under convolution if the convolution $f*g$ of two functions $f,g\in L^2(G)$ is defined by

$$f*g(y)=M_x f(x)g(x^{-1}y), \tag{2.13}$$

where $M_x h(x)$ is the normalized Haar integral of h.

Because of Eqs. (2.9–2.11) the quaternionic characters coincide with the real ones up to constant factors.

$$\chi^{\Lambda}_{\mathbb{R}}=\rho^{\Lambda}\chi^{\Lambda}_{\mathbb{Q}},\quad \rho^{\mathsf{A}}=1,\quad \rho^{\mathsf{B}}=2,\quad \rho^{\Gamma}=4. \tag{2.14}$$

The real characters satisfy

$$\chi^{\Lambda}_{\mathbb{R}}{}^{*}\chi^{\Lambda'}_{\mathbb{R}}=\delta_{\Lambda\Lambda'}\chi^{\Lambda}_{\mathbb{R}}(m^{\Lambda})^{-1}, \tag{2.15}$$

$$\rho^{\Lambda}m^{\Lambda}=\dim R^{\Lambda}(x)=\chi^{\Lambda}_{\mathbb{R}}(e). \tag{2.16}$$

Moreover

$$\chi^{\Lambda}_{\mathbb{R}}{}^{*}\psi^{\mathsf{B}}_{\mathbb{R}}=\delta_{\Lambda\mathsf{B}}\psi^{\mathsf{B}}_{\mathbb{R}}(m^{\mathsf{B}})^{-1},$$

$$\psi^{\mathsf{B}}_{\mathbb{R}}{}^{*}\psi^{\mathsf{B}}_{\mathbb{R}}=\delta_{\mathsf{B}\mathsf{B}'}\chi^{\mathsf{B}}_{\mathbb{R}}(m^{\mathsf{B}})^{-1}. \tag{2.17}$$

These equations imply the orthogonality relations

$$\langle\chi^{\Lambda}_{\mathbb{R}},\chi^{\Lambda'}_{\mathbb{R}}\rangle=\delta_{\Lambda\Lambda'}\rho^{\Lambda},\ \langle\chi^{\Lambda}_{\mathbb{R}},\psi^{\mathsf{B}}_{\mathbb{R}}\rangle=0,$$

$$\langle\psi^{\mathsf{B}}_{\mathbb{R}},\psi^{\mathsf{B}'}_{\mathbb{R}}\rangle=2\delta_{\mathsf{B}\mathsf{B}'}, \tag{2.18}$$

because the χs and the ψs are continuous functions, $\psi^{\mathsf{B}}(e)=0$, and

$$\langle f,g\rangle=\langle g,f\rangle=f*g(e)\quad\text{for } f,g\in L^2(G,\mathbb{R}). \tag{2.19}$$

The characters can be used to determine the type of a given irreducible representation $D^{\Lambda}_{\mathbb{F}}$ since the real numbers $t^{\Lambda}_{\mathbb{F}}$ defined by

$$M_x\chi^{\Lambda}_{\mathbb{F}}(x^2)=t^{\Lambda}_{\mathbb{F}}, \tag{2.20}$$

are positive for $\Lambda=\mathsf{A}$, zero for $\Lambda=\mathsf{B}$, and negative for $\Lambda=\Gamma$. We therefore arrive at the following classification scheme:

type	kind	χ test	labels	
$\mathbb{R}$	1st	+	A, α	
$\mathbb{C}$	3rd	0	B, β	(2.21)
$\mathbb{Q}$	2nd	−	Γ, γ	

3. CONSTRUCTION OF t REPS

It has been shown in Ref. 3 how to obtain t reps from a given set of complex unitary representations. There it also has been pointed out that real t reps can be found constructing a suitable basis, in the following called t basis, of the real group algebra $\mathbf{A}_{\mathbb{R}}(G)$. Here the elements $\mathbf{a}\in\mathbf{A}_{\mathbb{R}}(G)$ are defined as real linear combinations of group elements; more precisely

$$\mathbf{a}=M_x a(x)\mathbf{x},\ a\in L^2(G,\mathbb{R}), \tag{3.1}$$

where $x\to\mathbf{x}$ is the regular representation of G (cf. Ref. 2, Sec. 2). The goal is to find elements $\mathbf{e}^{\Lambda}_{JK}\in\mathbf{A}_{\mathbb{R}}(G)$ satisfying

$$(\mathbf{e}^{\Lambda}_{JK})^{\dagger}=\mathbf{e}^{\Lambda}_{KJ}, \tag{3.2}$$

$$\mathbf{e}_{JK}\mathbf{e}_{J'K'}=\delta_{\Lambda\Lambda'}\mathbf{e}^{\Lambda}_{JK'}, \tag{3.3}$$

and elements $\mathbf{f}^{\Lambda}_{S}\in\mathbf{A}_{\mathbb{R}}(G)$ behaving like the units of the field which determines the types of the representation. That is, if $\mathbf{e}^{\Lambda}$ is defined by

$$\mathbf{e}^{\Lambda}=\sum_{J}\mathbf{e}^{\Lambda}_{JJ}=(\mathbf{e}^{\Lambda})^{\dagger}, \tag{3.4}$$

then

$$\Lambda=\mathsf{A}:\ \mathbf{f}^{\mathsf{A}}_{0}=\mathbf{e}^{\mathsf{A}}; \tag{3.5}$$

$$\Lambda=\mathsf{B}:\ \mathbf{f}^{\mathsf{B}}_{0}=\mathbf{e}^{\mathsf{B}},\ \mathbf{f}^{\mathsf{B}}_{1}=\mathbf{i}^{\mathsf{B}},$$
$$(\mathbf{i}^{\mathsf{B}})^2=-\mathbf{e}^{\mathsf{B}},\ (\mathbf{i}^{\mathsf{B}})^{\dagger}=-\mathbf{i}^{\mathsf{B}},\ \mathbf{e}^{\mathsf{B}}\mathbf{i}^{\mathsf{B}}=\mathbf{i}^{\mathsf{B}}\mathbf{e}^{\mathsf{B}}=\mathbf{i}^{\mathsf{B}}; \tag{3.6}$$

$$\Lambda=\Gamma:\mathbf{f}^{\Gamma}_{0}=\mathbf{e}^{\Gamma},\ \mathbf{f}^{\Gamma}_{1}=\mathbf{i}^{\Gamma},\ \mathbf{f}^{\Gamma}_{2}=\mathbf{j}^{\Gamma},\ \mathbf{f}^{\Gamma}_{3}=\mathbf{k}^{\Gamma},$$
$$(\mathbf{i}^{\Gamma})^2=-\mathbf{e}^{\Gamma},\ (\mathbf{i}^{\Gamma})^{\dagger}=-\mathbf{i}^{\Gamma},\ \mathbf{e}^{\Gamma}\mathbf{i}^{\Gamma}=\mathbf{i}^{\Gamma}\mathbf{e}^{\Gamma}=\mathbf{i}^{\Gamma},$$
$$\mathbf{i}^{\Gamma}\mathbf{j}^{\Gamma}=-\mathbf{j}^{\Gamma}\mathbf{i}^{\Gamma}=\mathbf{k}^{\Gamma},\text{ and cyclic permutations of } i,j,k. \tag{3.7}$$

The e's and the f's are related by

$$\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda'}_{S}=\mathbf{f}^{\Lambda'}_{S}\mathbf{e}^{\Lambda}_{JK}$$
$$=0\quad\text{for } \Lambda\neq\Lambda'. \tag{3.8}$$

If the group G is finite the determination of a t basis $\{\mathbf{e}^{\Lambda}_{JK}, \mathbf{f}^{\Lambda}_{S}\}$ is a purely algebraic problem. In this case the elements $\mathbf{e}^{\Lambda}_{JJ}$ can be calculated computing and factoring (over **R**!) the minimal polynomials of a sufficiently large number of self-adjoint elements of $\mathbf{A}_{\mathbf{R}}(G)$. The finer and finer decomposition of self-adjoint idempotents ends up with idempotents $\mathbf{e}'$ ($= \mathbf{e}^{\Lambda}_{JJ}$) for which the only sef-adjoint elements of the form $\mathbf{e}'\mathbf{a}\mathbf{e}'$, $\mathbf{a}\in\mathbf{A}_{\mathbf{R}}(G)$, are the real multiples of $\mathbf{e}'$ (primitive idempotents). Once the $\mathbf{e}^{\Lambda}_{JJ}$'s are known the elements $\mathbf{e}^{\Lambda}_{JK}$, $J\neq K$, are determined from the elements $\mathbf{e}^{\Lambda}_{JJ}\mathbf{a}\mathbf{e}^{\Lambda}_{KK}\neq\mathbf{O}$ (for details see, e.g., Ref. 4). Finally the **f**'s are found by studying the structure of the division algebra consisting of all elements $\mathbf{e}^{\Lambda}_{00}\mathbf{a}\mathbf{e}^{\Lambda}_{00}$, $\mathbf{a}\in\mathbf{A}_{\mathbf{R}}(G)$. This algebra contains units $\mathbf{f}^{\Lambda}_{00,S}=\mathbf{e}^{\Lambda}_{00}\mathbf{a}_{S}\mathbf{e}^{\Lambda}_{00}$ which satisfy equations similar to one of the sets (3.5-7) but with $\mathbf{e}^{\Lambda}$ replaced by $\mathbf{e}^{\Lambda}_{00}$. From these elements the fs are obtained by

$$\mathbf{f}^{\Lambda}_{S}=\sum_{J}\mathbf{e}^{\Lambda}_{J0}\mathbf{f}^{\Lambda}_{00,S^{e}}\mathbf{e}^{\Lambda}_{0J}.$$

The same methods work for a compact continuous group G once the elements $\mathbf{e}^{\Lambda}$ are known. However, to determine these elements, forming a countable set in this case, nonalgebraic methods have to be used (functional anlalysis, differential equations, etc.). This remark also applies to the completeness relation.

$$\mathbf{a}\in\mathbf{A}_{\mathbf{R}}(G):\mathbf{a}=\sum_{\Lambda JKS}\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda}_{S}R^{\Lambda}_{JS,KO}(\mathbf{a}),$$

$$R^{\Lambda}_{JS,KO}(\mathbf{a})=(m^{\Lambda}\rho^{\Lambda})^{-1}\langle\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda}_{S},\mathbf{a}\rangle\in\mathbf{R}. \tag{3.9}$$

$$\mathbf{a},\mathbf{b}\in\mathbf{A}_{\mathbf{R}}(G):\langle\mathbf{a},\mathbf{b}\rangle=\langle\mathbf{b},\mathbf{a}\rangle=M_{x}a(x)b(x). \tag{3.10}$$

If the **e**'s and the **f**'s are known the corresponding real t reps,

$$R^{\Lambda}(\mathbf{x})=R^{\Lambda}(x), \tag{3.11}$$

can be obtained either from

$$\mathbf{x}\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda}_{S}=\sum_{J'S'}\mathbf{e}^{\Lambda}_{J'K}\mathbf{f}^{\Lambda}_{S'}R^{\Lambda}_{J'S',JS}(x) \tag{3.12}$$

or from

$$\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda}_{S}=\rho^{\Lambda}m^{\Lambda}M_{x}R^{\Lambda}_{JS,K0}(x)\mathbf{x} \tag{3.13}$$

and

$$R^{\mathsf{B}}_{JS,KT}(x)=R^{\mathsf{C}}_{ST}[C^{\mathsf{B}+}_{JK}(x)], \tag{3.14}$$

$$R^{\Gamma}_{JS,KT}(x)=R^{\mathsf{Q}}_{ST}[Q^{\Gamma}_{JK}(x)]. \tag{3.15}$$

Combined with (2.1–2.3) the last two equations do not only show the peculiar structure of the real t reps R^{B} and R^{Γ} but also indicate how to obtain the nonreal t reps from the real ones.

4. SEMIDIRECT PRODUCTS

As in the familiar complex case it may be convenient to construct the irreducible representations of a large group by means of irreducible representations of smaller groups (e.g., subgroups) which are easier to construct or even already known. The groups considered in this section are semidirect products

$$\mathsf{G}=G\circledS g;\quad G=\{E,X,Y,\ldots\}\ \text{compact},$$

$$g=\{e,x,y,\ldots\}\ \text{finite}. \tag{4.1}$$

We focus on the construction of a t basis of $\mathbf{A}_{\mathbf{R}}(\mathsf{G})$ which, as has been pointed out before, contains all relevant information and is of more interest in applications than the corresponding matrix representation. The applications we have in mind are real t reps of symmorphic space groups which are of interest in some problems of solid state theory (e.g., lattice dynamics, see Ref. 3). In the following discussion G is neither assumed to be abelian nor finite but a complete set of real t reps of G or, equivalently, a t basis of $\mathbf{A}_{\mathbf{R}}(G)$ is assumed to be known. The construction, described under Secs. 4A–4F, parallels the construction of complex unitary irreducible representations of semidirect products.[5,6]

A. Equivalence of t reps of G

The t reps of G may be partitioned into disjoint classes according to the equivalence relation

$$\Lambda\sim\Lambda'\ \text{iff}\ \mathbf{x}\mathbf{e}^{\Lambda}\mathbf{x}^{-1}=\mathbf{e}^{\Lambda'}\ \text{for some}\ x\in g. \tag{4.2}$$

It should be noted that the finite-dimensional two-sided ideals

$$\mathbf{A}^{\Lambda}_{\mathbf{R}}(G)=\mathbf{e}^{\Lambda}\mathbf{A}_{\mathbf{R}}(G)=\mathbf{A}_{R}(G)\mathbf{e}^{\Lambda} \tag{4.3}$$

and $\mathbf{A}^{\Lambda'}_{\mathbf{R}}(G)$ are isomorphic if $\Lambda\sim\Lambda'$. Accordingly if x intertwines Λ and Λ' as indicated in (4.2) and $\{\mathbf{e}^{\Lambda}_{JK},\mathbf{f}^{\Lambda}_{S}\}$ is a t basis of $\mathbf{A}^{\Lambda}_{\mathbf{R}}(G)$ so is $\{\mathbf{x}\mathbf{e}^{\Lambda}_{JK}\mathbf{x}^{-1},\mathbf{x}\mathbf{f}^{\Lambda}_{S}\mathbf{x}^{-1}\}$ of $\mathbf{A}^{\Lambda'}_{\mathbf{R}}(G)$. The first task in constructing a t basis of $\mathbf{A}_{\mathbf{R}}(\mathrm{G})$ is to determine the equivalence classes $\{\Lambda\}$ of t reps of G and to fix a set of representatives $\Lambda\in\{\Lambda\}$, one for each class.

B. Little cogroups

For each representative Λ the little cogroup g^{Λ} is the subgroup of g defined by

$$x\in g^{\Lambda}\ \text{if}\ \mathbf{x}\mathbf{c}^{\Lambda}\mathbf{x}^{-1}=\mathbf{c}^{\Lambda}\ \text{for all central elements of}\ \mathbf{A}^{\Lambda}_{\mathbf{R}}(G). \tag{4.4}$$

If $\Lambda=\mathsf{A}$ or Γ the central elements are all real multiples of $\mathbf{e}^{\Lambda}$. If however $\Lambda=\mathsf{B}$ there are two linearly independent central elements, viz. $\mathbf{e}^{\mathsf{B}}$ and $\mathbf{i}^{\mathsf{B}}$ [cf. Eqs. (3.3–3.8)]. In this case a second group, denoted by $\bar{g}^{\mathsf{B}}$ is defined by

$$x\in\bar{g}^{\mathsf{B}}\ \text{iff}\ \mathbf{x}\mathbf{e}^{\mathsf{B}}\mathbf{x}^{-1}=\mathbf{e}^{\mathsf{B}}\ \text{and}\ \mathbf{x}\mathbf{i}^{\mathsf{B}}\mathbf{x}^{-1}=\pm\mathbf{i}^{\mathsf{B}}. \tag{4.5}$$

If there does not exist an element $x\in g$ which transforms $\mathbf{i}^{\mathsf{B}}$ into $-\mathbf{i}^{\mathsf{B}}$ then $g^{\mathsf{B}}=\bar{g}^{\mathsf{B}}$. If such an element, say $\bar{x}$, exists then $\bar{g}^{\mathsf{B}}=\{g^{\mathsf{B}},\bar{x}g^{\mathsf{B}}\}=\{g^{\mathsf{B}},g^{\mathsf{B}}\bar{x}^{-1}\}$, i.e., g^{B} is a normal subgroup of $\bar{g}^{\mathsf{B}}$ of index 2. The group g^{B} is then called the proper little cogroup and $\bar{g}^{\mathsf{B}}$ is called the full little cogroup. We note in passing that the (proper) little cogroups are the same groups as encountered in the construction of complex representations.

C. Coverings of little cogroups

To each $x\in g^{\Lambda}$ an inner automorphism of the ideal $\mathbf{A}^{\Lambda}_{\mathbf{R}}(G)$ is assigned through the mapping $\mathbf{a}^{\Lambda}\rightarrow\mathbf{x}\mathbf{a}^{\Lambda}\mathbf{x}^{-1}$, transforming the t basis $\{\mathbf{e}^{\Lambda}_{JK},\mathbf{f}^{\Lambda}_{S}\}$ of $\mathbf{A}^{\Lambda}_{\mathbf{R}}(G)$ into the equivalent basis $\{\mathbf{x}\mathbf{e}^{\Lambda}_{JK}\mathbf{x}^{-1},\mathbf{x}\mathbf{f}^{\Lambda}_{S}\mathbf{x}^{-1}\}$. Since the automorphism is inner (cf. Ref. 2, Sec. 4) it is possible to find for each $x\in g^{\Lambda}$ and element $\mathbf{u}^{\Lambda}(x)\in\mathbf{A}^{\Lambda}_{\mathbf{R}}(G)$ such that

$$\mathbf{u}^{\Lambda}(x)^{\dagger}\mathbf{u}^{\Lambda}(x)=\mathbf{u}^{\Lambda}(x)\mathbf{u}^{\Lambda}(x)^{\dagger}=\mathbf{e}^{\Lambda}, \tag{4.6}$$

$$\mathbf{x}\mathbf{e}^{\Lambda}_{JK}\mathbf{x}^{-1}=\mathbf{u}^{\Lambda}(x)^{\dagger}\mathbf{e}^{\Lambda}_{JK}\mathbf{u}^{\Lambda}(x). \tag{4.7}$$

The explicit construction of an element $\mathbf{u}$ corresponding to an inner automorphism (here given by x) has been described in detail in Ref. 2, Secs. 4 and 6B. This construction yields $\mathbf{u}^{\Lambda}(e) = \mathbf{e}^{\Lambda}$ and

$$\mathbf{u}^{\Lambda}(x) = \mathbf{e}^{\Lambda} \quad \text{for all } x \in g^{\Lambda} \text{ if } \mathbf{e}^{\Lambda}_{KK} = \mathbf{e}^{\Lambda}. \tag{4.8}$$

The last condition is always satisfied for abelian G's. If G is not abelian it is sufficient to construct the elements $\mathbf{u}^{\Lambda}(x)$ for the generators $x,y,..$ and to put $\mathbf{u}^{\Lambda}(z) = \mathbf{u}^{\Lambda}(x)^m \mathbf{u}^{\Lambda}(y)^n \cdots$ if $z = x^m y^n \cdots$. Now if $x \in g^{\Lambda}$, $\mathbf{u}^{\Lambda}(x) \in \mathbf{A}^{\Lambda}_{\mathbb{R}}(G)$ is an element satisfying (4.6,7), and

$$\mathbf{F}^{\Lambda}_{\mathbb{R}} = \left\{ \sum_S \mathbf{f}^{\Lambda}_S a_S \,\middle|\, a_S \in \mathbb{R} \right\}, \tag{4.9}$$

then $\mathbf{u}^{\Lambda}(x)\mathbf{x}\mathbf{f}\mathbf{x}^{-1}\mathbf{u}^{\Lambda}(x)^{\dagger} \in \mathbf{F}^{\Lambda}_{\mathbb{R}}$ iff $\mathbf{f} \in \mathbf{F}^{\Lambda}_{\mathbb{R}}$. Moreover, since the mapping $\mathbf{f} \to \mathbf{u}^{\Lambda}(x)\mathbf{x}\mathbf{f}\mathbf{x}^{-1}\mathbf{u}^{\Lambda}(x)^{\dagger}$ is invertible it is automorphism of the division algebra $\mathbf{F}^{\Lambda}_{\mathbb{R}}$: for $\mathbf{F}^{\mathsf{A}}_{\mathbb{R}} \simeq \mathbb{R}$ and $\mathbf{F}^{\mathsf{B}}_{\mathbb{R}} \simeq \mathbb{C}$ it is the identical mapping ($\mathbf{i}^{\mathsf{B}}$ is invariant under g^{B}!), and for $\mathbf{F}^{\Lambda}_{\mathbb{R}} \simeq \mathbb{Q}$ it is an inner automorphism since all automorphisms of $\mathbb{Q}$ are inner. In this case there exists a unimodular quaternion $q \in \mathbb{Q}$ inducing this automorphism and therefore an element $\mathbf{q}^{\Gamma}(x) \in \mathbf{F}^{\Gamma}_{\mathbb{R}}$ such that

$$\mathbf{q}^{\Gamma}(x)^{\dagger}\mathbf{q}^{\Gamma}(x) = \mathbf{q}^{\Gamma}(x)\mathbf{q}^{\Gamma}(x)^{\dagger} = \mathbf{e}^{\Gamma}, \tag{4.10}$$

$$\mathbf{q}^{\Gamma}(x)^{\dagger}\mathbf{e}^{\Gamma}_{JK}\mathbf{q}^{\Gamma}(x) = \mathbf{e}^{\Gamma}_{JK}, \tag{4.11}$$

$$\mathbf{u}^{\Gamma}(x)\mathbf{x}\mathbf{f}^{\Gamma}_S\mathbf{x}^{-1}\mathbf{u}^{\Gamma}(x) = \mathbf{q}^{\Gamma}(x)^{\dagger}\mathbf{f}^{\Gamma}_S\mathbf{q}^{\Gamma}(x), \tag{4.12}$$

the second equation following from (3.8). Putting

$$\begin{aligned} \mathbf{w}^{\Lambda}(x) &= \mathbf{u}^{\Lambda}(x) \quad \text{for } \Lambda = \mathsf{A},\mathsf{B}, \\ &= \mathbf{q}^{\Lambda}(x)\mathbf{u}^{\Gamma}(x) \quad \text{for } \Lambda = \Gamma, \end{aligned} \tag{4.13}$$

we see that

$$\mathbf{x}\mathbf{a}^{\Lambda}\mathbf{x}^{-1} = \mathbf{w}^{\Lambda}(x)^{\dagger}\mathbf{a}^{\Lambda}\mathbf{w}^{\Lambda}(x) \quad \text{for all } \mathbf{a}^{\Lambda} \in \mathbf{A}^{\Lambda}_{\mathbb{R}}(G) \tag{4.14}$$

and that the factors in the definition of $\mathbf{x}^{\Lambda}$,

$$\mathbf{x}^{\Lambda} = \mathbf{w}^{\Lambda}(x)\mathbf{x} = \mathbf{x}\mathbf{w}^{\Lambda}(x), \tag{4.15}$$

may be interchanged because of $\mathbf{w}^{\Lambda}(x) \in \mathbf{A}^{\Lambda}_{\mathbb{R}}(G)$.

If $x,y \in g$ and $xy = z$ then $y^{-1}x^{-1}z$ induces the identical automorphism and $\mathbf{w}^{\Lambda}(y)\mathbf{w}^{\Lambda}(x)\mathbf{w}^{\Lambda}(z)^{\dagger}$ is an element of $\mathbf{A}^{\Lambda}_{\mathbb{R}}(G)$ commuting with all other elements of this ideal.

$$xy \in g: \mathbf{w}^{\Lambda}(y)\mathbf{w}^{\Lambda}(x)\mathbf{w}^{\Lambda}(xy)^{\dagger} = \mathbf{c}^{\Lambda}(x,y) \in \text{ center of } \mathbf{A}^{\Lambda}_{\mathbb{R}}(G). \tag{4.16}$$

Now if $\Lambda = \mathsf{A}$ or $\Lambda = \Gamma$ the center consists of multiples of $\mathbf{e}^{\Lambda}$ and $\mathbf{c}^{\Lambda}$ is unimodular because of (4.6,10). Hence

$$\Lambda = \mathsf{A},\Gamma\text{: } \mathbf{c}^{\Lambda}(x,y) = \pm \mathbf{e}^{\Lambda}. \tag{4.17}$$

For $\Lambda = \mathsf{B}$ one has

$$\begin{aligned} \mathbf{c}^{\mathsf{B}}(x,y) &= \mathbf{e}^{\mathsf{B}} \cos\phi(x,y) + \mathbf{i}^{\mathsf{B}} \sin\phi(x,y) \\ &= \exp\{\mathbf{i}^{\mathsf{B}}\phi(x,y)\}, \ \phi(x,y) \in [0,2\pi), \end{aligned} \tag{4.18}$$

i.e., the factors behave like unimodular complex numbers. It seems to be unknown which phases can occur in (4.18) but it has been shown already by I. Schur how to redefine the elements $\mathbf{w}$ so that the new factors behave like nth roots of unity, n being the order of g^{B}. The new elements $\tilde{\mathbf{w}}$ are given by[7]

$$\tilde{\mathbf{w}}^{\mathsf{B}}(x) = \exp\{-\mathbf{i}^{\mathsf{B}}\psi(x)\}\mathbf{w}^{\mathsf{B}}(x), \tag{4.19}$$

where the angles ψ have to be determined from

$$\psi(x) = |g^{\mathsf{B}}|^{-1} \sum_{y \in g^{\mathsf{B}}} \phi(x,y). \tag{4.20}$$

If the elements $\tilde{\mathbf{w}}$ are used instead of $\mathbf{w}$ in (4.16,4.18) the phases are found to satisfy

$$|g^{\mathsf{B}}|\tilde{\phi}(x,y) = \text{multiple of } 2\pi. \tag{4.21}$$

Moreover

$$\tilde{\phi}(x,y) = 0 \quad \text{if } g^{\mathsf{B}} \text{ is abelian.} \tag{4.22}$$

In the following it is always assumed that the $\mathbf{w}$'s are normalized according to (4.19), if $\Lambda = \mathsf{B}$, and that these normalized $\mathbf{w}$s are used in the definition(4.15) of the elements $\mathbf{x}^{\mathsf{B}}$.

Having determined the (normalized) elements $\mathbf{w}^{\Lambda}(x)$ for all $x \in g^{\Lambda}$ it is possible to define a group $\mathbf{pg}^{\Lambda}$ by

$$\mathbf{pg}^{\Lambda} = \text{multiplicative group generated by } \mathbf{x}^{\Lambda}, \ x \in g^{\Lambda}. \tag{4.23}$$

The center of $\mathbf{pg}^{\Lambda}$ contains the group

$$\mathbf{z}^{\Lambda} = \text{multiplicative group generated by } \mathbf{c}^{\Lambda}(x,y), \ x,y \in g^{\Lambda}, \tag{4.24}$$

and $\mathbf{pg}^{\Lambda}/\mathbf{z}^{\Lambda} \simeq g^{\Lambda}$; in this sense $\mathbf{pg}^{\Lambda}$ is a covering group of $\mathbf{g}^{\Lambda}$. The symbol $\mathbf{pg}^{\Lambda}$ has been chosen to indicate that, in an equivalent terminology, this group could have been called a projective representation of g^{Λ} with factor system $\mathbf{z}^{\Lambda}$. Since this factor system is at most of order 2 for $\Lambda = \mathsf{A},\Gamma$ and at most of order $|g^{\mathsf{B}}|$ for $\Lambda = \mathsf{B}$ $\mathbf{pg}^{\Lambda}$ is always a finite group.

D. t reps of the covering groups

Let pg^{Λ} be the abstract group defined by the isomorphism $pg^{\Lambda} \simeq \mathbf{pg}^{\Lambda}$. The the next step is to construct real t reps of the groups pg^{Λ} or t bases of the algebras $\mathbf{A}_{\mathbb{R}}(pg^{\Lambda})$. How to obtain these objects has been indicated in Sec. 3: If complex unitary irreducible representations of the group are known they can be used to form real t reps from which the t-bases are obtained by Eq. (3.13). Likewise, since pg^{Λ} is finite, the algebraic methods outlined before can be used to construct the t bases directly. Moreover if pg^{Λ} is itself a semidirect product the strategy described in this section may be used to reduce the construction of its t bases to that of simpler groups (see also Secs. 5 and 6).

Not all the t reps of a group pg^{Λ} are needed but only a subset depending on the factor system $\mathbf{z}^{\Lambda}$. For let $R^{\lambda}(\tilde{x})$ be a real t rep of pg^{Λ} and $\tilde{x}^{\Lambda}$ be the image of $\tilde{x}$ in $\mathbf{pg}^{\Lambda}$; then the elements

$$\mathbf{e}^{\lambda}_{jk}\mathbf{f}^{\lambda}_s = \rho^{\lambda} m^{\lambda} M_{\tilde{x}} R^{\lambda}_{js,k0}(\tilde{x})\tilde{\mathbf{x}}^{\Lambda} \tag{4.25}$$

may vanish for $pg^{\Lambda} \neq g^{\Lambda}$ since then the elements $\tilde{\mathbf{x}}^{\Lambda}$ are not linearly independent. All these linear dependences may be traced back to those of elements of $\mathbf{z}^{\Lambda}$, i.e., to relations of the form

$$\sum_{\tilde{\mathbf{x}} \in \mathbf{z}^{\Lambda}} \tilde{\mathbf{x}} a(\tilde{x}) = \mathbf{O}, \quad a(\tilde{x}) \in \mathbb{R}. \tag{4.26}$$

A representation λ of pg^{Λ} is of interest if, and only if, all Eqs. (4.26) with $R^{\lambda}(\tilde{x})$ substituted for $\tilde{\mathbf{x}}$ are identically satisfied. In terms of projective representations this means that only

those irreducible representations are allowed which have the same factor system as the original representations.

E. t reps of the little groups

For the semidirect product groups considered here the little groups may be defined as the semi-direct products $\mathsf{G}^{\Lambda} = G \circledS pg^{\Lambda}$. For a fixed Λ and an allowable λ Eqs. (3.2–3.8), (4.25), and (4.14 and 4.15) imply that the real linear combinations of the elements

$$\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda}_{S}\mathbf{e}^{\lambda}_{jk}\mathbf{f}^{\lambda}_{s} = \mathbf{e}^{\lambda}_{jk}\mathbf{f}^{\lambda}_{s}\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda}_{S} = \mathbf{e}^{\Lambda}_{JK}\mathbf{e}^{\lambda}_{jk}\mathbf{f}^{\Lambda}_{S}\mathbf{f}^{\lambda}_{s}, \tag{4.27}$$

form a two-sided ideal in $\mathbf{A}_{\mathbb{R}}(\mathsf{G}^{\Lambda})$. This ideal does not necessarily correspond to one single t rep of G^{Λ}, and even if it does, the elements (4.27) need not constitute a t basis. An obvious exception is the case where one of the two representations involved is of $\mathbb{R}$ type, e.g., $\Lambda = \mathsf{A}$. In this case the elements

$$\mathbf{e}^{\mathsf{A}\lambda}_{Jj,Kk} = \mathbf{e}^{\mathsf{A}}_{JK}\mathbf{e}^{\lambda}_{jk}, \quad \mathbf{f}^{\mathsf{A}\lambda}_{s} = \mathbf{f}^{\lambda}_{s}, \tag{4.28}$$

form a t basis and

$$\mathsf{A}\lambda \text{ is of the same type as } \lambda. \tag{4.29}$$

In all other cases the passage from (4.27) to t bases of minimal two-sided ideals involves only linear transformations of the elements $\mathbf{f}^{\Lambda}_{s}\mathbf{f}^{\lambda}_{s}$. If, for instance, $\Lambda\lambda = \mathsf{B}\beta$ one introduces the idempotents

$$\mathbf{e}^{[\mathsf{B}\beta]\kappa} = \tfrac{1}{2}[\mathbf{e}^{\mathsf{B}}\mathbf{e}^{\beta} + (-1)^{\kappa}\mathbf{i}^{\mathsf{B}}\mathbf{i}^{\beta}], \quad \kappa = 0,1, \tag{4.30}$$

and defines the units $\mathbf{i}^{\mathsf{B}\beta\kappa}$ by

$$\mathbf{i}^{\mathsf{B}\beta\kappa} = \mathbf{i}^{\mathsf{B}}\mathbf{e}^{[\mathsf{B}\beta]\kappa} = -(-1)^{\kappa}\mathbf{i}^{\beta}\mathbf{e}^{[\mathsf{B}\beta]\kappa}. \tag{4.31}$$

It is easily verified that these elements combined with

$$\mathbf{e}^{\mathsf{B}\beta\kappa}_{Jj,Kk} = \mathbf{e}^{\mathsf{B}}_{JK}\mathbf{e}^{\beta}_{jk}\mathbf{e}^{[\mathsf{B}\beta]\kappa}, \tag{4.32}$$

constitute t bases of the two inequivalent t reps $\mathsf{B}\beta\,0$ and $\mathsf{B}\beta\,1$, and that

$$\mathsf{B}\beta\kappa \text{ is of } \mathbb{C} \text{ type.} \tag{4.33}$$

Next consider $\Lambda\lambda = \mathsf{B}\gamma$. Choosing

$$\mathbf{e}^{\mathsf{B}\gamma}_{JjL,Kkl} = \mathbf{e}^{\mathsf{B}}_{JK}\mathbf{e}^{\gamma}_{jk}\mathbf{e}^{[\mathsf{B}\gamma]}_{Ll}, \tag{4.34}$$

$$\begin{aligned}
&\mathbf{e}^{[\mathsf{B}\gamma]}_{00} = \tfrac{1}{2}[\mathbf{e}^{\mathsf{B}}\mathbf{e}^{\gamma} + \mathbf{i}^{\mathsf{B}}\mathbf{i}^{\gamma}],\\
&\mathbf{e}^{[\mathsf{B}\gamma]}_{11} = \tfrac{1}{2}[\mathbf{e}^{\mathsf{B}}\mathbf{e}^{\gamma} - \mathbf{i}^{\mathsf{B}}\mathbf{i}^{\gamma}],\\
&\mathbf{e}^{[\mathsf{B}\gamma]}_{01} = \mathbf{e}^{[\mathsf{B}\gamma]}_{00}\mathbf{j}^{\gamma} = \mathbf{j}^{\gamma}\mathbf{e}^{[\mathsf{B}\gamma]}_{11},\\
&\mathbf{e}^{[\mathsf{B}\gamma]}_{10} = -\mathbf{j}^{\gamma}\mathbf{e}^{[\mathsf{B}\gamma]}_{00} = -\mathbf{e}^{[\mathsf{B}\gamma]}_{11}\mathbf{j}^{\gamma},
\end{aligned} \tag{4.35}$$

$$\mathbf{i}^{\mathsf{B}\gamma} = \mathbf{i}^{\mathsf{B}}, \tag{4.36}$$

one sees that the elements (4.34), (4.36) form a t basis and that

$$\mathsf{B}\gamma \text{ is of } \mathbb{C} \text{ type.} \tag{4.37}$$

Finally, if $\Lambda\lambda = \Gamma\gamma$, it is possible to define elements $\mathbf{e}^{[\Gamma\gamma]}_{Ll}$,

$$\begin{aligned}
&\mathbf{e}^{[\Gamma\gamma]}_{00} = \tfrac{1}{4}[\mathbf{e}^{\Gamma}\mathbf{e}^{\gamma} + \mathbf{i}^{\Gamma}\mathbf{i}^{\gamma} + \mathbf{j}^{\Gamma}\mathbf{j}^{\gamma} + \mathbf{k}^{\Gamma}\mathbf{k}^{\gamma}],\\
&\mathbf{e}^{[\Gamma\gamma]}_{01} = \mathbf{e}^{[\Gamma\gamma]}_{00}\mathbf{i}^{\gamma},\ \mathbf{e}^{[\Gamma\gamma]}_{02} = \mathbf{e}^{[\Gamma\gamma]}_{00}\mathbf{j}^{\gamma},\ \mathbf{e}^{[\Gamma\gamma]}_{03} = \mathbf{e}^{[\Gamma\gamma]}_{00}\mathbf{k}^{\gamma},\\
&\mathbf{e}^{[\Gamma\gamma]}_{12} = -\mathbf{i}^{\gamma}\mathbf{e}^{[\Gamma\gamma]}_{00}\mathbf{j}^{\gamma},\ \mathbf{e}^{[\Gamma\gamma]}_{13}\\
&\qquad = -\mathbf{i}^{\gamma}\mathbf{e}^{[\Gamma\gamma]}_{00}\mathbf{k}^{\gamma},\ \mathbf{e}^{[\Gamma\gamma]}_{23} = -\mathbf{j}^{\gamma}\mathbf{e}^{[\Gamma\gamma]}_{00}\mathbf{k}^{\gamma},\\
&\mathbf{e}^{[\Gamma\gamma]\dagger}_{Ll} = \mathbf{e}^{[\Gamma\gamma]}_{lL},\ \mathbf{e}^{[\Gamma\gamma]}_{LL} = \mathbf{e}^{[\Gamma\gamma]}_{L0}\mathbf{e}^{[\Gamma\gamma]}_{0L},
\end{aligned} \tag{4.38}$$

which can be used to define the t basis

$$\mathbf{e}^{\Gamma\gamma}_{JjL,Kkl} = \mathbf{e}^{\Gamma}_{JK}\mathbf{e}^{\gamma}_{jj}\mathbf{e}^{[\Gamma\gamma]}_{Ll}. \tag{4.39}$$

Its structure shows that

$$\Gamma\gamma \text{ is of } \mathbb{R} \text{ type.} \tag{4.40}$$

This exhausts all possibilities since the remaining three cases ($\Lambda\lambda = \mathsf{B}\alpha, \Gamma\alpha, \Gamma\beta$) are obtained from (4.27) and (4.34–4.36) simply by interchanging the roles of Λ and λ.

We note in passing that the t reps corresponding to the t bases (4.31,4.32), (4.34,4.36), and (4.39) can also be obtained from the tensor products of the t reps R^{Λ} and R^{λ}, i.e., from the matrices with elements

$$[R^{\Lambda}(X) \otimes R^{\lambda}(\tilde{x})]_{JSjs,KTkt} = R^{\Lambda}_{JS,KT}(X)R^{\lambda}_{js,kt}(\tilde{x}), \tag{4.41}$$

by orthogonal transformations involving only the indices Ss,Tt. For $\Lambda\lambda = \mathsf{B}\beta$ this reads

$$\begin{aligned}
&\delta_{\kappa\kappa'}R^{\mathsf{B}\beta\kappa}_{Jjs,Kkt}(X\tilde{x})\\
&\quad = \sum_{S'\bar{s}'T't'} R^{[\mathsf{B}\beta]}_{S'\bar{s}',\kappa s}R^{\mathsf{B}}_{JS',KT'}(X)R^{\beta}_{js',kt'}(\tilde{x})R^{[\mathsf{B}\beta]}_{T't',\kappa' t},
\end{aligned} \tag{4.42}$$

where $R^{[\mathsf{B}\beta]}$ is the matrix

$$R^{[\mathsf{B}\beta]} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 1 & 0\\ -1 & 0 & 0 & 1\\ 1 & 0 & 0 & 1\\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

Rows: $Tt = 00,10,01,11.$

Columns: $\kappa s = 00,01;10,11.$ (4.43)

Similar results hold for $\lambda = \gamma$, where

$$\begin{aligned}
&\delta_{mm'}R^{\Lambda\gamma}_{JjLS,Kkls}(X\tilde{x})\\
&\quad = \sum_{S'\bar{s}'T't'} R^{[\Lambda\gamma]}_{S'\bar{s}',mLS}R^{\Lambda}_{JS',KT'}(X)R^{\gamma}_{js',kt'}(\tilde{x})R^{[\Lambda\gamma]}_{T't',m'ls},
\end{aligned} \tag{4.44}$$

and m is an index labeling the identical copies of $R^{\Lambda\gamma}$ appearing in the decomposition of $R^{\Lambda} \otimes R^{\gamma}$.

$$\Lambda = \mathsf{B}\text{: } m = 0,1. \tag{4.45}$$

$$\Lambda = \Gamma\text{: } m = 0,1,2,3. \tag{4.46}$$

The transformation matrices are given by

$$R^{[\mathsf{B}\gamma]} = \frac{1}{\sqrt{2}}\begin{pmatrix}
0 & 0 & 0 & -1 & 0 & 1 & 0 & 0\\
0 & 0 & 1 & 0 & -1 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0\\
0 & -1 & 0 & 0 & 0 & 0 & 0 & -1\\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
-1 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$

Rows: $Tt = 00,10,01,11,02,12,03,12.$

Columns: $mls = 000,001,010,011;100,101,110,111.$

$$R^{[\Gamma\gamma]} = \tfrac{1}{2}\begin{pmatrix} E & I & J & K\\ -I & E & K & -J\\ -J & -K & E & I\\ -K & J & -I & E \end{pmatrix}, \tag{4.47}$$

$E = \bar{R}^{\mathbb{Q}}[1]$, $I = \bar{R}^{\mathbb{Q}}[i]$, $J = \bar{R}^{\mathbb{Q}}[j]$, $K = \bar{R}^{\mathbb{Q}}[k]$, $\bar{R}^{\mathbb{Q}}[q]$, see Ref. 2, Eq. (2.30).

Rows: $Tt = 00,10,20,30,01,..,31,02,..32,03,..,33.$

Columns: $ml = 00,01,02,03;10,..,13;20,..,23;30,..,33.$ (4.48)

Some more manipulations are needed if $\Lambda = \mathsf{B}$ and $\bar{g}^{\mathsf{B}} \neq g^{\mathsf{B}}$. Since g^{B} is a normal subgroup of $\bar{g}^{\mathsf{B}} = \{g^{\mathsf{B}}, \bar{x}g^{\mathsf{B}}\} = \{g^{\mathsf{B}}, g^{\mathsf{B}}\bar{x}^{-1}\}$ and the elements of this group map central elements of $\mathbf{A}^{\mathsf{B}}_{\mathbb{R}}(G)$ onto central elements the covering of the proper little cogroup, $\mathbf{pg}^{\mathsf{B}}$, can be extended to a covering of the full little cogroup, namely $\mathbf{p}\bar{\mathbf{g}}^{\mathsf{B}} = \{\mathbf{pg}^{\mathsf{B}}, \bar{\mathbf{x}}\mathbf{pg}^{\mathsf{B}}\} = \{\mathbf{pg}^{\mathsf{B}}, \mathbf{pg}^{\mathsf{B}}\bar{\mathbf{x}}^{-1}\}$. Therefore it is possible to define in an obvious way a proper and a full little group in this case. The (allowable) t reps of $\bar{\mathsf{G}}^{\mathsf{B}}$ may be obtained from those of G^{B} but their explict construction depends on whether the quantity τ defined by

$$\mathbf{e}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}\mathbf{e}^{\mathsf{B}\lambda(\kappa)}\bar{x}^{-1} = \tau^2\mathbf{e}^{\mathsf{B}\lambda(\kappa)}, \tag{4.49}$$

vanishes or not. If it does a t basis of $\bar{\mathsf{G}}^{\mathsf{B}}$ is given by the following expressions:

$\tau = 0$:

$$\mathbf{e}^{\mathsf{B}\lambda(\kappa)0}_{Jj(L)N,Kk(l)n} = \bar{\mathbf{x}}^N\mathbf{e}^{\mathsf{B}\lambda(\kappa)}_{Jj(L),Kk(l)}\bar{\mathbf{x}}^{-n}, \quad N,n = 0,1, \tag{4.50}$$

$$\mathbf{i}^{\mathsf{B}\lambda(\kappa)0} = \mathbf{i}^{\mathsf{B}\lambda(\kappa)} + \bar{\mathbf{x}}\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}^{-1}. \tag{4.51}$$

$\mathsf{B}\lambda(\kappa)0$ is of $\mathbb{C}$ type. (4.52)

If $\tau \neq 0$ the situation is slightly more complicated. The first thing to be observed in this case is that the mapping $\mathbf{a} \to \bar{\mathbf{x}}\mathbf{a}\bar{\mathbf{x}}^{-1}$, defined for all $\mathbf{a} \in \mathbf{A}^{\mathsf{B}\lambda(\kappa)}_{\mathbb{R}}(\mathsf{G}^{\mathsf{B}})$, is an automorphism of this ideal. This automorphism is an outer one because of (4.28) with $\mathsf{A} \to \alpha$, $\lambda \to \mathsf{B}$, (4.31), (4.36), and $\bar{\mathbf{x}}\mathbf{i}^{\mathsf{B}}\bar{\mathbf{x}}^{-1} = -\mathbf{i}^{\mathsf{B}}$. Therefore

$$\bar{\mathbf{x}}\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}^{-1} = -\mathbf{i}^{\mathsf{B}\lambda(\kappa)}. \tag{4.53}$$

There exists an element $\bar{\mathbf{u}}^{\mathsf{B}\lambda(\kappa)} \in \mathbf{A}^{\mathsf{B}\lambda(\kappa)}_{\mathbb{R}}(\mathsf{G}^{\mathsf{B}})$ (constructed by the same method as the elements $\mathbf{u}^{\Lambda}(x)$, $x \in g^{\Lambda}$, before) such that

$$\bar{\mathbf{u}}^{\mathsf{B}\lambda(\kappa)\dagger}\mathbf{e}^{\mathsf{B}\lambda(\kappa)}_{Jj(L),Kk(l)}\bar{\mathbf{u}}^{\mathsf{B}\lambda(\kappa)} = \bar{\mathbf{x}}\mathbf{e}^{\mathsf{B}\lambda(\kappa)}_{Jj(L),Kk(l)}\bar{\mathbf{x}}^{-1},$$
$$\bar{\mathbf{u}}^{\mathsf{B}\lambda(\kappa)\dagger}\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{u}}^{\mathsf{B}\lambda(\kappa)} = \mathbf{i}^{\mathsf{B}\lambda(\kappa)}. \tag{4.54}$$

Accordingly defining

$$\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)} = \bar{\mathbf{u}}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}, \tag{4.55}$$

one has

$$\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)\dagger} = \bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)\dagger}\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)} = \mathbf{e}^{\mathsf{B}\lambda(\kappa)}, \tag{4.56}$$

$$\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)}\mathbf{e}^{\mathsf{B}\lambda(\kappa)}_{Jj(L),Kk(l)}\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)\dagger} = \mathbf{e}^{\mathsf{B}\lambda(\kappa)}_{Jj(L),Kk(l)}, \tag{4.57}$$

$$\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)}\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)\dagger} = -\mathbf{i}^{\mathsf{B}\lambda(\kappa)}, \tag{4.58}$$

and

$$[\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)}]^2 = \exp\{\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\theta\}, \quad \theta \in [0,2\pi), \tag{4.59}$$

the last equation following from the fact that (4.49) is an element of $\mathbf{A}^{\mathsf{B}\lambda(\kappa)}_{\mathbb{R}}(\mathsf{G}^{\mathsf{B}})$ and induces the identical mapping. Combining (4.58) and (4.59) one finds $\exp\{+\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\theta\} = \exp\{-\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\theta\}$; therefore $\theta = 0$ or π and

$$\bar{\mathbf{x}}^{|\mathsf{B}\lambda(\kappa)|2} = \tau\mathbf{e}^{\mathsf{B}\lambda(\kappa)}, \quad \tau = \pm 1. \tag{4.60}$$

Taking into account Eqs. (4.56–4.58,4.60) the corresponding t bases of $\bar{\mathsf{G}}^{\mathsf{B}}$ are easily constructed.

$\tau = 1$:

$$\mathbf{e}^{\mathsf{B}\lambda(\kappa)(-1)}_{Jj(L),Kk(l)} = \mathbf{e}^{\mathsf{B}\lambda(\kappa)}_{Jj(L),Kk(l)}; \tag{4.61}$$

$$\mathbf{i}^{\mathsf{B}\lambda(\kappa)(-1)} = \mathbf{i}^{\mathsf{B}\lambda(\kappa)}, \quad \mathbf{j}^{\mathsf{B}\lambda(\kappa)(-1)} = \bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)},$$
$$\mathbf{k}^{\mathsf{B}\lambda(\kappa)(-1)} = \mathbf{i}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)}; \tag{4.62}$$

$\mathsf{B}\lambda(\kappa)(-1)$ is of $\mathbb{Q}$ type. (4.63)

$\tau = +1$:

$$\mathbf{e}^{\mathsf{B}\lambda(\kappa)(+1)}_{Jj(L)N,Kk(l)n} = \mathbf{e}^{\mathsf{B}\lambda(\kappa)}_{Jj(L),Kk(l)}\mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{Nn}; \tag{4.64}$$
$$\mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{00} = (1/2)[\mathbf{e}^{\mathsf{B}\lambda(\kappa)} + \mathbf{i}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)}],$$
$$\mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{11} = (1/2)[\mathbf{e}^{\mathsf{B}\lambda(\kappa)} - \mathbf{i}^{\mathsf{B}\lambda(\kappa)}\bar{\mathbf{x}}^{\mathsf{B}\lambda(\kappa)}],$$
$$\mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{10} = \mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{00}\mathbf{i}^{\mathsf{B}\lambda(\kappa)} = -\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{11},$$
$$\mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{01} = -\mathbf{i}^{\mathsf{B}\lambda(\kappa)}\mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{00} = \mathbf{e}^{[\mathsf{B}\lambda(\kappa)]}_{11}\mathbf{i}^{\mathsf{B}\lambda(\kappa)}; \tag{4.65}$$

$\mathsf{B}\lambda(\kappa)(+1)$ is of $\mathbb{R}$ type. (4.66)

F. t reps of the semidirect product group

The final step is to induce the t reps of $\mathsf{G} = G \circledS g$ from the t reps of the (full) little group. To this end g has to be decomposed into cosets with respect to g^{Λ} (or $\bar{g}^{\mathsf{B}}$, respectively) and a fixed set of coset representatives has to be chosen First assume that

$\Lambda \neq \mathsf{B}$ or $\Lambda = \mathsf{B}$, $\bar{g}^{\mathsf{B}} = g^{\mathsf{B}}$:

$$g = \{y^{(0)}g^{\Lambda}, y^{(1)}g^{\Lambda}, \ldots\}, \quad y^{(0)} = e. \tag{4.67}$$

A t basis of G is then given by

$$\mathbf{e}^{\{\Lambda\lambda(\kappa)\}}_{Jj(L)P,Kk(l)p} = \mathbf{y}^{(P)}\mathbf{e}^{\Lambda\lambda(\kappa)}_{Jj(L),Kk(l)}\mathbf{y}^{(p)-1}, \tag{4.68}$$

$$\mathbf{f}^{\{\Lambda\lambda(\kappa)\}}_{s} = \sum_{p}\mathbf{y}^{(p)}\mathbf{f}^{\Lambda\lambda(\kappa)}_{s}\mathbf{y}^{(p)-1}, \tag{4.69}$$

and

$$\{\Lambda\lambda(\kappa)\} \text{ is of the same type as } \Lambda\lambda(\kappa). \tag{4.70}$$

The remaining cases are handled in an analogous manner.

$\Lambda = \mathsf{B}$, $\bar{g}^{\mathsf{B}} \neq g^{\mathsf{B}}$:

$$g = \{y^{(0)}\bar{g}^{\mathsf{B}}, y^{(1)}\bar{g}^{\mathsf{B}}, \ldots\}, \quad y^{(0)} = e. \tag{4.71}$$

$$\mathbf{e}^{\{\mathsf{B}\lambda(\kappa)\tau\}}_{Jj(L)(N)P,Kk(l)(n)p} = \mathbf{y}^{(P)}\mathbf{e}^{\mathsf{B}\lambda(\kappa)\tau}_{Jj(L)(N),Kk(l)(n)}\mathbf{y}^{(p)-1}, \tag{4.72}$$

$$\mathbf{f}^{\{\mathsf{B}\lambda(\kappa)\tau\}}_{s} = \sum_{p}\mathbf{y}^{(p)}\mathbf{f}^{\mathsf{B}\lambda(\kappa)\tau}_{s}\mathbf{y}^{(p)-1}. \tag{4.73}$$

The type of $\{\mathsf{B}\lambda(\kappa)\tau\}$ is given by $\tau\,(=0,\pm 1)$. (4.74)

This concludes the construction of t reps of semidirect products. That the t bases obtained this way satisfy Eq. (3.2) is obvious from their construction ($\mathbf{x}^{\dagger} = \mathbf{x}^{-1}$). That (3.3) holds true follows from

$$\Lambda \neq \mathsf{B}\text{: } \mathbf{a}^{\lambda}\mathbf{x}\mathbf{b}^{\Lambda} = \mathbf{O} \quad \text{if} \quad x \notin g^{\Lambda}; \tag{4.75}$$

$$\Lambda = \mathsf{B}\text{: } \mathbf{a}^{\mathsf{B}}\mathbf{x}\mathbf{b}^{\mathsf{B}} = \mathbf{O} \quad \text{if} \quad x \notin \bar{g}^{\mathsf{B}}; \tag{4.75}$$

these equations also show that the fs of the t bases of G behave like the fs of the t bases of the little groups G^{Λ} (or $\bar{\mathsf{G}}^{\mathsf{B}}$) which explains propositions (4.70,4.74). That this methodis exhaustive may be shown by successively proving the completeness of the bases $\{\mathbf{e}^{\Lambda}_{JK}\mathbf{f}^{\Lambda}_{S}\mathbf{x} \mid \text{all reps } \Lambda, \text{ all } x \in g\}$ and $\{\mathbf{y}'\mathbf{e}^{\Lambda\lambda(\kappa)(\tau)}_{Jj(L)(N),Kk(l)(n)}\mathbf{f}^{\Lambda\lambda(\kappa)(\tau)}_{s}\mathbf{y}^{-1} \mid \text{all representatives } \Lambda, \text{ all allowable representations } \lambda(\kappa)(\tau), \text{ all coset representatives } y,y'\}$.

5. DIRECT PRODUCTS

The above considerations are greatly simplified if G is a direct product of the form

$$\mathrm{G} = G \times g, \quad G,g \text{ compact.} \tag{5.1}$$

In this case $Xx = xX$ for all $X \in G$, $x \in g$, $\{\Lambda\} = \Lambda$, and $g^{\Lambda} = pg^{\Lambda} = g$; therefore the construction outlined above reduces to Sec. 4E once the t reps of G and g are known. If these representations are denoted by Λ and λ one finds [cf. Eqs. (4.29,4.33,4.37)]

$$\begin{aligned} &A\alpha, \Gamma\gamma \text{ are of } \mathbb{R} \text{ type;} \\ &A\beta, B\alpha, B\beta\kappa, B\gamma, \Gamma\beta \text{ are of } \mathbb{C} \text{ type;} \\ &A\gamma, \Gamma\alpha \text{ are of } \mathbb{Q} \text{ type.} \end{aligned} \tag{5.2}$$

This is in agreement with the character test (2.20) because the character of the representation $R^{\Lambda} \otimes R^{\lambda}$ containing the t rep $R^{\Lambda\lambda(\kappa)}$ is $\chi_{\mathbf{R}}^{\Lambda \otimes \lambda}(Xx) = \chi_{\mathbf{R}}^{\Lambda}(X)\chi_{\mathbf{R}}^{\lambda}(x)$.

6. INDUCTION FROM NORMAL SUBGROUPS

It is also possible to generalize the methods of Sec. 4 to include the construction of t reps by induction from t reps of a normal subgroup. Let G, G, and g be related by

$$\begin{aligned} &G = \text{compact normal subgroup of G,} \\ &\mathrm{G}/G \simeq g \text{ (finite),} \end{aligned} \tag{6.1}$$

and let "g" be a fixed set of coset representatives with respect to G.

$$\mathrm{G} = \{Ge, Gx, Gy, \cdots\}, \quad \text{"}g\text{"} = \{e,x,y,\cdots\}. \tag{6.2}$$

The product of two coset representatives is then given by

$$yx = Z[x,y]z[x,y], \quad Z[x,y] \in G, \ z[x,y] \in \text{"}g\text{"}, \tag{6.3}$$

where the elements Z and z are uniquely determined by x and y. Assume furthermore that the scheme of Sec. 4 has been followed up to Eq. (4.15) but with g replaced by "g" everywhere, and that the elements $\mathbf{w}^{\Lambda}(x) \in \mathbf{A}_{\mathbf{R}}^{\Lambda}(G)$ satisfying (4.14) have been constructed for all $x \in$ "g". Then

$$\begin{aligned} &\mathbf{x}^{\Lambda}\mathbf{y}^{\Lambda} = \mathbf{c}^{\Lambda}(x,y)\mathbf{z}^{\Lambda}, \quad \mathbf{c}^{\Lambda}(x,y) = \mathbf{w}^{\Lambda}(y)\mathbf{w}^{\Lambda}(x)\mathbf{Z}\mathbf{w}^{\Lambda}(z)^{\dagger}, \\ &Z = Z[x,y], \quad z = z[x,y], \end{aligned} \tag{6.4}$$

because of (4.14) and (6.3). But

$$\begin{aligned} &\mathbf{y}^{-1}\mathbf{x}^{-1}\mathbf{Z}[x,y]\mathbf{z}[x,y]\mathbf{a}^{\Lambda}\mathbf{z}[x,y]^{-1}\mathbf{Z}[x,y]^{-1}\mathbf{x}\mathbf{y} \\ &\quad = \mathbf{c}^{\Lambda}(x,y)\mathbf{a}^{\Lambda}\mathbf{c}^{\Lambda}(x,y)^{\dagger} = \mathbf{a}^{\Lambda} \quad \text{for all } \mathbf{a}^{\Lambda} \in \mathbf{A}_{\mathbf{R}}^{\Lambda}(G), \end{aligned} \tag{6.5}$$

which again implies

$$\mathbf{c}^{\Lambda}(x,y) \in \text{center of } \mathbf{A}_{\mathbf{R}}^{\Lambda}(G). \tag{6.6}$$

Hence the conclusions following (4.16) are again valid and the further scheme of Sec. 4 can be applied as it stands. Combined with the content of this section this method shows, among other things, how to construct t reps of both symmorphic and nonsymmorphic space groups.

[1] F. J. Dyson, J. Math. Phys. **6**, 1199 (1962). This article not only covers most of the results of Ref. 3 (and some of Ref. 2) but is also more general since it contains the algebraic classification of corepresentations and their commutants for arbitrary magnetic groups.

[2] P. Kasperkovitz, J. Math. Phys. **22**, 2417 (1981).

[3] P. Kasperkovitz and G. Kahl, J. Math. Phys. **22**, 2404 (1981).

[4] R. Dirl and P. Kasperkovitz, *Gruppentheorie* (Vieweg, Braunschweig, 1977), pp. 33–37, 56–62.

[5] L. Jansen and M. Boon, *Theory of Finite Groups* (North-Holland, Amsterdam, 1967) pp. 153–161.

[6] R. Dirl, J. Math. Phys. **18**, 2065 (1977).

[7] C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras* (Wiley, New York, 1962), pp. 358–365.

Type-adapted subduction matrices

P. Kasperkovitz
Institut für Theoretische Physik, Technische Universität, A-1040 Wien, Karlsplatz 13, Austria

(Received 7 August 1981; accepted for publication 16 September 1981)

If an irreducible representation is restricted to a subgroup it becomes reducible in general. The matrices transforming this reducible representation into a direct sum of irreducible constituents are called subduction matrices. Their structure is discussed for real, complex, and quaternionic representations where all these representations are assumed to show a peculiar structure characteristic for the type of this representation (character test $+, 0, -$). The choice of these type-adapted representations, a convention possible for all compact groups, considerably reduces the number of parameters needed to fix a subduction matrix.

PACS numbers: 02.20. + b

1. INTRODUCTION

In the preceding paper[1] (referred to as I) it has been shown how to obtain type-adapted representations (t reps) of a group from (projective) t reps of a normal subgroup. Here we consider the inverse problem: How does a given t rep Λ of a group G decompose into t reps λ of a subgroup g (which need not be normal) if Λ is restricted to g? The essential result of this paper is that the matrix which transforms the (reducible) representation $D^{\Lambda}\downarrow g$ into a direct sum of t reps D^{λ} can be put into a form which is adapted to both Λ and all the λ's contained in this representation. That is, this so-called subduction matrix can always be chosen to reflect the internal algebraic structure (i.e., the type) of the representations involved. Similarly to I, the problem is essentially solved if it is solved for real t reps because the complex and quaternionic subduction matrices are obtained from the real ones according to some simple algebraic rules (substitutions, transformations) depending only on the types of the representations involved. It turns out that this method reduces the number of real parameters needed to fix a complex subduction matrix. As a byproduct we also find that the multiplicities of the representations λ contained in a representation Λ have to be even in some instances or even multiples of four, if noncomplex representations are considered.

If $G = g \times g$ (direct product) the results of this paper can be combined with those of I to find convenient Clebsch Gordan matrices. For complex representations this problem has been discussed extensively by R. Dirl.[2] Although the reasoning differs our results agree where they overlap. But apart from the different setting (noncomplex representations versus corepresentations, general subductions versus Clebsch Gordan series) the present approach, at least in the author's opinion, makes the principle from which these results emerge more transparent. It is simply the fact that every irreducible representation of a compact group is absolutely irreducible over one of the three fields: the reals, the complex numbers, or the quaternions.

In the following the notation of I is used and (I.n.m) means Eq. (n.m) of I. For both G and g a complete set of t reps labeled by Λ and λ, respectively, is assumed to be given.

2. MULTIPLICITIES

The multiplicities $m_{\mathbb{F}}^{\Lambda\lambda}$ are the nonnegative integers appearing in the decomposition

$$S_{\mathbb{F}}^{\Lambda(g)\dagger} D_{\mathbb{F}}^{\Lambda}(x) S_{\mathbb{F}}^{\Lambda(g)} = \sum \oplus m_{\mathbb{F}}^{\Lambda\lambda} D_{\mathbb{F}}^{\lambda}(x);$$

$$D_{\mathbb{F}}^{\Lambda} = t \text{ rep of } G, \quad D_{\mathbb{F}}^{\lambda} = t \text{ rep of } g; \quad G \supset g. \tag{2.1}$$

Every $m_{\mathbb{F}}^{\Lambda\lambda}$ is uniquely determined by the characters $\chi_{\mathbb{F}}^{\Lambda}$ and $\chi_{\mathbb{F}}^{\lambda}$.

$\mathbb{F} \neq \mathbb{C}$:

$$m_{\mathbb{F}}^{\Lambda\lambda} = M_x \chi_{\mathbb{F}}^{\Lambda}(x) \chi_{\mathbb{F}}^{\lambda}(x) / M_x \chi_{\mathbb{F}}^{\lambda}(x) \chi_{\mathbb{F}}^{\lambda}(x), \tag{2.2}$$

$\mathbb{F} = \mathbb{C}$:

$$m_{\mathbb{C}}^{\Lambda\lambda} = M_x \chi_{\mathbb{C}}^{\Lambda}(x) * \chi_{\mathbb{C}}^{\lambda}(x) = M_x \chi_{\mathbb{C}}^{\Lambda}(x) \chi_{\mathbb{C}}^{\lambda}(x)^*. \tag{2.3}$$

Equations (I.2.4), (I.2.6), and (I.2.3) imply trivial identities like

$$m_{\mathbb{C}}^{A\beta+} = m_{\mathbb{C}}^{A\beta-}, \quad m_{\mathbb{C}}^{B+\beta+} = m_{\mathbb{C}}^{B-\beta-}, \text{ etc.} \tag{2.4}$$

Taking into account

$$2\chi_{\mathbb{C}}^{B\pm} = \chi_{\mathbb{R}}^{B} \pm i\psi_{\mathbb{R}}^{B}, \tag{2.5}$$

and using Eqs. (I.2.9–11), (I.2.18), one finds the following interrelations of multiplicities, some of which are less obvious.

$$\begin{aligned}
&m_{\mathbb{R}}^{A\alpha} = m_{\mathbb{C}}^{A\alpha} = m_{\mathbb{Q}}^{A\alpha}, \\
&m_{\mathbb{R}}^{A\beta} = m_{\mathbb{C}}^{A\beta\pm} = \tfrac{1}{2} m_{\mathbb{Q}}^{A\beta}, \\
&m_{\mathbb{R}}^{A\gamma} = \tfrac{1}{2} m_{\mathbb{C}}^{A\gamma} = \tfrac{1}{4} m_{\mathbb{Q}}^{A\gamma}, \\
&m_{\mathbb{R}}^{B\alpha} = 2m_{\mathbb{C}}^{B\pm\alpha} = 2m_{\mathbb{Q}}^{B\alpha}, \\
&m_{\mathbb{R}}^{B\beta} = m_{\mathbb{C}}^{B\pm\beta\pm} + m_{\mathbb{C}}^{B\pm\beta\mp} = m_{\mathbb{Q}}^{B\beta}, \\
&m_{\mathbb{R}}^{B\gamma} = m_{\mathbb{C}}^{B\pm\gamma} = \tfrac{1}{2} m_{\mathbb{Q}}^{B\gamma},
\end{aligned} \tag{2.6}$$

$$\begin{aligned}
&m_{\mathbb{R}}^{\Gamma\alpha} = 2m_{\mathbb{C}}^{\Gamma\alpha} = 4m_{\mathbb{Q}}^{\Gamma\alpha}, \\
&m_{\mathbb{R}}^{\Gamma\beta} = 2m_{\mathbb{C}}^{\Gamma\beta\pm} = 2m_{\mathbb{Q}}^{\Gamma\beta}, \\
&m_{\mathbb{R}}^{\Gamma\gamma} = m_{\mathbb{C}}^{\Gamma\gamma} = m_{\mathbb{Q}}^{\Gamma\gamma}.
\end{aligned} \tag{2.7}$$

(2.8)

3. REAL SUBDUCTION MATRICES

The matrices appearing in Eq. (2.1) for $\mathbb{F} = \mathbb{R}$,

$$S_{\mathbb{R}}^{\Lambda(g)} = R^{\Lambda(g)}, \tag{3.1}$$

can be constructed combining the so-called projection technique[3–5] with a generalized Schmidt process.[6,7] The basic idea of the projection technique is to write (2.1) in the form

$$R^{\Lambda}(x) R^{\Lambda(g)} = R^{\Lambda(g)} \left[\sum_{\lambda} \oplus m_{\mathbb{R}}^{\Lambda\lambda} R^{\lambda}(x) \right], \tag{3.2}$$

 0022-2488/83/010008-06$02.50

and to consider the columns of $R^{\Lambda(g)}$ as (othonormalized) vectors, $R^{\Lambda}(x)$ as an operator acting onto these vectors, and $[\Sigma \oplus m_{\mathbf{R}}^{\Lambda\lambda} R^{\lambda}(x)]$ as a matrix representation of the operators $R^{\Lambda}(x)$. This alone would not be too advantageous but forming the appropriate linear combinations the real vector space of dimension $\rho^{\lambda} m^{\lambda}$ spanned by the columns of $R^{\Lambda(g)}$ can be shown to carry not only a representation of the group g but also if its real group algebra $\mathbf{A}_{\mathbf{R}}(g)$.

$$R^{\Lambda}(\mathbf{a}) = M_x a(x) R^{\Lambda}(x), \tag{3.3}$$

$$\sum_{\lambda} \oplus m_{\mathbf{R}}^{\Lambda\lambda} R^{\lambda}(\mathbf{a}) = \sum_{\lambda} \oplus m_{\mathbf{R}}^{\Lambda\lambda} M_x a(x) R^{\lambda}(x). \tag{3.4}$$

Now the elements of a t basis have an extremely simple matrix representation, viz.

$$R^{\lambda}_{js,k0}(\mathbf{e}^{\lambda}_{j'k'}) = \delta_{\lambda\lambda'} \delta_{s0} \delta_{jj'} \delta_{kk'}, \tag{3.5}$$

$$R^{\lambda}_{js,k0}(\mathbf{f}^{\lambda}_{s'}) = \delta_{\lambda\lambda'} \delta_{jk} \delta_{ss'}, \tag{3.6}$$

the remaining matrix elements ($t \neq 0$) being related to (3.5, 3.6) by Eqs. (I.2.1) and (I.2.2). The corresponding matrices acting from the left on the column vectors are

$$E^{(\Lambda)\lambda}_{jk} = R^{\Lambda}(\mathbf{e}^{\lambda}_{jk}), \tag{3.7}$$

$$F^{(\Lambda)\lambda}_{s} = R^{\Lambda}(\mathbf{f}^{\lambda}_{s}), \quad \text{e.g.,} \quad I^{(\Lambda)\beta} = R^{\Lambda}(\mathbf{i}^{\beta}). \tag{3.8}$$

Equations (3.2)–(3.5) show that each column vector is an eigenvector of one of the projection operators $E^{(\Lambda)\lambda}_{jj}$ and that the columns belonging to a certain λ may be partitioned into $m_{\mathbf{R}}^{\Lambda\lambda}$ groups, the members of the groups being transformed into each other under the action of the shift operators $E^{(\Lambda)\lambda}_{jk}$, $j \neq k$, and $F^{(\Lambda)\lambda}_{s}$, $s \neq 0$.

The peculiar feature of the shift operators $F^{(\Lambda)\lambda}_{s}$ is that they commute with the projection operators $E^{(\Lambda)\lambda}_{jk}$ [cf. Eq. (I.3.8)]. Moreover the column vectors $F^{(\Lambda)\lambda}_{s} v$, $s = 0, \ldots, \rho^{\lambda} - 1$, obtained from a given column vector v by application of the operators $F^{(\Lambda)\lambda}_{s}$ are mutually orthogonal and of the same norm as v. These facts may be used to construct sets of orthonormalized eigenvectors of a given projection operator; more generally, if one tries to find a complete set of orthonormalized eigenvectors by a Schmidt process, this process may be varied in such a way that in each step a pair or a quadruple of vectors is obtained instead of a single one. For $\Lambda \neq \mathsf{A}$ there exist, however, even more commuting shift operators to implement this procedure. If $\Lambda = \mathsf{B}$ it is

$$I^{\mathsf{B}} = R^{\mathsf{B}}(\mathbf{i}^{\mathsf{B}}); \tag{3.9}$$

if $\Lambda = \Gamma$ it is the operators $\bar{I}^{\Gamma}, \bar{J}^{\Gamma}, \bar{K}^{\Gamma}$ defined by

$$\begin{aligned} \bar{Q}^{\Gamma}_{JS,KT} &= (\bar{E}^{\Gamma} a + \bar{I}^{\Gamma} b + \bar{J}^{\Gamma} c + \bar{K}^{\Gamma} d)_{JS,KT} \\ &= \delta_{JK} \bar{R}^{\mathsf{Q}}_{ST}[a + ib + jc + kd] = \delta_{JK} \bar{R}^{\mathsf{Q}}_{ST}[q], \end{aligned} \tag{3.10}$$

$$\bar{R}^{\mathsf{Q}}[a + ib + jc + kd] = \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}. \tag{3.11}$$

The operators (3.9, 3.10) commute with all operators $R^{\Lambda}(X)$, $X \in G$, and hence with all operators $R^{\lambda}(\mathbf{a})$, $\mathbf{a} \in \mathbf{A}_{\mathbf{R}}(g)$. The existence of these operators indicates that the subduction matrix can be adapted to the type of Λ whereas the operators (3.8) are responsible for its adaption to λ. To avoid linear dependences members of both sets have to be combined into additional projection operators, the rest being used as shift operators. How this can be done is implicitly contained in Eqs. (2.7, 2.8); it is stated explicitly in Eqs. (3.19)–(3.26) below.

Before entering into the details of the construction of $R^{\Lambda(g)}$ let us first fix the notation.

$$\dim R^{\Lambda(g)} = \dim R^{\Lambda} = \rho^{\Lambda} m^{\Lambda}; \quad \rho^{\mathsf{A}} = 1, \rho^{\mathsf{B}} = 2, \rho^{\Gamma} = 4. \tag{3.12}$$

Labeling of the elements of $R^{\Lambda(g)}$:

$$\begin{aligned} &\text{row index } JS: && J = 0, \ldots, m^{\Lambda} - 1; \\ &&& S = 0, \ldots, \rho^{\Lambda} - 1; \\ &\text{column index } jtm: && j = 0, \ldots, m^{\lambda} - 1; \\ &&& t = 0, \ldots, \tau^{\Lambda\lambda} - 1; \\ &&& m = 0, \ldots, \mu^{\Lambda\lambda} - 1. \end{aligned} \tag{3.13}$$

Now let R^{λ} be a representation of g contained in $R^{\Lambda}\downarrow g$, i.e., $m_{\mathbf{R}}^{\Lambda\lambda} \neq 0$. The rows of $R^{\Lambda(g)}$ belonging to this λ are then obtained according to the following scheme:

(i) A subspace V^{λ} of the real vector space V consisting of all columns with $\rho^{\Lambda} m^{\Lambda}$ components is characterized by a projection matrix $P^{\Lambda\lambda}$, i.e.,

$$V^{\lambda} = \{ v_0^{\lambda} | v_0^{\lambda} \in V, P^{\Lambda\lambda} v_0^{\lambda} = v_0^{\lambda} \}. \tag{3.14}$$

The matrices $P^{\Lambda\lambda}$ are given below for all combinations of types.

(ii) An orthonormalized basis of V^{λ} is constructed by means of a generalized Schmidt process. In each step a $\tau^{\Lambda\lambda}$-tuple of vectors is constructed (cf. Ref. 7, Sec. 6A). The members of a set are related by

$$v_{0t}^{\lambda} = S_t^{\Lambda\lambda} v_{00}^{\lambda}, \quad t = 0, \ldots, \tau^{\Lambda\lambda} - 1, \tag{3.15}$$

the matrices $S_t^{\Lambda\lambda}$ being specified below, and there exist $\mu^{\Lambda\lambda}$ such sets. One ends up with

$$\tau^{\Lambda\lambda} \mu^{\Lambda\lambda} = \rho^{\lambda} m_{\mathbf{R}}^{\Lambda\lambda}, \tag{3.16}$$

orthonormalized column vectors $v_{0t}^{\lambda m}$ ($t = 0, \ldots, \tau^{\Lambda\lambda} - 1$; $m = 0, \ldots, \mu^{\Lambda\lambda} - 1$).

(iii) The rest of the columns belonging to λ is obtained applying the matrices $E^{(\Lambda)\lambda}_{j0}$, Eq. (3.7), onto the vectors $v_{0t}^{\lambda m}$ constructed in (ii), i.e.,

$$v_{jt}^{\lambda m} = E^{(\Lambda)\lambda}_{j0} v_{0t}^{\lambda m}, \quad j = 0, \ldots, m^{\lambda} - 1. \tag{3.17}$$

The matrix $R^{\Lambda(g)}$ is obtained by constructing successively the columns belonging to the different λ's (with $m_{\mathbf{R}}^{\Lambda\lambda} \neq 0$) and collecting them into a square matrix.

The matrices $P^{\Lambda\lambda}$ and $S_t^{\Lambda\lambda}$ needed for the explicit construction are given in the following equations:

$\Lambda\lambda = \mathsf{A}\alpha$:

$$\begin{aligned} &\text{(i)} \quad P^{\mathsf{A}\alpha} = E^{(\mathsf{A})\alpha}_{00}. \\ &\text{(ii)} \quad \tau^{\mathsf{A}\alpha} = 1\text{: } S_0^{\mathsf{A}\alpha} = E^{(\mathsf{A})\alpha}. \end{aligned} \tag{3.18}$$

$\Lambda\lambda = \mathsf{A}\beta$:

$$\begin{aligned} &\text{(i)} \quad P^{\mathsf{A}\beta} = E^{(\mathsf{A})\beta}_{00}. \\ &\text{(ii)} \quad \tau^{\mathsf{A}\beta} = 2\text{: } S_0^{\mathsf{A}\beta} = E^{(\mathsf{A})\beta}, S_1^{\mathsf{A}\beta} = I^{(\mathsf{A})\beta}. \end{aligned} \tag{3.19}$$

$\Lambda\lambda = \mathsf{A}\gamma$:

(i) $P^{\mathsf{A}\gamma} = E_{00}^{(\mathsf{A})\gamma}$.

(ii) $\tau^{\mathsf{A}\gamma} = 4$: $S_0^{\mathsf{A}\gamma} = E^{(\mathsf{A})\gamma}, S_1^{\mathsf{A}\gamma} = I^{(\mathsf{A})\gamma}$, $S_2^{\mathsf{A}\gamma} = J^{(\mathsf{A})\gamma}, S_3^{\mathsf{A}\gamma} = K^{(\mathsf{A})\gamma}$. (3.20)

$\Lambda\lambda = \mathsf{B}\alpha$:

(i) $P^{\mathsf{B}\alpha} = E_{00}^{(\mathsf{B})\alpha}$.

(ii) $\tau^{\mathsf{B}\alpha} = 2$: $S_0^{\mathsf{B}\alpha} = E^{\mathsf{B}}, S_1^{\mathsf{B}\alpha} = I^{\mathsf{B}}$. (3.21)

$\Lambda\lambda = \mathsf{B}\beta$:

(i) first $P^{\mathsf{B}\beta} = P^{\mathsf{B}\beta(+)}$, then $P^{\mathsf{B}\beta} = P^{\mathsf{B}\beta(-)}$, where $P^{\mathsf{B}\beta(\pm)} = \frac{1}{2}[E^{\mathsf{B}} \mp I^{\mathsf{B}} I^{(\mathsf{B})\beta}] E_{00}^{(\mathsf{B})\beta}$. (3.22)

(ii) $\tau^{\mathsf{B}\beta} = 2$: $S_0^{\mathsf{B}\beta} = E^{(\mathsf{B})\beta}, S_1^{\mathsf{B}\beta} = I^{(\mathsf{B})\beta}$.

$\Lambda\lambda = \mathsf{B}\gamma$:

(i) $P^{\mathsf{B}\gamma} = \frac{1}{2}[E^{\mathsf{B}} - I^{\mathsf{B}} I^{(\mathsf{B})\gamma}] E_{00}^{(\mathsf{B})\gamma}$.

(ii) $\tau^{\mathsf{B}\gamma} = 4$: $S_0^{\mathsf{B}\gamma} = E^{(\mathsf{B})\gamma}, S_1^{\mathsf{B}\gamma} = I^{(\mathsf{B})\gamma}$, $S_2^{\mathsf{B}\gamma} = J^{(\mathsf{B})\gamma}, S_3^{\mathsf{B}\gamma} = K^{(\mathsf{B})\gamma}$. (3.23)

$\Lambda\lambda = \Gamma\alpha$:

(i) $P^{\Gamma\alpha} = E_{00}^{(\Gamma)\alpha}$.

(ii) $\tau^{\Gamma\alpha} = 4$: $S_0^{\Gamma\alpha} = \bar{E}^{\Gamma}, S_1^{\Gamma\alpha} = -\bar{I}^{\Gamma}$, $S_2^{\Gamma\alpha} = -\bar{J}^{\Gamma}, S_3^{\Gamma\alpha} = -\bar{K}^{\Gamma}$. (3.24)

$\Lambda\lambda = \Gamma\beta$:

(i) $P^{\Gamma\beta} = \frac{1}{2}[E^{\Gamma} + \bar{I}^{\Gamma} I^{(\Gamma)\beta}] E_{00}^{(\Gamma)\beta}$.

(ii) $\tau^{\Gamma\beta} = 4$: $S_0^{\Gamma\beta} = \bar{E}^{\Gamma}, S_1^{\Gamma\beta} = -\bar{I}^{\Gamma}$, $S_2^{\Gamma\beta} = -\bar{J}^{\Gamma}, S_3^{\Gamma\beta} = -\bar{K}^{\Gamma}$. (3.25)

$\Lambda\lambda = \Gamma\gamma$:

(i) $P^{\Gamma\gamma} = \frac{1}{4}[E^{\Gamma} + \bar{I}^{\Gamma} I^{(\Gamma)\gamma} + \bar{J}^{\Gamma} J^{(\Gamma)\gamma} + \bar{K}^{\Gamma} K^{(\Gamma)\gamma}] E_{00}^{(\Gamma)\gamma}$.

(ii) $\tau^{\Gamma\gamma} = 4$: $S_0^{\Gamma\gamma} = \bar{\bar{E}}^{\Gamma}, S_1^{\Gamma\gamma} = -\bar{I}^{\Gamma}$, $S_2^{\Gamma\beta} = -\bar{J}^{\Gamma}, S_3^{\Gamma\beta} = -\bar{K}^{\Gamma}$. (3.26)

4. COMPLEX AND QUATERNIONIC SUBDUCTION MATRICES

Next to be shown is how the complex and quaternionic subduction matrices,

$$S_{\mathbb{C}}^{\Lambda(g)} = C^{\Lambda(g)}, \quad S_{\mathbb{Q}}^{\Lambda(g)} = Q^{\Lambda(g)}, \tag{4.1}$$

can be obtained from the real matrices $R^{\Lambda(g)}$. This step involves only substitutions or linear transformations given by simple complex or quarternionic matrices. That no more effort is needed may be traced back to the fact that all irreducible representations of a compact group can be obtained from the real ones by successively extending $\mathbb{R}$ to $\mathbb{C}$ and $\mathbb{Q}$ (see Ref. 6, Sec. 3). In the first step each irreducible representation of $\mathbb{C}$ or $\mathbb{Q}$ type decomposes into two complex irreducible representations of half the dimension which are inequivalent (complex conjugate) for $\Lambda = \mathsf{B}$ but equivalent (identical) for $\Lambda = \Gamma$. If $\mathbb{C}$ is extended to $\mathbb{Q}$ the representations $C^{\mathsf{B}+}$ and $C^{\mathsf{B}-}$ become equivalent whereas C^{Γ} splits into two copies of the quaternionic representation Q^{Γ}.

Extending the reals has two aspects since both Λ and λ may become reducible. The second effect is obviously harmless: if a representation D^{λ} becomes reducible it only has to be transformed into a direct sum of its irreducible constituents. But this is a standard procedure because all that needs to be done in this case is to diagonalize elements of the group algebra which behave like the units of the extension field; which elements are appropriate is clearly indicated by the type of λ. At first sight the reducibility of D^{Λ} seems to pose more serious problems but in this case the subduction matrix may always be put into such a form that the same transformation decomposes both D^{Λ} and $S^{\Lambda(g)}$. This is not surprising for the following reasons: Every representation of G may be restricted to a representation of g which in turn may be transformed into a direct sum of irreducible representations by a matrix of the same dimension. If the representation field of Λ is fixed and only representations of g irreducible over this field are considered the elements of the subduction matrix may be chosen from the same field. Therefore starting from the absolutely irreducible representations R^{A}, $C^{\mathsf{B}+}$, Q^{Γ} one ends up with direct sums of representations of g irreducible over $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Q}$, respectively, if the corresponding subduction matrices are real, complex, and quaternionic. If the representation field is an extension of the field from which the elements of the matrices D^{Λ} ($= R^{\mathsf{A}}$ or $C^{\mathsf{B}+}$ or Q^{Γ}), $S^{\Lambda(g)}$, and $\Sigma \oplus m^{\Lambda\lambda} D^{\lambda}$ are taken, then the reducible ones of the matrices D^{λ} must be transformed into irreducible representations as described above. If, on the other hand, the representation field is a proper subfield of the field over which Λ is absolutely irreducible, the matrices D^{Λ} ($= R^{\mathsf{A}}$ or $C^{\mathsf{B}+}$ or Q^{Γ}), $S^{\Lambda(g)}$, and $\Sigma \oplus m^{\Lambda\lambda} D^{\lambda}$ can be "blown up" by one of the substitutions (I.2.1–2.3) resulting in larger matrices of a peculiar structure. Up to minor changes, discussed below, this approach has already been taken into account in the construction of the real subduction matrice $R^{\Lambda(g)}$.

The details of the construction of the nonreal subduction matrices are best understood treating the three types of Λ separately. For $\Lambda = \mathsf{A}$ the real representation is absolutely irreducible so that only the possible reducibility of the representations λ has to be taken into account. This is done by the following definitions:

$$C^{\mathsf{A}(g)} = R^{\mathsf{A}(g)} C^{\mathsf{A}}, \quad Q^{\mathsf{A}(g)} = C^{\mathsf{A}(g)} Q^{\mathsf{A}}, \tag{4.2}$$

$$C^{\mathsf{A}}_{\lambda jtm,\lambda' j' t' m'} = \delta_{\lambda\lambda'} \cdot \delta_{jj'} \cdot \delta_{mm'} C^{\lambda}_{tt'}, \tag{4.3}$$

$$C^{\alpha} = 1,$$

$$C^{\beta} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix},$$

$$C^{\gamma} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 1 & -1 \\ -i & i & -i & -i \\ 1 & -1 & -1 & -1 \\ -i & -i & i & -i \end{pmatrix}, \tag{4.4}$$

$$Q_{\lambda jtm,\lambda' j' t' m'} = \delta_{\lambda\lambda'} \cdot \delta_{jj'} \cdot \delta_{mm'} Q^{\lambda}_{tt'}, \tag{4.5}$$

$$Q^{\alpha} = 1, \quad Q^{\beta} = \begin{pmatrix} 1 & 0 \\ 0 & j \end{pmatrix}, \quad Q^{\gamma} = Q \oplus Q,$$

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ j & k \end{pmatrix}. \tag{4.6}$$

It is easily checked that C^{β} transforms $R^{\mathbb{C}}[a + ib]$ into a direct sum of $a \pm ib$. Accordingly R^{β} decomposes into $C^{\beta+}$

$\oplus C^{\beta-}$, the two representations being interlocked since they belong to $t=0$, 1. The transformation Q^{β} carries $C^{\beta-}$ into $C^{\beta+}$ to make explicit that these two representations are equivalent over Q. Similarly the matrix C^{γ} transforms $R^{Q}[q]$ into a direct sum of two copies of $C^{Q}[q]$ which are further reduced to $q\oplus q\oplus q\oplus q$ by Q^{γ}. Note that all this is in agreement with Eqs. (2.6).

Next consider $\Lambda=\mathsf{B}$. Here the subduction matrix $R^{\mathsf{B}(g)}$ had been chosen in such a way that the irreducible representations R^{λ} appearing in the decomposition have the form indicated in Eq. (3.6). In fact Eqs. (3.22) ensure that the elements $\mathbf{e}^{\beta}a+\mathbf{i}^{\beta}b$ are always represented by matrices

$$R^{\beta}(\mathbf{e}^{\beta}a+\mathbf{i}^{\beta}b)=\sum\oplus R^{\mathrm{C}}[a+ib]. \tag{4.7}$$

Likewise conventions (3.23) had been chosen to obtain representations where

$$R^{\gamma}(\mathbf{e}^{\gamma}a+\mathbf{i}^{\gamma}b+\mathbf{j}^{\gamma}c+\mathbf{k}^{\gamma}d)=\sum\oplus R^{Q}[a+ib+jc+kd]. \tag{4.8}$$

This was necessary for real t reps but is inconvenient for complex ones because in general the matrix $R^{\mathsf{B}(g)}$ does not commute with

$$R^{\mathsf{B}}(\mathbf{i}^{\mathsf{B}})=I^{\mathsf{B}}=\sum\oplus R^{\mathrm{C}}[i]. \tag{4.9}$$

Now we are looking for a real subduction matrix $\tilde{R}^{\mathsf{B}(g)}$ commuting with I^{B},

$$\tilde{R}^{\mathsf{B}(g)}I^{\mathsf{B}}=I^{\mathsf{B}}\tilde{R}^{\mathsf{B}(g)}, \tag{4.10}$$

because all matrices with this property are composed of submatrices of the form R^{C} and hence decomposed into two complex conjugate matrices by the same transformation which transforms I^{B} into $\Sigma\oplus[(+i)\oplus(-i)]$. Such a matrix $\tilde{R}^{\mathsf{B}(g)}$ is obtained from $R^{\mathsf{B}(g)}$ by multiplication with a simple orthogonal matrix R, viz.

$$\tilde{R}^{\mathsf{B}(g)}=R^{\mathsf{B}(g)}R, \tag{4.11}$$

$$R_{\lambda jtm,\lambda' j' t' m'}=\delta_{\lambda\lambda'}\delta_{jj'}\delta_{mm'}r^{\lambda}_{tm}, \tag{4.12}$$

$r^{\alpha}_{tm}=1$,

$r^{\beta}_{tm}=\mu^{t}$, where $\mu(=\pm1)$ is determined for each m from $P^{\mathsf{B}\beta\mu}v^{\beta m}_{00}=v^{\beta m}_{00}$,

$$r^{\gamma}_{tm}=\begin{Bmatrix}+1 & \text{for} & t\neq 2\\ -1 & \text{for} & t=2\end{Bmatrix}. \tag{4.13}$$

It follows from (4.11–4.13) and (3.21–3.23) that the columns of $\tilde{R}^{\mathsf{B}(g)}$ can be combined into pairs $v, I^{\mathsf{B}}v$; this implies (4.10). It is also evident from (4.11–4.13) and (3.6) that the representations $\tilde{R}^{\lambda}$ appearing in (3.2) if $\tilde{R}^{\mathsf{B}(g)}$ is used instead of $R^{\mathsf{B}(g)}$ are no longer t reps for $\lambda=\beta(-1)$, γ but have the following form:

$$R^{\beta(\mu)}(\mathbf{e}^{\beta}a+\mathbf{i}^{\beta}b)=\sum\oplus R^{\mathrm{C}}[a+\mu ib], \tag{4.14}$$

$$R^{\gamma}(\mathbf{e}^{\gamma}a+\mathbf{i}^{\gamma}b+\mathbf{j}^{\gamma}c+\mathbf{k}^{\gamma}d)=\sum\oplus\begin{pmatrix}R^{\mathrm{C}}[a+ib] & R^{\mathrm{C}}[c+id]\\ R^{\mathrm{C}}[-c+id] & R^{\mathrm{C}}[a-ib]\end{pmatrix}. \tag{4.15}$$

The matrices R^{B}, $\tilde{R}^{\mathsf{B}(g)}$, and $\Sigma\oplus m_{\mathrm{R}}{}^{\lambda}\tilde{R}^{\lambda}$ all commute with I^{B}, which is equivalent to saying that they are composed of submatrices of the form R^{C}. The corresponding complex representations and subduction matrices are therefore easily obtained by replacing each submatrix $R^{\mathrm{C}}[a+ib]$ by the complex number $a+ib$. That is, if the columns of $\tilde{R}^{\mathsf{B}(g)}$ are relabeled according to

$$\begin{aligned}\lambda=\alpha,\beta:&\quad t=0\to u=0,\ v=0\\ &\quad t=1\to u=0,\ v=1,\\ \lambda=\gamma:&\quad t=0\to u=0,\ v=0\\ &\quad t=1\to u=0,\ v=1\\ &\quad t=2\to u=1,\ v=0\\ &\quad t=3\to u=1,\ v=1,\end{aligned} \tag{4.16}$$

then

$$\tilde{R}^{\mathsf{B}(g)}_{JR,\lambda juvm}=R^{\mathrm{C}}_{Rv}[C^{\mathsf{B}+(g)}_{J,\lambda jum}]. \tag{4.17}$$

This is the subduction matrix for the representation $C^{\mathsf{B}+}\downarrow g$, where $C^{\mathsf{B}+}$ is given by

$$R^{\mathsf{B}}_{JR,J'R'}(x)=R^{\mathrm{C}}_{RR'}[C^{\mathsf{B}+}_{JJ'}(x)]. \tag{4.18}$$

The subduction matrix for the representation $C^{\mathsf{B}-}$ is

$$C^{\mathsf{B}-(g)}=[C^{\mathsf{B}+(g)}]^{*}C^{\mathsf{B}}, \tag{4.19}$$

$$C^{\mathsf{B}}_{\lambda jum,\lambda' j' u' m'}=\delta_{\lambda\lambda'}\delta_{jj'}\delta_{mm'}C^{\lambda}_{uu'}, \tag{4.20}$$

$$C^{\alpha}=C^{\beta}=1,\quad C^{\gamma}=R^{\mathrm{C}}[j]. \tag{4.21}$$

The only purpose of the matrix C^{B} is to bring the representation γ into the form required for complex t reps, i.e., to ensure that

$$C^{\gamma}(\mathbf{e}^{\gamma}a+\mathbf{i}^{\gamma}b+\mathbf{j}^{\gamma}c+\mathbf{k}^{\gamma}d)=\sum\oplus C^{Q}[a+ib+jc+kd]. \tag{4.22}$$

Since $C^{\mathsf{B}+}$ is an absolutely irreducible representation of G the quaternionic subduction matrix is obtained from the complex one by multiplying it with a quaternionic matrix, which transforms $C^{\beta-}$ into $C^{\beta+}$ and decomposes C^{γ}.

$$Q^{\mathsf{B}(g)}=C^{\mathsf{B}+(g)}Q^{\mathsf{B}}, \tag{4.23}$$

$$Q^{\mathsf{B}}_{\lambda jum,\lambda' j' u' m'}=\delta_{\lambda\lambda'}\delta_{jj'}\delta_{mm'}Q^{\lambda}_{uu'}, \tag{4.24}$$

$$Q^{\alpha}=Q^{\beta+}=1,\ Q^{\beta-}=j,\quad Q^{\gamma}=Q,\quad Q \text{ see (4.6).} \tag{4.25}$$

The simplest case is $\Lambda=\Gamma$. We recall the peculiar structure of R^{Γ}, which indicates that this representation is essentially a quaternionic one:

$$R^{\Gamma}_{JR,J'R'}(X)=R^{Q}_{RR'}[Q^{\Gamma}_{JJ'}(X)]. \tag{4.26}$$

That the matrices $R^{\Gamma}(X)$ are composed of four-dimensional submatrices of the form R^{Q} is equivalent to the commutation relation

$$R^{\Gamma}(X)\bar{R}^{\Gamma}(q)=\bar{R}^{\Gamma}(q)R^{\Gamma}(X),\quad X\in G,\ q\in\mathbb{Q}, \tag{4.27}$$

$$\bar{R}^{\Gamma}(q)=\sum\oplus\bar{R}^{Q}[q]=\bar{Q}^{\Gamma},\quad \bar{Q}^{\Gamma} \text{ see (3.10)} \tag{4.28}$$

[cf. Ref. 7, Eq. (2.15)]. Because of (3.24–3.26) $R^{\Gamma(g)}$ is also composed of four-dimensional submatrices of the form R^{Q} so that

$$R^{\Gamma(g)}\bar{R}^{\Gamma}(q)=\bar{R}^{\Gamma}(q)R^{\Gamma(g)},\quad q\in\mathbb{Q}, \tag{4.29}$$

that is to say $R^{\Gamma(g)}$ is also essentially quaternionic.

$$R^{\Gamma(g)}_{JR,\lambda jtm} = R^{Q}_{Rt}\,[Q^{\Gamma(g)}_{J,\lambda jm}] \tag{4.30}$$

Since both the matrices $R^{\Gamma}(x)$, $x \in g$, and the subduction matrix $R^{\Gamma(g)}$ are transformed into quaternionic matrices of quarter dimension by the substitution $R^{Q}[q] \to q$ so is the direct sum $\Sigma \oplus m^{\Gamma\lambda}_{R} R^{\lambda}(x)$, showing the same structure as R^{Γ} and $R^{\Gamma(g)}$ because of (3.2) and (4.27, 4.29). Now

$$\sum \oplus m^{\Gamma\alpha}_{R} R^{\alpha}(x) = \sum \oplus m^{\Gamma\alpha}_{Q}[\oplus 4R^{\alpha}(x)],$$

$$[\oplus 4R^{\alpha}(x)]_{jt,j't'} = R^{Q}_{tt'}\,[R^{\alpha}_{j0,j'0}(x)], \tag{4.31}$$

$$\sum \oplus m^{\Gamma\beta}_{R} R^{\beta}(x) = \sum \oplus m^{\Gamma\beta}_{Q}[\oplus 2R^{\beta}(x)],$$

$$[\oplus 2R^{\beta}(x)]_{jt,j't'} = R^{Q}_{tt'}\,[R^{\beta}_{j0,j'0}(x) + iR^{\beta}_{j1,j'0}(x)], \tag{4.32}$$

so that

$$\sum \oplus m^{\Gamma\lambda}_{Q} Q^{\lambda}(x) = \sum \oplus m^{\Gamma\alpha}_{Q} R^{\alpha}(x) \oplus \sum \oplus m^{\Gamma\beta}_{Q} C^{\beta+}(x) \oplus \sum \oplus m^{\Gamma\gamma}_{Q} Q^{\gamma}(x). \tag{4.33}$$

Once the quaternionic matrices are known the corresponding complex matrices are obtained by substituting $q \to C^{Q}[q]$ [see Eq. (I.2.3)]:

$$C^{\Gamma}_{JT,J'T'}(X) = C^{Q}_{TT'}\,[Q^{\Gamma}_{JJ'}(X)], \tag{4.34}$$

$$C^{\Gamma(g)}_{JT,\lambda jtm} = C^{Q}_{Tt}\,[Q^{\Gamma(g)}_{J,\lambda jm}]. \tag{4.35}$$

Equations (4.33–4.35) imply that the representations appearing in $\Sigma \oplus m^{\Gamma\lambda}_{C} C^{\lambda}(x)$ are all complex t reps. Note that the representations $\beta+$ and $\beta-$ are interlocked since they belong to $t = 0, 1$.

5. REDUCTION OF REAL PARAMETERS

If one is only interested in complex subduction matrices it may seem a bit fancy to construct them via the real ones. However, the advantage of the approach considered here does not result from these details (which certainly may be substituted by equivalent ones) but from its spirit: If a matrix representation is chosen in type-adapted form the number of real parameters fixing the matrices is, on the average, only $\frac{2}{3}$ the amount needed to fix a nonadapted representation. Likewise comparing the subduction from a type-adapted representation to type-adapted representations with the subduction from a nonadapted representation to nonadapted representations of the subgroup one finds again a reduction of real parameters, this time to as low as $\frac{5}{9}$, averaged over all types of subductions. To see how this comes about consider, for instance, the subduction $A\beta\pm$. For non-adapted representations C^{A} is not real nor are $C^{\beta+}$ and $C^{\beta-}$ complex conjugate representations. If the projection method [based on the complex group algebra $\mathbf{A}_{C}(g)$] is used to determine the rows of $C^{A(g)}$ belonging to $\beta\pm$ one has to find $m^{A\beta+}_{C} + m^{A\beta-}_{C} (= 2\mu^{A\beta})$ orthonormalized vectors each having m^{A} $(= \dim C^{A})$ complex components. In the method proposed here only $\mu^{A\beta}$ $[= m^{A\beta}_{R}/2$, cf. (3.16, 3.19)] orthonormalized vectors with m^{A} real components have to be determined. In this case the ratio Tt/Nn is equal to $\frac{1}{4}$, if the numbers Tt and Nn are defined as follows:

number of real parameters needed to fix the columns of $C^{\Lambda(g)}$ belonging to λ

$$= Tt, \text{ if both } \Lambda \text{ and } \lambda \text{ are } t \text{ reps},$$
$$= Nn, \text{ if neither } \Lambda \text{ nor } \lambda \text{ are } t \text{ reps}. \tag{5.1}$$

For $\Lambda\lambda = A\beta\pm$ a reduction by $\frac{1}{2}$ may be attributed to the fact that C^{A} has been chosen to be real and $C^{\beta\pm}$ to be complex conjugate. A further reduction by $\frac{1}{2}$ is due to the fact that the real and imaginary parts of the matrix elements of $C^{A(g)}$ belonging to $\beta+$ (or $\beta-$) are related by the matrix $I^{(A)\beta}$ which is uniquely determined by A and β up to the sign.

Collecting these results for all pairs $\Lambda\lambda$ we arrive at the following table:

$\Lambda\lambda$	$A\alpha$	$A\beta\pm$	$A\gamma$	$B\pm\alpha$	$B\pm\beta\pm$	$B\pm\gamma$	$\Gamma\alpha$	$\Gamma\beta\pm$	$\Gamma\gamma$	average
Tt/Nn	1/2	1/4	1/4	1/1	1/2	1/2	1/2	1/2	1/1	5/9

(5.2)

It furthermore should be noted that the conventions which are always necessary to fix a subduction matrix completely reduce here to a choice and/or calculation of real numbers. That the matrices and vectors used in this method are always real might be of interest for numerical calculations. Moreover if the ambiguity inherent to this kind of problem is removed in the manner proposed here three related problems are solved in one run. Finally it is pointed out that the reasons to use t reps are even more stringent if one is interested in the noncomplex representations. Here no true alternative seems to exist. In principle one could construct successively irreducible subspaces (of column vectors) starting from cyclic representations, i.e., from the linear hulls of sets $\{D(x)v | x \in g\}$. In this approach the Schmidt process is the only mean to construct bases since at this stage no shift operators are known. But if one tries to pass from a nonadapted representation to linear combinations of these matrices suited to characterize invariant subspaces and to construct orthonormalized bases problems arise both for the real and the quaternionic representations.

For the real representations this is due to two facts: (i) Contrary to the complex case there exist no simple rules how to obtain projection and shift matrices if the matrix representation of the group is nonadapted. (ii) If $\lambda \neq \alpha$, matrices commuting with the matrices representing group elements must exist but their form is not obvious for nonadapted representations. Thus if a representation λ $(\neq \alpha)$ were given in nonadapted form one would have to find first the algebra of commuting matrices (which is isomorphic to $\mathbb{C}$ for $\lambda = \beta$ and to $\mathbb{Q}$ for $\lambda = \gamma$), and then to transform it into a peculiar form by

an orthogonal transformation. For $\lambda = \beta$ this would mean to find a matrix transforming the matrix $I^{\beta}_{\text{non}} = R^{\beta}_{\text{non}}(\mathbf{i}^{\beta})$ into the form (4.9); for $\lambda = \gamma$ it were matrices $\overline{R}^{\gamma}_{\text{non}}(q)$, $q = i, j$, that would have to be brought into the form (4.28). These orthogonal transformations transform the nonadapted representations into t reps which in turn allow to define projection and shift matrices. But even then a systematic explanation of the multiplicities in case $\Lambda\lambda = \mathsf{B}\alpha$, $\Gamma\alpha$, $\Gamma\beta$ is still missing until the algebra of matrices commuting with $R^{\Lambda}(X)$, $X \in G$, has been determined.

For quaternionic representations the situation is quite similar. A short reflection shows that the only linear combinations of group elements which certainly leave a subspace invaraint, if it is invariant under the action of the group, must have real coefficients. Hence it is only the real group algebra which can be used to find the desired subspaces (cf. Ref. 6, Sec. 3). This algebra is not obvious for nonadapted representations $\lambda \neq \gamma$. To recognize their type and the corresponding t basis these quaternionic representations have to be transformed into real ($\lambda = \alpha$) or complex form ($\lambda = \beta$). Fortunately one need not worry about how to find the hyperunitary matrices needed for these transformations since quaternionic representations are hardly given from the outset.

[1]P. Kasperkovitz, "Type-adapted representations of semidirect product groups", J. Math. Phys. **24**, 1 (1983), preceding paper.
[2]R. Dirl, On the uniqueness and reality of Clebsch Gordan coefficients. I. Ordinary representations. II. Corepresentations (submitted for publication).
[3]S. Schindler and R. Mirman, J. Math. Phys. **18**, 1678 (1977).
[4]P. M. van den Broek and J. F. Cornwell, Phys. Stat. Sol. B **90**, 211 (1978).
[5]R. Dirl, J. Math. Phys. **20**, 659 (1979).
[6]P. Kasperkovitz, J. Math. Phys. **22**, 2417 (1981).
[7]P. Kasperkovitz and G. Kahl, J. Math. Phys. **22**, 2404 (1981).

A new mathematical function connected with boundary value problems in kinetic transport theory

Zhe-ming Liu
Peking Feng-Yuan Institute of Mechanical and Electric Engineering, Building No. 66, 2–8, Wan-yuan Lu Rd., Peking, People's Republic of China

(Received 14 January 1981; accepted for publication 4 September 1981)

A new mathematical function connected with solving the Maxwell transport equation by applying the bimodal two-stream relaxing distribution is defined. This new function gives a more correct description of the direct nonequilibrium effect in gas molecular distribution on the macroscopic transferring of moment flux. In this paper, the differential equation satisfied by this function, its recurrence relations, and series or asymptotic expansions in various conditions are formulated. The degrees of approximations for these expansions are discussed.

PACS numbers: 02.30. + g, 51.10. + y

I. INTRODUCTION AND GENERAL CONSIDERATION

The distribution function of monoatomic molecules for dilute gases can be described by the Boltzmann integro-differential equation in the region of densities from the free-molecule realm to the continuous medium (bicollisions among molecules still play a dominant role), i.e.,

$$\frac{\partial f}{\partial t} + \xi \cdot \nabla_r f = \int\int\int (f'f'_1 - ff_1) v b \, db \, d\epsilon \, d\xi_1, \tag{1}$$

where f is the distribution function of molecules, $v = |\xi_1 - \xi|$ the relative velocity between two colliding molecules, b the parameter of collisions, and ϵ the collisional azimuthal angle.

The right-hand side of Eq. (1) is called a collisional term. Because of its nonlinearity, to solve Eq. (1) under certain boundary conditions is very difficult. Fortunately, in many real problems of physics one is interested not in distribution function itself, but in its some lower moments, e.g., gas density, temperature, flow velocity, shear stress, and heat flux. In order to obtain correct values of these macroscopic quantities, the moment method, model equation, and variational method are widely used. The moment method is a powerful instrument, especially for the nonlinear problems. Multiplying both sides of the Boltzmann equation (1) by the velocity function $\phi(\xi)$ and integrating it for all possible velocities of molecules, one gets the following Maxwell transport equation or moment equation:

$$\frac{\partial}{\partial t} \int \phi(\xi) f \, d\xi + \nabla_r \cdot \int \xi \phi(\xi) f \, d\xi = \Delta\phi, \tag{2}$$

where

$$\Delta\phi = \int\int\int\int (\phi' - \phi) f f_1 v b \, db \, d\epsilon \, d\xi_1 \, d\xi.$$

The differences between Eqs. (1) and (2) consist in that for the latter there is a possibility of not necessarily having to find the precise value of the distribution function point-by-point, but to put stress on computing the moments of the distribution function in some average sense. It is possible to construct a suitable form for the distribution function in advance in order to approximately solve Eq. (2) in the sense of correct macroscopic parameters of gases. However, this functional form must be determined in such a way that it can represent the physical nature of the problem, reflect the effect of soid boundaries, and make the mathematical treatment easy to carry out. Based on this idea, a bimodal two-stream relaxing distribution suggested by the author in Ref. 1 can be introduced here according to the following considerations:

(i) The distribution function should be discontinuous along the normal direction with respect to the surface of solid walls. This is particularly important for rarefied gases and (or) in strong nonlinear problems or near the solid boundaries within the region about the mean free path of molecules.

(ii) In the nonlinear cases, there must be a bimodal character emerging in the distribution function of the molecules. For example, this character needs to be accounted for in the problems like shock wave structure, heat transfer with large temperature gradiant, etc.

(iii) The influence of solid boundaries on the distribution function of gas molecules should be divided into two parts: The first is direct influence, i.e., reflected molecules from solid walls directly reach certain place in the gas field and constitute some part of local molecular ensemble. Molecules of this part may be described by the relaxing term for the distribution function of reflected molecules from corresponding solid surfaces, which is to be decayed exponentially along their trajectories due to molecular collisions. The second is indirect influence, i.e., contributions to the distribution function of molecules in that same place from collisions between the reflected molecules and the other gas molecules and from many-times collisions among the molecules which had been collided (directly or indirectly) with reflected molecules in their histories.

The above division is very important, because the nature of the velocity distribution between these two sets of molecules is quite different, in particular for the nonlinear cases.

Therefore, the total distribution function of gas molecules can be described by the following formula, which is

 0022-2488/83/010014-08$02.50

called the bimodal two-stream relaxing distribution:

$$f(\xi,\mathbf{r}) = a_1(\mathbf{r})f_1(\xi,\mathbf{r}) + a_2(\mathbf{r})f_2(\xi,\mathbf{r}) + \exp\left[-\frac{K_1|\mathbf{r}-\mathbf{r}_{w_1}|}{|\xi|}\right] f_{w_1}(\xi,\mathbf{r}_{w_1}), \quad \text{for } \xi\cdot\mathbf{n}_{w_1} > 0 \tag{3}$$

$$f(\xi,\mathbf{r}) = a_3(\mathbf{r})f_1(\xi,\mathbf{r}) + a_4(\mathbf{r})f_2(\xi,\mathbf{r}) + \exp\left[-\frac{K_2|\mathbf{r}-\mathbf{r}_{w_2}|}{|\xi|}\right] f_{w_2}(\xi,\mathbf{r}_{w_2}), \quad \text{for } \xi\cdot\mathbf{n}_{w_2} > 0$$

where $\mathbf{r}$ is a radius vector in a gas field, $\mathbf{r}_{w_1}$, $\mathbf{r}_{w_2}$ the original radius vectors of gas molecules on the surfaces of walls, $\mathbf{n}_{w_1}$, $\mathbf{n}_{w_2}$ the normals of two oppositely faced solid surfaces towards the gas, ξ the velocity of a gas molecule, $f_{w_1}(\xi,\mathbf{r}_{w_1})$ and $f_{w_2}(\xi,\mathbf{r}_{w_2})$ the distribution functions of reflected molecules from the solid boundaries, $a_1(\mathbf{r})$, $a_2(\mathbf{r})$, $a_3(\mathbf{r})$, and $a_4(\mathbf{r})$ the space influence functions (indirect) of solid boundaries, and $f_1(\xi,\mathbf{r})$, $f_2(\xi,\mathbf{r})$ may be selected as local Maxwellian distribution which contains several other space functions determined by moment equations.

By applying this distribution, formula (3), to the heat-conduction problem between two parallel plates, much better results than existing theory have been obtained in Ref. 1, including the total heat transfer and the temperature variation along the axis perpendicular to the walls.

In order for the problem to be solved, the number of moment equations is required to be equal to the number of unknown space influence functions.

For summational invariants of collisions, namely, the mass m, moment $m\xi$, and energy $m|\xi|^2/2$ of a molecule, the $\Delta\phi_s \equiv 0$ $(s = 1,2,3,4,5)$, which is independent of what the form of the distribution function is. Thus, for the steady problems there are five moment equations which may be selected:

$$\nabla_r \cdot \int \xi\phi_s(\xi) f\, d\xi = 0, \quad (s = 1,2,3,4,5). \tag{4}$$

However, besides that, at least one moment equation should be constructed, in which the velocity function $\phi_i(\xi)$ differs from above five collisional invariants $\phi_s(\xi)$.

For the $\phi_i(\xi)$ selected, the $\Delta\phi_i(\xi)$ is evaluated easily for the Maxwell molecules by (2), which possesses the following general form:

$$\Delta\phi_i(\xi) = \sum_k C_{ik} n/a_{ik}, \tag{5}$$

where C_{ik} are constants depending on the dynamics of collisions between molecules, $1/a_{ik}$ the moment flux in gases, which are constants in steady problems, $n = \int f\, d\xi$ the number density of molecules per unit volume.

Thus, the additional moment equations have the following form:

$$\nabla_r \cdot \int \xi\phi_i(\xi) f\, d\xi = \sum_k C_{ik} n/a_{ik}. \tag{6}$$

Introducing (3) into (6) and (4), one gets a system of nonhomogeneous ordinal differential equations of first order together with five algebraic equations for the space-influence functions.

Integrating these simultaneous differential equations, we obtain a new mathematical function indicating the direct effect of solid boundaries, where reflected molecules have Maxwellian distribution, on transferring the moment flux in gases. This new function is defined as

$$L_n(x,a) = \int_0^\infty \frac{u^{n+1}}{u+a} e^{-u^2 - x/u}\, du, \quad (n = 0,1,2,3,\cdots) \tag{7}$$

where x and a are two non-negative real numbers.

It is clear from the above discussion that the parameter $1/a$ characterizes the flux of macroscopic physical quantities of gases which in turn represents the degree of nonlinearity of the physical problem. When $1/a$ approaches zero, the net effect of solid boundaries on the distribution of gas molecules tends to zero, i.e., $L_n(x,\infty) \to 0$, because the system between solid boundaries and gases is in the complete thermal and (or) dynamic equilibrium states. But when $1/a$ is raised up approximately to infinity, the direct effect of solid boundaries on the distribution function of gas molecules becomes extremely strong and the physical flux gets very large. The increase of $1/a$ from zero to a large value shows that the physical problem changes from the linear one into the nonlinear one.

Similarily, in the transport problems of other neutral particles function (7) may also appear.

2. GENERAL PROPERTIES

The function $L_n(x,a)$ satisfies the following differential equation and recurrence relations:

$$ax\frac{\partial^4 L_n}{\partial x^4} - [x + (n-1)a]\frac{\partial^3 L_n}{\partial x^3} + (n-1)\frac{\partial^2 L_n}{\partial x^2} + 2a\frac{\partial L_n}{\partial x} - 2L_n = 0, \tag{8}$$

$$\frac{\partial L_n}{\partial x} = -L_{n-1}, \tag{9}$$

$$2L_n = axL_{n-4} + [x + (n-1)a]L_{n-3} + (n-1)L_{n-2} - 2aL_{n-1}. \tag{10}$$

Formula (9) is the simple result of the differentiating definition of $L_n(x,a)$. Integrating both $(n-1)L_{n-2}(x,a)$ and $(n-1)aL_{n-3}(x,a)$ by parts and introducing them into (10), one is in a position to prove that (10) is an identity. Equation (8) is the direct result of (9) and (10).

When $a = 0$ in (7),

$$L_n(x,0) = \int_0^\infty u^n e^{-u^2 - x/u} du \equiv J_n(x). \tag{11}$$

It can be seen from this that the case discussed by Abramowitz *et al.*[2,3] (i.e., $J_n(x)$) is a special case of $L_n(x,a)$.

The main differences in physical meaning between $J_n(x)$ and $L_n(x,a)$ may be explained as follows:

The function $J_n(x)$ represents the direct contribution of reflected molecules from soid boundaries to the local mass, momentum, and energy of gases, i.e., collisional invariants, carried by the reflected molecules themselves. $J_n(x)$ appears in the equations of conservation (4), which are independent of whether the moment–flux exists or not. However, $L_n(x,a)$ energes after integrating Eq. (6), which denotes the existence

of moment-flux in gases. Therefore, $L_n(x,a)$ is really connected with kinetic transport problems, which answer the question what part of the direct contributions of reflected molecules from solid boundaries to the transferring the moment-flux in the local gas field is. As the moment-flux becomes extremely large, $L_n(x,a)$ is in the same mathematical form as $J_n(x)$, but their physical meaning is quite different as shown by the above discussion.

3. EXPANSIONS FOR VARIOUS a AND x

A. The case of large a and small x

First of all, it can be pointed out that for all the values of a one finds following identity, proved by induction:

$$\frac{u^{n+1}}{u+a} = (-a)^{n+1}\frac{1}{u+a} + \sum_{s=0}^{n}(-a)^{n-s}u^s. \tag{12}$$

Then, we have

$$L_n(0,a) = \int_0^\infty \frac{u^{n+1}}{u+a}e^{-u^2}du = (-a)^{n+1}\int_0^\infty \frac{1}{u+a}e^{-u^2}du + \sum_{s=0}^{n}(-a)^{n-s}J_s(0). \tag{13}$$

Obviously, in the induction of (13), interchanging the order between integration and summation is permisible.

The asymptotic expansion of the first term on the right-hand side in (13) can be obtained by expanding the denominator of the integrand in descending powers of a[4]:

$$\int_0^\infty \frac{1}{u+a}e^{-u^2}du = \frac{1}{2}\sum_{r=0}^{\infty}\frac{(-1)^r}{a^{r+1}}\Gamma\left(\frac{r+1}{2}\right). \tag{14}$$

Thus

$$L_n(0,a) = (-a)^{n+1}\frac{1}{2}\sum_{r=0}^{\infty}\frac{(-1)^r}{a^{r+1}}\Gamma\left(\frac{r+1}{2}\right) + \frac{1}{2}\sum_{s=0}^{n}(-a)^{n-s}\Gamma\left(\frac{s+1}{2}\right) = (-a)^{n+1}\frac{1}{2}\sum_{r=n+1}^{\infty}\frac{(-1)^r}{a^{r+1}}\Gamma\left(\frac{r+1}{2}\right). \tag{15}$$

Let us now turn to discussing the expansion of $L_n(x,a)$. This can be obtained by use of the Laplace transform. The Laplace transform of $L_n(x,a)$ is defined as $\mathscr{L}\{L_n\} = \int_0^\infty e^{-tx}\times L_n(x,a)dx$, i.e.,

$$\mathscr{L}\{L_n\} = \int_0^\infty e^{-tx}\int_0^\infty \frac{u^{n+1}}{u+a}e^{-u^2-x/u}\,du\,dx. \tag{16}$$

Since (7) is absolutely convergent, the order of integration above may be changed. Consequently,

$$\mathscr{L}\{L_n\} = \frac{1}{t}\int_0^\infty \frac{1}{u+a}\frac{u^{n+2}}{u+1/t}e^{-u^2}du. \tag{17}$$

Formula (17) can be further rewritten by using (12) for $u^{n+2}/(u+1/t)$:

$$\mathscr{L}\{L_n\} = (-1)^n\frac{1}{t^{n+3}}\int_0^\infty \frac{1}{(u+a)(u+1/t)}e^{-u^2}du - \sum_{s=-1}^{n}(-1)^{n-s+1}\frac{1}{t^{n-s+1}}L_s(0,a). \tag{18}$$

When $a>0$, $L_{-1}(0,a)$ in (18) is convergent. The first term in the right-hand side in (18) may be changed:

$$(-1)^n\frac{1}{t^{n+3}}\int_0^\infty \frac{1}{(u+a)(u+1/t)}e^{-u^2}du = (-1)^{n+1}\frac{1}{a}\frac{1}{t^{n+2}(t-1/a)}L_{-1}(0,a) + (-1)^n\frac{1}{a}\frac{1}{t^{n+2}(t-1/a)}\int_0^\infty \frac{1}{u+1/t}e^{-u^2}du. \tag{19}$$

The series expansion of the integral on the right-hand side in (19) is[4]

$$\int_0^\infty \frac{1}{u+1/t}e^{-u^2}du = \sum_{r=0}^{\infty}\frac{(-1)^r\{\frac{1}{2}\psi(r+1)+\ln t\}}{r!t^{2r}} + \pi^{1/2}\sum_{r=0}^{\infty}\frac{(-2)^r}{1\cdot3\cdot5\cdots(2r+1)t^{2r+1}}, \tag{20}$$

where $\psi(r+1) = -\gamma + \Sigma_{m=1}^{r}(1/m)$, $\psi(1) = -\gamma$, and $\gamma = 0.577\,215\cdots$ is Euler's constant.

Substituting (19) and (20) into (18), we have

$$\mathscr{L}\{L_n\} = (-1)^{n+1}\frac{1}{a}\frac{1}{t^{n+2}(t-1/a)}L_{-1}(0,a) + (-1)^n\frac{1}{a}\left[\sum_{r=0}^{\infty}\frac{(-1)^r\{\frac{1}{2}\psi(r+1)+\ln t\}}{r!t^{2(r+1)+n}(t-1/a)} + \pi^{1/2}\sum_{r=0}^{\infty}\frac{(-2)^r}{1\cdot3\cdot5\cdots(2r+1)t^{2(r+1)+n+1}(t-1/a)}\right] - \sum_{s=-1}^{n}(-1)^{n-s+1}\frac{1}{t^{n-s+1}}L_s(0,a). \tag{21}$$

Using the theorem of convolution

$$\mathscr{L}^{-1}\left[\frac{1}{t^N}f(t)\right] = \left(\int_0^x ds\right)^N F(x), \tag{22}$$

and

$$\mathscr{L}^{-1}\left[\frac{1}{t-1/a}\right] = e^{x/a}, \tag{23}$$

we deduce

$$\mathscr{L}^{-1}\left[\frac{1}{t^N(t-1/a)}\right] = a^N\left(e^{x/a} - \sum_{p=0}^{N-1}\frac{1}{p!}(x/a)^p\right). \tag{24}$$

In the same way, using (22) and

$$\mathscr{L}^{-1}\left[\frac{\ln t}{t-1/a}\right] = e^{x/a}(\ln(1/a) + E_1(x/a)), \tag{25}$$

we obtain

$$\mathscr{L}^{-1}\left[\frac{\ln t}{t^N(t-1/a)}\right] = a^N\left[e^{x/a}(\ln(1/a)+E_1(x/a)) - \sum_{s=0}^{N-1}\frac{1}{s!}(x/a)^s(\psi(s+1)-\ln x)\right], \tag{26}$$

where $E_1(x/a) = \int_{x/a}^\infty (e^{-v}/v)\,dv$ is the exponential integral Its series expansion takes the form[5]:

$$E_1(x/a) = -\gamma - \ln(x/a) - \sum_{n=1}^{\infty}\frac{(-1)^n}{nn!}(x/a)^n. \tag{27}$$

In the deduction of (26) attention has been paid to the

following integral[5]:

$$\int_0^{x/a} e^s E_1(s)\,ds = e^{x/a}E_1(x/a) + \gamma + \ln x/a. \tag{28}$$

Formula (26) may be further rewritten to simplify the procedure for getting the expansion of $L_n(x,a)$. Using the serious expansion of $E_1(x/a)$ and the definition of the function $\psi(s)$, we deduce that

$$e^{x/a}(\ln(1/a) + E_1(x/a)) = \sum_{s=0}^{\infty} \frac{1}{s!}(x/a)^s(\psi(s+1) - \ln x). \tag{29}$$

Thus, (26) may be further rewritten as

$$\mathscr{L}^{-1}\left[\frac{\ln t}{t^N(t-1/a)}\right] = a^N \sum_{s=N}^{\infty} \frac{1}{s!}(x/a)^s(\psi(s+1) - \ln x). \tag{30}$$

Substituting (24), and (30) into (21), we obtain following expansion for $L_n(x,a)$ in the case of large a and small x:

$$L_n(x,a) = \sum_{s=-1}^{n} \frac{1}{(n-s)!}(-x)^{n-s}L_s(0,a) + (-a)^{n+1}L_{-1}(0,a) \times \sum_{s=n+2}^{\infty} \frac{1}{s!}(x/a)^s + (-1)^n \sum_{s=2}^{\infty} \frac{1}{a^{s-1}}\Omega_{n+s}(x), \tag{31}$$

where

$$\Omega_{n+s}(x) = \sum_{r=0}^{\infty} \frac{(-1)^r\{\frac{1}{2}\psi(r+1) + \psi(2r+n+s+1) - \ln x\}}{r!(2r+n+s)!}x^{2r+n+s} + \pi^{1/2}\sum_{r=0}^{\infty} \frac{(-2)^r}{1\cdot 3\cdot 5\cdots(2r+1)[(2r+n+s+1)!]}x^{2r+n+s+1}. \tag{32}$$

We shall discuss the order of approximation of the expansion (31) in detail as follows:

Because the series (32) in an alternating convergent one, if we sum its terms through κ, the remainder satisfies

$$R^{(\kappa+1)}_{\Omega_{n+s}} < \left|(-1)^{\kappa+1}\frac{\frac{1}{2}\psi(\kappa+2) + \psi(2\kappa+n+s+3) - \ln x}{(\kappa+1)!(2\kappa+n+s+2)!}x^{2\kappa+n+s+2} + \pi^{1/2}\frac{(-2)^{\kappa+1}}{1\cdot 3\cdot 5\cdots(2\kappa+3)[(2\kappa+n+s+3)!]}x^{2\kappa+n+s+3}\right|. \tag{33}$$

Therefore, (31) may be expressed as

$$\begin{aligned} L_n(x,a) = {} & \sum_{s=-1}^{n} \frac{1}{(n-s)!}(-x)^{n-s}L_s(0,a) + (-a)^{n+1}L_{-1}(0,a)\sum_{s=n+2}^{\infty}\frac{1}{s!}(x/a)^s \\ & + (-1)^n\sum_{r=0}^{\kappa}\frac{(-1)^r}{r!}a^{2r+n+1}\sum_{s=2}^{\infty}\psi(2r+n+s+1)\frac{(x/a)^{2r+n+s}}{(2r+n+s)!} + (-1)^n\sum_{s=2}^{\infty}\frac{1}{a^{s-1}}R^{(\kappa+1)}_{\Omega_{n+s}} \\ & + (-1)^n\sum_{r=0}^{\kappa}\frac{(-1)^r}{r!}a^{2r+n+1}\{\tfrac{1}{2}\psi(r+1) - \ln x\}\sum_{s=2}^{\infty}\frac{(x/a)^{2r+n+s}}{(2r+n+s)!} \\ & + (-1)^n\pi^{1/2}\sum_{r=0}^{\kappa}\frac{(-2)^r}{1\cdot 3\cdot 5\cdots(2r+1)}a^{2r+n+2}\sum_{s=2}^{\infty}\frac{(x/a)^{2r+n+s+1}}{(2r+n+s+1)!}. \end{aligned} \tag{34}$$

Since the last two terms in the right-hand side of (34) can be calculated precisely, only the first four infinite series in (34), namely $L_s(0,a)$ in the first term, $L_{-1}(0,a)$ in the second term, and the third and fourth terms, remain to be discussed in estimating the error of approximation.

First, we shall determinate the error of approximation of the first and second terms in (34), which is caused by $L_s(0,a)$ having the asymptotic expansion

$$L_s(0,a) = (-a)^{s+1}\frac{1}{2}\sum_{r=s+1}^{\infty}\frac{(-1)^r}{a^{r+1}}\Gamma\left(\frac{r+1}{2}\right). \tag{35}$$

Observing that the series in (35) is an alternating one and summing its terms through $r = \kappa$ (which must be larger than $s+1$), we find that the remainder of series (35) satisfies

$$R^{(\kappa+1)}_{L_s(0,a)} < \left|(-1)^{s+\kappa}\frac{1}{4}\frac{\kappa}{a^{\kappa-s+1}}\Gamma\left(\frac{\kappa}{2}\right)\right|. \tag{36}$$

Therefore, when we take $r = \kappa$ in $L_s(0,a)$ and $L_{-1}(0,a)$, the remainder in the first and second terms is to be estimated as

$$\begin{aligned} & R^{(\kappa+1)}_{1+2}(x,a) \\ & < \left|\sum_{s=-1}^{n}(-1)^{n+\kappa}\frac{\kappa}{4}\Gamma(\kappa/2)\frac{1}{(n-s)!}(x/a)^{n-s}\frac{1}{a^{\kappa-n+1}} + (-1)^{n+\kappa}\frac{\kappa}{4}\Gamma(\kappa/2)\frac{1}{a^{\kappa-n+1}}\sum_{s=n+2}^{\infty}\frac{1}{s!}(x/a)^s\right| \\ & = \left|(-1)^{n+\kappa}\frac{\kappa}{4}\Gamma(\kappa/2)\frac{1}{a^{\kappa-n+1}}e^{x/a}\right| \end{aligned} \tag{37}$$

Second, we discuss the third term in (34), which may be rewritten as

$$(-1)^n \sum_{r=0}^{\kappa} \frac{(-1)^r}{r!} a^{2r+n+1} \sum_{s=2}^{\infty} \psi(2r+n+s+1) \frac{(x/a)^{2r+n+s}}{(2r+n+s)!}$$
$$= (-1)^n \sum_{r=0}^{\kappa} \frac{(-1)^r}{r!} a^{2r+n+1} \sum_{s=2}^{l} \psi(2r+n+s+1) \frac{(x/a)^{2r+n+s}}{(2r+n+s)!} + R_3^{(l+1)}(x,a), \tag{38}$$

where the remainder is

$$\begin{aligned} R_3^{(l+1)}(x,a) = \Bigg| (-1)^n \sum_{r=0}^{\kappa} \frac{(-1)^r}{r!} a^{2r+n+1} \Bigg[\Bigg(-\gamma + 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{2r+n+l+1} \Bigg) \frac{(x/a)^{2r+n+l+1}}{(2r+n+l+1)!} \\ + \Bigg(-\gamma + 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{2r+n+l+1} + \frac{1}{2r+n+l+2} \Bigg) \frac{(x/a)^{2r+n+l+2}}{(2r+n+l+2)!} \\ + \Bigg(-\gamma + 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{2r+n+l+1} + \frac{1}{2r+n+l+2} + \frac{1}{2r+n+l+3} \Bigg) \frac{(x/a)^{2r+n+l+3}}{(2r+n+l+3)!} \\ + \Bigg(-\gamma + 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{2r+n+l+1} + \frac{1}{2r+n+l+2} + \frac{1}{2r+n+l+3} + \frac{1}{2r+n+l+4} \Bigg) \frac{(x/a)^{2r+n+l+4}}{(2r+n+l+4)!} + \cdots \Bigg] \Bigg| \\ = \Bigg| (-1)^n \sum_{r=0}^{\kappa} \frac{(-1)^r}{r!} a^{2r+n+1} \Bigg[\Bigg(-\gamma + 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{2r+n+l+1} \Bigg) \Bigg(e^{x/a} - \sum_{p=0}^{2r+n+l} \frac{1}{p!} (x/a)^p \Bigg) \\ + \frac{1}{2r+n+l+2} \frac{(x/a)^{2r+n+l+2}}{(2r+n+l+2)!} + \Bigg(\frac{1}{2r+n+l+2} + \frac{1}{2r+n+l+3} \Bigg) \frac{(x/a)^{2r+n+l+3}}{(2r+n+l+3)!} \\ + \Bigg(\frac{1}{2r+n+l+2} + \frac{1}{2r+n+l+3} + \frac{1}{2r+n+l+4} \Bigg) \frac{(x/a)^{2r+n+l+4}}{(2r+n+l+4)!} + \cdots \Bigg] \Bigg|. \end{aligned} \tag{39}$$

Owing to the relations

$$\begin{aligned} &\frac{1}{2r+n+l+2} \frac{(x/a)^{2r+n+l+2}}{(2r+n+l+2)!} \\ &\quad = \frac{x/a}{2r+n+l+2} \frac{(x/a)^{2r+n+l+1}}{(2r+n+l+2)!} \Bigg(\frac{1}{2r+n+l+2} + \frac{1}{2r+n+l+3} \Bigg) \frac{(x/a)^{2r+n+l+3}}{(2r+n+l+3)!} \\ &\quad < \Bigg(\frac{x/a}{2r+n+l+2} \Bigg)^2 \frac{(x/a)^{2r+n+l+1}}{(2r+n+l+2)!} + \frac{x/a}{2r+n+l+3} \frac{(x/a)^{2r+n+l+2}}{(2r+n+l+3)!}, \\ &\Bigg(\frac{1}{2r+n+l+2} + \frac{1}{2r+n+l+3} + \frac{1}{2r+n+l+4} \Bigg) \frac{(x/a)^{2r+n+l+4}}{(2r+n+l+4)!} \\ &\quad < \Bigg(\frac{x/a}{2r+n+l+2} \Bigg)^3 \frac{(x/a)^{2r+n+l+1}}{(2r+n+l+2)!} + \Bigg(\frac{x/a}{2r+n+l+3} \Bigg)^2 \frac{(x/a)^{2r+n+l+2}}{(2r+n+l+3)!} + \frac{x/a}{2r+n+l+4} \frac{(x/a)^{2r+n+l+3}}{(2r+n+l+4)!}, \\ &\quad \cdots < \cdots \\ &\quad \cdots < \cdots, \end{aligned} \tag{40}$$

we have

$$\begin{aligned} R_3^{(l+1)}(x,a) < \Bigg| (-1)^n \sum_{r=0}^{\kappa} \frac{(-1)^r}{r!} a^{2r+n+1} \Bigg\{ \psi(2r+n+l+2) \Bigg(e^{x/a} - \sum_{p=0}^{2r+n+l} \frac{1}{p!} (x/a)^p \Bigg) \\ + \Bigg[\frac{x/a}{2r+n+l+2} + \Bigg(\frac{x/a}{2r+n+l+2} \Bigg)^2 + \Bigg(\frac{x/a}{2r+n+l+2} \Bigg)^3 + \cdots \Bigg] \frac{(x/a)^{2r+n+l+1}}{(2r+n+l+2)!} \\ + \Bigg[\frac{x/a}{2r+n+l+3} + \Bigg(\frac{x/a}{2r+n+l+3} \Bigg)^2 + \Bigg(\frac{x/a}{2r+n+l+3} \Bigg)^3 + \cdots \Bigg] \frac{(x/a)^{2r+n+l+2}}{(2r+n+l+3)!} \\ + \Bigg[\frac{x/a}{2r+n+l+4} + \Bigg(\frac{x/a}{2r+n+l+4} \Bigg)^2 + \Bigg(\frac{x/a}{2r+n+l+4} \Bigg)^3 + \cdots \Bigg] \frac{(x/a)^{2r+n+l+3}}{(2r+n+l+4)!} + \cdots \Bigg\} \Bigg| \\ = \Bigg| (-1)^n \sum_{r=0}^{\kappa} \frac{(-1)^r}{r!} a^{2r+n+1} \Bigg[\psi(2r+n+l+2) \Bigg(e^{x/a} - \sum_{p=0}^{2r+n+l} \frac{1}{p!} (x/a)^p \Bigg) \\ + \frac{1}{2r+n+l+2-x/a} \frac{(x/a)^{2r+n+l+2}}{(2r+n+l+2)!} + \frac{1}{2r+n+l+3-x/a} \frac{(x/a)^{2r+n+l+3}}{(2r+n+l+3)!} \\ + \frac{1}{2r+n+l+4-x/a} \frac{(x/a)^{2r+n+l+4}}{(2r+n+l+4)!} + \cdots \Bigg] \Bigg| \\ < \Bigg| (-1)^n \sum_{r=0}^{\kappa} \frac{(-1)^r}{r!} a^{2r+n+1} \Bigg[\psi(2r+n+l+2) \Bigg(e^{x/a} - \sum_{p=0}^{2r+n+l} \frac{1}{p!} (x/a)^p \Bigg) \end{aligned}$$

$$+\frac{1}{2r+n+l+2-x/a}\left(e^{x/a}-\sum_{p=0}^{2r+n+l+1}\frac{1}{p!}(x/a)^p\right)\Bigg]\Bigg|$$

$$<\left|(-1)^n\sum_{r=0}^{\kappa}\frac{(-1)^r}{r!}a^{2r+n+1}\left[\psi(2r+n+l+2)\left(e^{x/a}-\sum_{p=0}^{2r+n+l}\frac{1}{p!}(x/a)^p\right)\right.\right.$$

$$\left.\left.+\frac{1}{2r+n+l+2}\left(e^{x/a}-\sum_{p=0}^{2r+n+l}\frac{1}{p!}(x/a)^p\right)\right]\right|$$

$$=\left|(-1)^n\sum_{r=0}^{\kappa}\frac{(-1)^r}{r!}a^{2r+n+1}\psi(2r+n+l+3)\left(e^{x/a}-\sum_{p=0}^{2r+n+l}\frac{1}{p!}(x/a)^p\right)\right|. \tag{41}$$

Finally, we are going to deal with the fourth term in (34), which may be estimated as follows:

$$\left|(-1)^n\sum_{s=2}^{\infty}\frac{1}{a^{s-1}}R^{(\kappa+1)}_{\Omega_{n+s}}\right|$$

$$=\left|(-1)^{n+\kappa+1}\frac{1}{(\kappa+1)!}a^{2\kappa+n+3}\sum_{s=2}^{\infty}\psi(2\kappa+n+s+3)\frac{(x/a)^{2\kappa+n+s+2}}{(2\kappa+n+s+2)!}\right.$$

$$+(-1)^{n+\kappa+1}\frac{1}{(\kappa+1)!}a^{2\kappa+n+3}(\tfrac{1}{2}\psi(\kappa+2)-\ln x)\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+3}\frac{1}{p!}(x/a)^p\right)$$

$$\left.+(-1)^{n+\kappa+1}\pi^{1/2}\frac{2^{\kappa+1}}{1\cdot3\cdot5\cdots(2\kappa+3)}a^{2\kappa+n+4}\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+4}\frac{1}{p!}(x/a)^p\right)\right|$$

$$<\left|(-1)^{n+\kappa+1}\frac{1}{(\kappa+1)!}a^{2\kappa+n+3}\psi(2\kappa+n+6)\left(e^{x/a}-\sum_{p=0}^{2r+n+l+3}\frac{1}{p!}(x/a)^p\right)\right.$$

$$+(-1)^{n+\kappa+1}\frac{1}{(\kappa+1)!}a^{2\kappa+n+3}(\tfrac{1}{2}\psi(\kappa+2)-\ln x)\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+3}\frac{1}{p!}(x/a)^p\right)$$

$$\left.+(-1)^{n+\kappa+1}\pi^{1/2}\frac{2^{\kappa+1}}{1\cdot3\cdot5\cdots(2\kappa+3)}a^{2\kappa+n+4}\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+4}\frac{1}{p!}(x/a)^p\right)\right|$$

$$=\left|(-1)^{n+\kappa+1}\frac{1}{(\kappa+1)!}a^{2\kappa+n+3}(\tfrac{1}{2}\psi(\kappa+2)+\psi(2\kappa+n+6)-\ln x)\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+3}\frac{1}{p!}(x/a)^p\right)\right.$$

$$\left.+(-1)^{n+\kappa+1}\pi^{1/2}\frac{2^{\kappa+1}}{1\cdot3\cdot5\cdots(2\kappa+3)}a^{2\kappa+n+4}\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+4}\frac{1}{p!}(x/a)^p\right)\right|. \tag{42}$$

In the deduction of (42) we have used (41), in which we have taken $l=3$.

Summing (37), (41), and (42), we obtain the total error of approximation for the expansion of $L_n(x,a)$ under the condition of large a and small x, i.e.,

$$E^{(k+1),(l+1)}_{L_n(x,a)}<\left|(-1)^{n+\kappa}\tfrac{1}{4}\kappa\Gamma(\kappa/2)\frac{1}{a^{\kappa-n+1}}e^{x/a}\right.$$

$$+(-1)^n\sum_{r=0}^{\kappa}\frac{(-1)^r}{r!}a^{2r+n+1}\psi(2r+n+l+3)\left(e^{x/a}-\sum_{p=0}^{2r+n+l}\frac{1}{p!}(x/a)^p\right)$$

$$+(-1)^{n+\kappa+1}\frac{1}{(\kappa+1)!}a^{2\kappa+n+3}(\tfrac{1}{2}\psi(\kappa+2)+\psi(2\kappa+n+6)-\ln x)\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+3}\frac{1}{p!}(x/a)^p\right)$$

$$\left.+(-1)^{n+\kappa+1}\pi^{1/2}\frac{2^{\kappa+1}}{1\cdot3\cdot5\cdots(2\kappa+3)}a^{2\kappa+n+4}\left(e^{x/a}-\sum_{p=0}^{2\kappa+n+4}\frac{1}{p!}(x/a)^p\right)\right|, \tag{43}$$

where $r=\kappa$ and $s=l$ are taken in the expansion (31) for $L_n(x,a)$.

B. The case of large a and large x

In this case, the denominator in the integrand of (7) can be expanded in descending powers of a. Then, the asymptotic expansion of $L_n(x,a)$ is

$$L_n(x,a)=\sum_{r=0}^{\kappa}\frac{(-1)^r}{a^{r+1}}\int_0^{\infty}u^{n+r+1}e^{-u^2-x/u}\,du+\frac{(-1)^{\kappa+1}}{a^{\kappa+1}}\int_0^{\infty}\frac{u^{n+\kappa+2}}{u+a}e^{-u^2-x/u}\,du$$

$$=\sum_{r=0}^{\kappa}\frac{(-1)^r}{a^{r+1}}J_{n+r+1}(x)+\frac{(-1)^{\kappa+1}}{a^{\kappa+1}}L_{n+\kappa+1}(x,a)$$

$$=\sum_{r=0}^{\kappa}\frac{(-1)^r}{a^{r+1}}J_{n+r+1}(x)+R^{(\kappa+1)}_{L_n}(x,a). \tag{44}$$

The remainder in (44) satisfies

$$R_{L_n}^{(\kappa+1)}(x,a) < \frac{1}{a^{\kappa+1}} J_{n+\kappa+1}(x). \tag{45}$$

The asymptotic expansion of $J_{n+\kappa+1}(x)$ has been studied by Abramowitz *et al.*[2,3] Thus, it is not necessary to discuss it further here.

C. The case of small a and small x

First of all notice that for small a the following expression can be obtained by use of (12):

$$L_n(0,a) = (-a)^{n+1}\left[\sum_{r=0}^{\infty}\frac{(-1)^r}{r!}\left\{\tfrac{1}{2}\psi(r+1)+\ln\frac{1}{a}\right\}a^{2r} + \pi^{1/2}\sum_{r=0}^{\infty}\frac{(-2)^r}{1\cdot3\cdot5\cdots(2r+1)}a^{2r+1}\right] + \tfrac{1}{2}\sum_{s=0}^{n}(-a)^{n-s}\Gamma\left(\frac{s+1}{2}\right), \quad (n=0,1,2,3,\cdots). \tag{46}$$

We now turn to discussing the expansion of $L_n(x,a)$ for small a and small x by use of the Laplace transform.

Formula (17) may be rewritten as

$$\mathscr{L}\{L_n\} = \frac{1}{t}\int_0^{\infty}\frac{u}{u+1/t}\frac{u^{n+1}}{u+a}e^{-u^2}\,du. \tag{47}$$

Using (12) for $u^{n+1}/(u+a)$, we obtain

$$\mathscr{L}\{L_n\} = (-a)^{n+1}\frac{1}{t}\int_0^{\infty}\frac{u}{(u+1/t)(u+a)}e^{-u^2}\,du + \frac{1}{t}\sum_{s=0}^{n}(-a)^{n-s}\int_0^{\infty}\frac{u^{s+1}}{u+1/t}e^{-u^2}\,du. \tag{48}$$

With the aid of (12) for $u^{s+1}/(u+1/t)$, we have

$$\mathscr{L}\{L_n\} = (-a)^{n+1}\frac{1}{t}\int_0^{\infty}\frac{u}{(u+1/t)(u+a)}e^{-u^2}\,du + (-1)^{n+1}\sum_{s=0}^{n}a^{n-s}\frac{1}{t^{s+2}}\int_0^{\infty}\frac{1}{u+1/t}e^{-u^2}\,du + \sum_{s=0}^{n}\sum_{r=0}^{s}(-1)^{n-r}a^{n-s}\frac{1}{t^{s-r+1}}J_r(0). \tag{49}$$

Since the value of t in the integrand of

$$(-a)^{n+1}\int_0^{\infty}\frac{1}{t}\frac{u}{(u+1/t)(u+a)}e^{-u^2}\,du \tag{50}$$

can be regarded as very large, (49) may be rewritten as

$$\mathscr{L}\{L_n\} = (-a)^{n+1}\frac{1}{t}\int_0^{\infty}\frac{1}{u+a}e^{-u^2}\,du + (-1)^{n+1}\sum_{s=0}^{n}a^{n-s}\frac{1}{t^{s+2}}\int_0^{\infty}\frac{1}{u+1/t}e^{-u^2}\,du + \sum_{s=0}^{n}\sum_{r=0}^{n}(-1)^{n-r}a^{n-s}\frac{1}{t^{s-r+1}}J_r(0). \tag{51}$$

Using the transforms

$$\mathscr{L}^{-1}\left[\frac{1}{t^{\kappa+1}}\right] = \frac{x^{\kappa}}{\kappa!},$$
$$\mathscr{L}^{-1}\left[\frac{\ln t}{t^{\kappa+1}}\right] = \frac{x^{\kappa}}{\kappa!}(\psi(\kappa+1)-\ln x), \tag{52}$$
$$\mathscr{L}^{-1}\left[\frac{1}{t^{\kappa}}\int_0^{\infty}\frac{1}{u+1/t}e^{-u^2}\,du\right] = \Omega_{\kappa-1}(x),$$

we obtain

$$L_n(x,a) = (-a)^{n+1}L_{-1}(0,a) + (-1)^{n+1}\sum_{s=0}^{n}a^{n-s}\Omega_{s+1}(x) + \frac{1}{2}\sum_{s=0}^{n}\sum_{r=0}^{s}(-1)^{n-r}a^{n-s}\frac{1}{(s-r)!}x^{s-r}\Gamma\left(\frac{r+1}{2}\right), \tag{53}$$

where

$$L_{-1}(0,a) = \sum_{r=0}^{\infty}\frac{(-1)^r}{r!}\left\{\frac{1}{2}\psi(r+1)+\ln\frac{1}{a}\right\}a^{2r} + \pi^{1/2}\sum_{r=0}^{\infty}\frac{(-2)^r}{1\cdot3\cdot5\cdots(2r+1)}a^{2r+1}. \tag{54}$$

Because the first two terms on the right-hand side in (53) are alternating convergent series, if we sum all its terms through $r=\kappa$, the remainder of $L_n(x,a)$ in (53) satisfies

$$R_{L_n(x,a)}^{(\kappa+1)} < \left|(-a)^{n+1}\left[\frac{(-1)^{\kappa+1}}{(\kappa+1)!}\left\{\frac{1}{2}\psi(\kappa+2)+\ln\frac{1}{a}\right\}a^{2\kappa+2} + \pi^{1/2}\frac{(-2)^{\kappa+1}}{1\cdot3\cdot5\cdots(2\kappa+3)}a^{2\kappa+3}\right] + (-1)^{n+1}\sum_{s=0}^{n}a^{n-s}\left[\frac{(-1)^{\kappa+1}}{(\kappa+1)!} \times\{\tfrac{1}{2}\psi(\kappa+2)+\psi(2\kappa+s+4)-\ln x\}\frac{x^{2\kappa+s+3}}{(2\kappa+s+3)!} + \pi^{1/2}\frac{(-2)^{\kappa+1}}{1\cdot3\cdot5\cdots(2\kappa+3)}\frac{x^{2\kappa+s+4}}{(2\kappa+s+4)!}\right]\right|. \tag{55}$$

D. The case of small a and large x

From (7) and (12) we have

$$L_n(x,a) = \int_0^{\infty}\frac{u^{n+1}}{u+a}e^{-u^2-x/u}\,du = (-a)^n\int_0^{\infty}\frac{u}{u+a}e^{-u^2-x/u}\,du + \sum_{s=1}^{n}(-a)^{n-s}J_s(x). \tag{56}$$

Since a is small and x is large and the relation

$$e^{-a/u} \leqslant \frac{u}{u+a} \leqslant 1 \tag{57}$$

holds for $0 \leqslant u \leqslant \infty$, we obtain the following two approximate

expressions for $L_n(x,a)$ by taking $e^{-a/u}$ and 1, respectively, in place of $u/(u+a)$ involved in the first term on right-hand side of (56):

$$L_n(x,a) = (-a)^n J_0(x+a) + \sum_{s=1}^{n} (-a)^{n-s} J_s(x), \tag{58}$$

$$L_n(x,a) = \sum_{s=0}^{n} (-a)^{n-s} J_s(x). \tag{59}$$

The errors of these two approximate expressions both satisfy

$$E_{L_n(x,a)} \leqslant a^n (J_0(x) - J_0(x+a)). \tag{60}$$

[1]Liu Zhe-ming, Acta Mech. Sinica **3**, 259 (1979).
[2]M. Abromowitz, J. Math. Phys. **32**, 188 (1953).
[3]M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions, Appl. Math. Ser.* **55** (U.S. Govt. Printing Office, Washington D.C., 1966).
[4]E. T. Goodwin and J. Staton, Quart J. Mech. **1**, 319 (1948).
[5]V. Kourganoff, *Basic Method in Transfer Problems* (Oxford U.P., London, 1952).

Null field solutions of the wave equation and certain generalizations

C. B. Collins
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada

(Received 10 November 1981; accepted for publication 11 August 1982)

The ordinary wave equation in 3 + 1 dimensions $\Box\phi = 0$,

$\Box \equiv -\partial^2/\partial t^2 + \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$

admits null field solutions, characterized by $\nabla\phi\cdot\nabla\phi = 0$,

$\nabla\phi\cdot\nabla\phi \equiv -(\partial\phi/\partial t)^2 + (\partial\phi/\partial x)^2 + (\partial\phi/\partial y)^2 + (\partial\phi/\partial z)^2$

with $\nabla\phi \neq \mathbf{0}$. It is shown that the general null field solution can be obtained from a knowledge of the "time-transported" solutions, i.e., those solutions of the form $\phi = t - \psi(x,y,z)$, where ψ satisfies both Laplace's equation and the eikonal equation in a Euclidean space. We obtain all second-order scalar wave equations of form $f(\phi,\phi_{;i},\phi_{;i;j}\phi^{;i;j}) = 0$ (in arbitrary dimension and involving a single potential function ϕ) for which the above technique applies. These equations are shown to be equivalent to the family of quasilinear third-order equations $\nabla\phi\cdot\nabla(\Box\phi) + K(\Box\phi)^2 = 0$, where K is a constant. Some null solutions of these equations are considered, and related to previous works. The results are applied to determine all shear-free hypersurface-orthogonal null geodesic congruences in Minkowski space–time, and some brief comments are made on complex solutions and on more general wave equations.

PACS numbers: 02.20.Jr, 04.20.Jb

I. INTRODUCTION

Friedlander[1,2] has considered simple progressing wave solutions of the scalar wave equation in 3 + 1 dimensions. For future reference and for the sake of brevity, we shall discuss the more general situation in n-dimensional Minkowski space–time, in which there are coordinates (x^i) $(i = 0,1,2,...,n-1; n \geqslant 2)$ such that the metric is

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + \cdots + (dx^{n-1})^2 = \eta_{ij}\,dx^i dx^j, \tag{1.1}$$

where $\eta_{ij} = \mathrm{diag}(-1,+1,+1,...,+1)$, with $0 \leqslant i,j \leqslant n$, and where the summation convention is employed on repeated indices. The wave equation is then

$$\Box\phi = 0, \quad \Box = -\frac{\partial^2}{(\partial x^0)^2} + \sum_{\alpha=1}^{n-1}\frac{\partial^2}{(\partial x^\alpha)^2}$$

or

$$\phi_{;i}{}^{;i} = 0 \tag{1.2}$$

in tensorial form, a semicolon indicating a (covariant[3]) derivative with respect to the metric (1.1). The simple progressing wave solutions of (1.2) have the special form $\phi = Uf(S)$, where f is arbitrary and S is not identically constant; they generalize the well-known d'Alembert solution, in which $U \equiv 1$ and $S = x^0 \pm x^1$ (and, for example, $n = 4$). Substituting $\phi = Uf(S)$ in (1.2) and recognizing that f is arbitrary leads to the overdetermined system of partial differential equations

$$\nabla S\cdot\nabla S = 0 \Leftrightarrow -\left(\frac{\partial S}{\partial x^0}\right)^2 + \sum_{\alpha=1}^{n-1}\left(\frac{\partial S}{\partial x^\alpha}\right)^2 = 0 \Leftrightarrow S_{;i}S^{;i} = 0, \tag{1.3a}$$

$$2\nabla S\cdot\nabla U + U\Box S = 0 \Leftrightarrow 2S_{;i}U^{;i} + US_{;i}{}^{;i} = 0, \tag{1.3b}$$

and

$$\Box U = 0 \Leftrightarrow U_{;i}{}^{;i} = 0, \tag{1.3c}$$

where the operators ∇ and $\Box$ are with respect to the metric (1.1). In his original paper,[1] Friedlander considered the special case (with $n = 4$): $U = V(x^1,x^2,...,x^{n-1})$ and $S = x^0 - \tau(x^1,x^2,...,x^{n-1})$. Then the system (1.3) becomes

$$\begin{aligned} &\nabla\tau\cdot\nabla\tau = 1, \\ &2\nabla\tau\cdot\nabla V + V\nabla^2\tau = 0, \\ &\nabla^2 V = 0, \end{aligned} \tag{1.4}$$

where the operators ∇ and ∇^2 are with respect to the $(n-1)$-dimensional Euclidean metric $ds^2 = (dx^1)^2 + (dx^2)^2 + \cdots + (dx^{n-1})^2$ induced on the hypersurfaces $\{x^0 = \text{const}\}$. However, in his later book,[2] Friedlander shows that, somewhat surprisingly, the general solution of the system (1.3) can (for $n = 4$) be reduced locally to the special case (1.4), in the sense that knowledge of the general solution of (1.4) is sufficient to determine implicitly the general solution of (1.3). We now provide a description of Friedlander's procedure, which will be of use later on. However, this description will be in tensorial notation, thereby rendering the procedure more transparent; it will automatically involve the generalization from 4 to n dimensions.

The procedure involves a change of coordinates. By (1.3a), we have $\partial S/\partial x^0 \neq 0$, for otherwise the equipotential hypersurface $\{S = \text{const}\}$ would not have a well-defined normal (corresponding to the existence of a caustic). We can therefore consider the coordinate transformation

$$\left.\begin{aligned} &X^0 = S(x^0,x^1,x^2,...,x^{n-1}) \\ &X^\alpha = x^\alpha \quad (\alpha = 1,2,...,n-1) \end{aligned}\right\} \Leftrightarrow \begin{cases} x^0 = \tau(X^0,X^1,X^2,...,X^{n-1}) \\ x^\alpha = X^\alpha \quad (\alpha = 1,2,...,n-1) \end{cases} \tag{1.5}$$

the right side of (1.5) being determined by the inverse function theorem. We now examine the (symmetric) metric $ds^2 = g_{ij}\,dX^i dX^j$ in the new coordinate system. Here $g_{ij} = \eta_{kl}\,(\partial x^k/\partial X^i)\,(\partial x^l/\partial X^j)$, and hence

$$ds^2 = -(\tau_i dX^i)^2 + \sum_{\alpha=1}^{n-1}(dX^\alpha)^2, \tag{1.6}$$

where $\tau_i \equiv \partial\tau/\partial X^i$. Moreover, the inverse of g_{ij} is

 0022-2488/83/010022-07$02.50

$g^{ij} = \eta^{kl}(\partial X^i/\partial x^k)(\partial X^j/\partial x^l)$ and hence

$$g^{ij}\frac{\partial}{\partial X^i}\frac{\partial}{\partial X^j} = -\frac{2\tau_\alpha}{\tau_0}\frac{\partial}{\partial X^0}\frac{\partial}{\partial X^\alpha} + \sum_{\alpha=1}^{n-1}\frac{\partial}{\partial X^\alpha}\frac{\partial}{\partial X^\alpha}. \tag{1.7}$$

The expression (1.7) follows either by direct inversion of g_{ij} from (1.6), or more simply from the facts that

$$g^{00} = \eta^{kl}\frac{\partial X^0}{\partial x^k}\frac{\partial X^0}{\partial x^l} = \nabla S\cdot\nabla S = 0$$

by (1.3a), and

$$g^{0\alpha} = \eta^{kl}\frac{\partial X^0}{\partial x^k}\frac{\partial X^\alpha}{\partial x^l} = \eta^{kl}\frac{\partial S}{\partial x^k}\delta^\alpha_l$$
$$= \eta^{k\alpha}\frac{\partial S}{\partial x^k} = \delta^k_\alpha\frac{\partial S}{\partial x^k} = \frac{\partial S}{\partial x^\alpha};$$

yet from (1.5), $0 = \partial x^0/\partial x^\alpha = \tau_0\,\partial S/\partial x^\alpha + \tau_\alpha$, implying $g^{0\alpha} = -\tau_\alpha/\tau_0$. Since $1 = \partial x^0/\partial x^0 = \tau_0\,\partial S/\partial x^0$, (1.3a) is equivalent to

$$\sum_{\alpha=1}^{n-1}\tau_\alpha^2 = 1 \tag{1.8}$$

and since $g_{ij} = \eta_{kl}(\partial x^k/\partial X^i)(\partial x^l/\partial X^j)$, we have

$$\det g_{ij} = (\det\eta_{kl})\left(\det\frac{\partial x^m}{\partial X^i}\right)^2 = -\tau_0^2. \tag{1.9}$$

We now reexpress Eqs. (1.3) with respect to the new coordinates X^i. Equation (1.3a) becomes

$$g^{ij}\frac{\partial S}{\partial X^i}\frac{\partial S}{\partial X^j} = 0 \Leftrightarrow g^{00} = 0,$$

which is identically satisfied (by virtue of the fact that the new coordinates have been adapted to this condition). For any quantity U we have

$$\Box U = \frac{1}{\sqrt{-\det g_{kl}}}\frac{\partial}{\partial X^i}\left(g^{ij}\sqrt{-\det g_{kl}}\,\frac{\partial U}{\partial X^j}\right)$$

or, using (1.9),

$$\Box U = \frac{1}{\tau_0}\sum_{\alpha=1}^{n-1}(\tau_0 U_{\alpha\alpha} - U_0\tau_{\alpha\alpha} - 2\tau_\alpha U_{\alpha 0}), \tag{1.10}$$

where a subscript i denotes partial differentiation with respect to X^i. It follows that Eq. (1.3c) is equivalent to

$$\sum_{\alpha=1}^{n-1}(\tau_0 U_{\alpha\alpha} - U_0\tau_{\alpha\alpha} - 2\tau_\alpha U_{\alpha 0}) = 0. \tag{1.11}$$

Substituting S for U in (1.10) and noting that $S_0 = 1$ and $S_\alpha = 0$ shows that Eq. (1.3b) is equivalent to

$$\sum_{\alpha=1}^{n-1}(2\tau_\alpha U_\alpha + U\tau_{\alpha\alpha}) = 0, \tag{1.12}$$

and, writing $V = U\tau_0$, this becomes

$$\sum_{\alpha=1}^{n-1}(2\tau_\alpha V_\alpha + V\tau_{\alpha\alpha}) = 0, \tag{1.13}$$

where use is made of the fact that $\Sigma_{\alpha=1}^{n-1}\tau_\alpha\tau_{0\alpha} = 0$, which follows from (1.8). Differentiating (1.12) with respect to X^0, and eliminating $\tau_{\alpha\alpha}$ between the resulting expression and Eq. (1.11), results in

$$\sum_{\alpha=1}^{n-1}(\tau_0 U_{\alpha\alpha} + 2\tau_{0\alpha}U_\alpha + U\tau_{0\alpha\alpha}) = 0,$$

which, with $V = U\tau_0$, shows that Eq. (1.3c) is equivalent to

$$\sum_{\alpha=1}^{n-1}V_{\alpha\alpha} = 0. \tag{1.14}$$

In summary, we have two functions V and τ which satisfy the equations

$$\sum_{\alpha=1}^{n-1}\tau_\alpha^2 = 1, \tag{1.8}$$

$$\sum_{\alpha=1}^{n-1}(2\tau_\alpha V_\alpha + V\tau_{\alpha\alpha}) = 0, \tag{1.13}$$

and

$$\sum_{\alpha=1}^{n-1}V_{\alpha\alpha} = 0, \tag{1.14}$$

where a subscript α denotes partial differentiation with respect to X^α ($\alpha = 1,2,\ldots,n-1$). These equations are of precisely the same form as those of the special subsystem (1.4), except that now the operators ∇ and ∇^2 refer to an *associated* $(n-1)$-dimensional Euclidean metric $(dX^1)^2 + (dX^2)^2 + \cdots + (dX^{n-1})^2$. Therefore, once the special subsystem (1.4) is solved, we have the solution in the general case. However, in making the appropriate transcription, it must be remembered that τ is a function of X^0, as well as of X^α ($\alpha = 1,2,\ldots,n-1$), and that $V = U\tau_0$.

This technique involving a change of coordinates is powerful, and it is natural to explore the extent to which we can apply it in order to obtain the most general solution of a system of equations from a very special solution. The above example, viz., system (1.3), involves three partial differential equations for two unknowns U and S. For simplicity, we will now consider instead systems involving two partial differential equations for one unknown ϕ. In Sec.II, we start by investigating "null field" solutions of the ordinary wave equation in 3 + 1 dimensions. These are solutions ϕ of the wave equation $\Box\phi = 0$, which also satisfy the equation $\nabla\phi\cdot\nabla\phi = 0$; they are of physical importance since they are linked to pure radiation fields and provide, in accordance with special relativity, the limiting case at which disturbances can propagate. In order to generalize the investigation, we then consider in Sec. III null field solutions of certain generalized wave equations in n-dimensional Minkowski space–time. Specifically, we prove the following

Theorem 1: Suppose that in n-dimensional Minkowski space–time ($n \geqslant 2$), in which there are coordinates (x^i) ($i = 0,1,2,\ldots,n-1$) such that the metric is

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + \cdots + (dx^{n-1})^2, \tag{1.15}$$

the partial differential equation

$$f(\phi,\phi_{;i},\phi_{;i;j}\phi^{;i;j}) = 0 \tag{1.16a}$$

admits null field solutions ϕ satisfying

$$\nabla\phi\cdot\nabla\phi \equiv \phi_{;i}\phi^{;i} = 0, \quad \nabla\phi \not\equiv 0. \tag{1.16b}$$

Further, suppose that the restriction (1.16a) is nontrivial and that the system (1.16) admits a time-transported null field solution of the form $\phi = x^0 - \tau(x^1,x^2,\ldots,x^{n-1})$ in some coordinate system (x^i) in which the metric is of form (1.15). If the system resulting from the coordinate transformation

$$\left.\begin{aligned} &X^0 = \phi\,(x^0,x^1,x^2,...,x^{n-1}) \\ &X^\alpha = x^\alpha \quad (\alpha = 1,2,...,n-1) \end{aligned}\right\}$$
$$\Leftrightarrow \begin{cases} x^0 = \tau(X^0,X^1,X^2,...,X^{n-1}) \\ x^\alpha = X^\alpha \quad (\alpha = 1,2,...,n-1) \end{cases}$$

in Euclidean $(n-1)$-dimensional space with metric $(dX^1)^2 + (dX^2)^2 + \cdots + (dX^{n-1})^2$ is explicitly independent of X^0, and is of precisely the same form as the system obtained from (1.16) with the substitution $\phi = x^0 - \tau(x^1,x^2,...,x^{n-1})$, in the Euclidean $(n-1)$-dimensional space with metric $(dx^1)^2 + (dx^2)^2 + \cdots + (dx^{n-1})^2$, then the partial differential equation (1.16a) is equivalent to the third-order quasilinear equation

$$D(\Box\phi) + K(\Box\phi)^2 = 0,$$

where K is a constant, and $D \equiv {}_{;i}\phi^{;i}$ denoted differentiation along the normals to the null hypersurfaces $\{\phi = \text{const}\}$. In this case, the general solution of the system (1.16) is obtainable from the most general time-transported solution.

In Sec. IV we apply our results to the construction of all shear-free hypersurface-orthogonal null geodesic congruences in Minkowski space–time. Various remarks are made in Sec. V, relating the results of the present work to those of previous articles, and concerning generalizations to the complexified case and to the case where the function f in (1.16a) depends not only on ϕ, $\phi_{;i}{}^{;i}$ and $\phi_{;i;j}\phi^{;i;j}$, but also on some covariantly constant vector field A^i.

Throughout, some familiarity with the geometric technique due to Friedlander[1,2] and extended by Collins[4,5] would be helpful. In this technique, a Gaussian coordinate system adapted to the equipotential surfaces is introduced, and curvature line parameters, related to the extrinsic curvature of the equipotential surfaces, are employed. Thus the entire description of the associated differential equation is in terms of coordinates which are geometrically significant.

II. NULL FIELD SOLUTIONS OF THE ORDINARY WAVE EQUATION

We first consider the concepts of a null field solution and of a time-transported solution. Suppose that we are dealing with the ordinary wave equation in 3 + 1 dimensional Minkowski space–time, i.e., $n = 4$ in Sec. I. A scalar function ϕ on a region of space–time which is not identically constant ($\nabla\phi \neq \mathbf{0}$) locally defines a system of hypersurfaces $\{\phi = \text{const}\}$, the normal at any point to which is (parallel to) $\nabla\phi$. Because of the indefinite metric, this normal vector, at any point, satisfies one of the conditions $\nabla\phi\cdot\nabla\phi > 0$, $\nabla\phi\cdot\nabla\phi < 0$, or $\nabla\phi\cdot\nabla\phi = 0$. If there is an open set in which $\nabla\phi\cdot\nabla\phi \equiv 0$ and $\nabla\phi \neq \mathbf{0}$, we call ϕ a *null field* (cf. Friedlander,[2] who has a different sign convention for the metric). In this case the null hypersurfaces $\{\phi = \text{const}\}$ are generated by a (unique) null geodesic congruence[2] (cf. Lemma 2.1 of Ref. 6). At any point, the normal to such a hypersurface is both orthogonal *and* tangent to the hypersurface, and tangential to a null geodesic in the generating congruence.

Given any Killing vector,[7] ξ, a null geodesic congruence with tangent vector $\mathbf{k}$ is *invariant under the action of* ξ if and only if the Lie derivative

$$\pounds_\xi \mathbf{k} = \mathbf{0} \Leftrightarrow [\xi,\mathbf{k}] = \mathbf{0} \Leftrightarrow k_{i;j}\xi^j - \xi_{i;j}k^j = 0 \qquad (2.1)$$

(cf. Ref. 6). We will say that a null-geodesic congruence is *time-transported* if it is invariant under the action of a time-like translational Killing vector. Such congruences were considered previously by Cox[8] and Collins,[6] by whom they were called "time-invariant." However, this latter nomenclature suggests time independence, and so here we prefer the terminology of time transportation. Following Ref. 6, if ϕ is a null field, $\phi_{;i}$ satisfies $\phi_{;i}\phi^{;i} = 0$, from which $\phi_{;i;j}\phi^{;i} = 0$, and hence $\phi_{;j;i}\phi^{;i} = 0$, i.e., the congruence tangent to $\phi_{;i}$ is a hypersurface-orthogonal affinely parametrized null geodesic congruence. If the congruence is time-transported, we say that the null field ϕ is *time-transported*. We can choose coordinates (x^i) in (1.1) such that $\xi = \partial/\partial x$ and, by (2.1), $\phi_{;i;j}\xi^j = 0$, which implies that there is a constant b and a real function $\tau(x^1,x^2,x^3)$ such that

$$\phi = bx^0 - \tau(x^1,x^2,x^3), \qquad (2.2)$$

and, since ϕ is null,

$$\nabla\tau\cdot\nabla\tau = b^2, \qquad (2.3)$$

where the operator ∇ refers to the metric induced on any hypersurface orthogonal to ξ, $ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2$. As in Sec. I, we may assume that $b = \partial\phi/\partial x^0 \neq 0$ (in order to obtain a well-defined normal to the hypersurface $\{\phi = \text{const}\}$), in which case

$$\bar\phi = \phi/b = x^0 - \bar\tau(x^1,x^2,x^3),$$

where $\bar\tau = \tau/b$. Thus, instead of ϕ we can consider $\bar\phi$ satisfying

$$\left.\begin{aligned} &\bar\phi_{;i}\bar\phi^{;i} = 0, \\ &\bar\phi = x^0 - \bar\tau(x^1,x^2,x^3), \end{aligned}\right\} \qquad (2.4)$$

i.e., without loss of generality, $b = 1$ in (2.2) and (2.3). In the following, we assume that (2.4) holds, and drop the barred notation. Note that time-transported solutions, satisfying (2.4), are *not* time-independent.

If ϕ is an arbitrary null field solution of the ordinary wave equation, then

$$\left.\begin{aligned} &\phi_{;i}\phi^{;i} = 0, \\ &\phi_{;i}{}^{;i} = 0. \end{aligned}\right\} \qquad (2.5)$$

If we were to consider special simple progressive solutions of the ordinary wave equation (1.2) of form $\phi = f(S)$, with f arbitrary and S not identically constant, i.e., if in Sec. I we have $U \equiv 1$, then equations (1.3) become

$$\left.\begin{aligned} &\nabla S\cdot\nabla S = 0, \\ &\Box S = 0, \end{aligned}\right\} \qquad (2.6)$$

which is equivalent to (2.5). The general solution of (2.6) is therefore obtainable first by considering the special time-transported solutions of form $S = t - \tau(x,y,z)$, where $(t,x,y,z) \equiv (x^0,x^1,x^2,x^3)$, so

$$\left.\begin{aligned} &\nabla\tau\cdot\nabla\tau = 1, \\ &\nabla^2\tau = 0, \end{aligned}\right\} \qquad (2.7)$$

where the operators ∇ and ∇^2 refer to the three-dimensional Euclidean space with metric $dx^2 + dy^2 + dz^2$, and then by performing the coordinate transformation technique de-

scribed in Sec. I. Now the general solution of (2.7) is[4]

$$\tau = lx + my + nz + A,$$

where l, m, n and A are constants satisfying $l^2 + m^2 + n^2 = 1$. Therefore, the general solution of (2.6) is given by solving [cf. (1.5), (1.8), (1.10), and (1.11) with $U \equiv 1$]

$$t = l(S)x + m(S)y + n(S)z + A(S)$$

for S, where $l(S), m(S), n(S)$, and $A(S)$ are *functions* satisfying $l^2 + m^2 + n^2 = 1$. Each hypersurface $\{S = \text{const}\}$ is a plane, but the orientation of the planes $\{S = \text{const}\}$ is the same only if the functions l, m, and n are constant. In the general case, the waves are plane-fronted, whereas in the case where l, m, and n are constants, the waves are plane-fronted with parallel rays, or *pp* waves.[9]

This result generalizes that of Ref. 6, in which only time-transported solutions were considered. It is itself extended to space–times of arbitrary dimension in Sec. V. Some brief comments on the global aspects of the result are made in Sec. IV.

III. PROOF OF THEOREM 1

We begin by invoking the change of coordinates (1.5). We note that putting $U = \phi = X^0$ in (1.10) yields

$$\Box\phi = \phi_{;i}{}^{;i} = \frac{-1}{\tau_0}\sum_{\alpha=1}^{n-1}\tau_{\alpha\alpha},$$

whereas $\phi_{;i;j}\phi^{;i;j} = (\phi_{;i;j}\phi^{;i})^{;j} - (\phi_{;i;j}{}^{;j})\phi^{;i} = -\phi_{;j}{}^{;j}{}_{;i}\phi^{;i}$ $= -D(\Box\phi)$, where $D \equiv {}_{;i}\phi^{;i}$. This latter result may be reexpressed in the coordinates (X^i) as

$$\begin{aligned}\phi_{;i;j}\phi^{;i;j} &= -(\Box\phi)_{;i}\phi_{;j}g^{ij} = -(\Box\phi)_{,\beta}g^{\beta 0}\\ &= -\sum_{\beta=1}^{n-1}\frac{\tau_\beta}{\tau_0}\frac{\partial}{\partial X^\beta}\left(\frac{1}{\tau_0}\sum_{\alpha=1}^{n-1}\tau_{\alpha\alpha}\right)\\ &= -\frac{1}{\tau_0}\sum_{\beta=1}^{n-1}\sum_{\alpha=1}^{n-1}\left(\frac{\tau_\beta\tau_{\alpha\alpha\beta}}{\tau_0} - \frac{\tau_\beta\tau_{0\beta}\tau_{\alpha\alpha}}{\tau_0^2}\right)\\ &= -\frac{1}{\tau_0^2}\nabla\tau\cdot\nabla(\nabla^2\tau),\end{aligned}$$

where the operators ∇ and ∇^2 refer to the Euclidean $(n-1)$-dimensional space with metric $(dX^1)^2 + (dX^2)^2 + \cdots + (dX^{n-1})^2$, and we have again used the fact that $\Sigma_{\alpha=1}^{n-1}\tau_\alpha\tau_{0\alpha} = 0$, as follows from (1.8). The partial differential equation $f = 0$, with $\phi_{;i}\phi^{;i} = 0$ in force, reduces to

$$f(\phi,\phi_{;i}{}^{;i},\phi_{;i;j}\phi^{;i;j}) = 0 \Leftrightarrow f\left(X^0, -\frac{\nabla^2\tau}{\tau_0}, -\frac{\nabla\tau\cdot\nabla(\nabla^2\tau)}{\tau_0^2}\right) = 0 \tag{3.1}$$

together with

$$\sum_{\alpha=1}^{n-1}\tau_\alpha^2 = 1,$$

which is equivalent to (1.16b). On the other hand, if we seek time-transported solutions with $\phi = x^0 - \tau(x^1,x^2,\ldots,x^{n-1})$, the equations (1.16) reduce to

$$f(x^0 - \tau(x^1,x^2,\ldots,x^{n-1}), -\nabla^2\tau, -\nabla\tau\cdot\nabla(\nabla^2\tau)) = 0, \tag{3.2}$$

together with $\nabla\tau\cdot\nabla\tau = 1$, where here the operators refer to the Euclidean $(n-1)$-dimensional space with metric $ds^2 = (dx^1)^2 + \cdots + (dx^{n-1})^2$. If Eq. (3.1) is to be independent of X^0, and of precisely the same form as (3.2), then, writing $f = f(u,v,w)$, we have $\partial f/\partial u = 0$ and writing $f = g(v,w)$, we obtain

$$g\left(-\frac{\nabla^2\tau}{\tau_0}, -\frac{\nabla\tau\cdot\nabla(\nabla^2\tau)}{\tau_0^2}\right) = 0 \Leftrightarrow g\left(-\nabla^2\tau, -\nabla\tau\cdot\nabla(\nabla^2\tau)\right) = 0$$

for all τ_0, and the partial differential equation (1.16a) is

$$g(\Box\phi, -D(\Box\phi)) = 0.$$

If $v \equiv \Box\phi \neq 0$, we write $y = w/v^2$ and $h(v,y) \equiv g(v,w)$, in which case

$$h\left(-\frac{\nabla^2\tau}{\tau_0}, -\frac{\nabla\tau\cdot\nabla(\nabla^2\tau)}{(\nabla^2\tau)^2}\right) = 0 \Leftrightarrow h\left(-\nabla^2\tau, -\frac{\nabla\tau\cdot\nabla(\nabla^2\tau)}{(\nabla^2\tau)^2}\right) = 0 \tag{3.3}$$

and the partial differential equation is

$$h\left(\Box\phi, -\frac{D(\Box\phi)}{(\Box\phi)^2}\right) = 0. \tag{3.4}$$

If h is independent of its second variable then, since by assumption f is nontrivial, (3.4) implies that $\Box\phi = c$, a constant. In that case (3.3) shows that $\nabla^2\tau = -c$, and so $\tau_{|\alpha|\beta}\tau^{|\alpha|\beta} = \tau_{|\alpha|\beta}\tau^{|\beta|\alpha} = (\tau_{|\alpha|\beta}\tau^{|\beta})^{|\alpha} - \tau_{|\alpha|\beta}{}^{|\alpha}\tau^{|\beta}$ $= -\nabla\tau\cdot\nabla(\nabla^2\tau) = 0$, where a vertical stroke (|) denotes covariant differentiation with respect to the metric induced on a hypersurface $\{x^0 = \text{const}\}$. Since this metric is positive-definite, we have $\tau_{|\alpha|\beta} = 0$, and, *a fortiori*, $\nabla^2\tau = \tau_{|\alpha}{}^{|\alpha} = 0$, i.e., $c = 0$, which contradicts the assumption that $\Box\phi \neq 0$. Thus, if $\Box\phi \neq 0$, Eq. (3.3) shows that $[\nabla\tau\cdot\nabla(\nabla^2\tau)/(\nabla^2\tau)^2]$ $= l(-\nabla^2\tau/\tau_0) = l(-\nabla^2\tau)$ for some function l, and for all τ_0, whence l is constant. Thus *either*

$$\Box\phi \equiv 0$$

or

$$D(\Box\phi) + K(\Box\phi)^2 = 0 \quad \text{with} \quad \Box\phi \neq 0,$$

where K is a constant. It is clear that $K \neq 0$, since otherwise $\nabla\tau\cdot\nabla(\nabla^2\tau) = 0$, which as we have seen, requires $\Box\phi \equiv 0$. We now combine the two possibilities, and Theorem 1 is proved. ■

IV. SHEAR-FREE HYPERSURFACE-ORTHOGONAL NULL GEODESIC CONGRUENCES IN MINKOWSKI SPACE–TIME

As previously shown,[6] a shear-free hypersurface-orthogonal null geodesic congruence in (four-dimensional) Minkowski space–time with metric

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$$

is tangential to a null vector $\mathbf{k}$ for which there exists a real function $g(t,x,y,z) \neq 0$ such that $k_i = g_{;i}$, $g_{;i}g^{;i} = 0$, $g_{;i;j}g^{;j} = 0$ and $(g_{;i}{}^{;i})^2 = 2g_{;i;j}g^{;i;j} = -2g^{;i}(g_{;j}{}^{;j})_{;i}$, i.e.,

$$\begin{aligned}&\nabla g\cdot\nabla g = g_{;i}g^{;i} = 0,\\ &(\Box g)^2 = -2\nabla g\cdot\nabla(\Box g).\end{aligned} \tag{4.1}$$

The general *time-transported* solution to (4.1) is, without loss of generality, of form $g = t - \tau(x,y,z)$ with

$$(\nabla\tau)^2 = 1,$$

$$(\nabla^2\tau)^2 = -2\nabla\tau\cdot\nabla(\nabla^2\tau), \tag{4.2}$$

where in (4.2) the operators ∇ and ∇^2 refer to the three-dimensional Euclidean space with metric $dx^2 + dy^2 + dz^2$, and whose solution is[6]:

(i) $\tau = lx + my + nz + A$, where l, m, n, and A are constants satisfying $l^2 + m^2 + n^2 = 1$ (the expansion scalar $g_{;i}{}^{;i} = -\nabla^2\tau$ of the congruence vanishes, and the hypersurfaces $\{g = \text{const}\}$ are null hyperplanes) or

(ii) $\tau = \pm[(x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2]^{1/2} + A$, when x_1, y_1, z_1, and A are constants (the expansion scalar $g_{;i}{}^{;i} = -\nabla^2\tau$ is $2/(\tau - A)$, and the hypersurfaces $\{g = \text{const}\}$ are null cones).

Now system (4.1) is a particular case of the class of systems referred to in Theorem 1, and therefore the *general* solution of (4.1) is given by solving for g:

$$\text{(i)}\quad t = l(g)x + m(g)y + n(g)z + A(g), \tag{4.3a}$$

where $l(g)$, $m(g)$, $n(g)$, and $A(g)$ are functions satisfying

$$l^2(g) + m^2(g) + n^2(g) = 1,$$

$$\text{(ii)}\quad t = \pm\{[x - x_1(g)]^2 + [y - y_1(g)]^2 + [z - z_1(g)]^2\}^{1/2} + A(g), \tag{4.3b}$$

where $x_1(g)$, $y_1(g)$, $z_1(g)$, and $A(g)$ are functions.

We therefore have

Theorem 2: The most general shear-free hypersurface-orthogonal null geodesic congruence in Minkowski space–time is generated by (the normals to) either light cones emanating from a single line $(t,x,y,z) = (t_1(g), x_1(g), y_1(g), z_1(g))$, or a system of null planes.

This theorem appears to be fairly well known (cf. the comments in Ref. 10, where it is incorrectly stated), but as far as I am aware there is no well-known standard reference to a proof. The theorem may be deduced as a corollary to Kerr's theorem,[8,11–15] which provides the form of the most general analytic shear-free null geodesic congruence in Minkowski space–time; some further comments relating to this can be found in Ref. 6. For a global application of the theorem, it is necessary that the equipotentials of g in (4.2) not intersect, which can only be achieved by requiring that the null planes all be parallel in (4.3a), i.e., that l, m, and n are constants, and by choosing either sign in (4.3b) and requiring that the curve $(t,x,y,z) = (t_1(g), x_1(g), y_1(g), z_1(g))$ be timelike (otherwise the equipotentials intersect or the set of null geodesics is not a congruence filling space–time, or both).

V. MISCELLANEOUS RESULTS

In this section we consider special cases and generalizations of Theorem 1.

(i) $\Box\phi = 0$; $n = 4$: The results of Sec. II follow immediately.

(ii) $\Box\phi = 0$; $n = 3$; This is a special case of (i) above, where we now write $(x^0,x^1,x^2) = (t,x,y)$, so

$$-\frac{\partial^2\phi}{\partial t^2} + \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0$$

and

$$-\left(\frac{\partial\phi}{\partial t}\right)^2 + \left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 = 0.$$

We may solve this by complexifying, writing $t = iz$. Then

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0$$

and

$$\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2 = 0.$$

This system was solved earlier,[4] using different methods. The general solution was given by $\phi \equiv$const or by solving

$$l(\phi)x + m(\phi)y + n(\phi)z = \psi(\phi),$$

where $l(\phi)$, $m(\phi)$, $n(\phi)$, and $\psi(\phi)$ are functions satisfying $l^2(\phi) + m^2(\phi) + n^2(\phi) = 0$, and l, m, and n are not all zero. Substituting back for t in favor of z, we have

$$l(\phi)x + m(\phi)y - in(\phi)t = \psi(\phi).$$

Suppose $n(\phi)\equiv 0$. Then $\partial\phi/\partial t = 0$ and $\phi = \phi(x \pm iy)$, i.e., ϕ is not real. If, however, $n(\phi)\not\equiv 0$, we obtain $t = \tilde{l}(\phi)x + \tilde{m}(\phi)y + A(\phi)$, where $\tilde{l} = -il/n$, $\tilde{m} = -im/n$, $A = i\psi/n$, and $\tilde{l}^2 + \tilde{m}^2 = 1$, and real solutions for $\phi = \phi(t,x,y)$ will exist. These solutions are in agreement with those of the special case of (i) above, when $n = 3$.

(iii) $\Box\phi \neq 0$; $n = 4$: We have

$$\nabla\phi\cdot\nabla(\Box\phi) + K(\Box\phi)^2 = 0,$$

$$\nabla\phi\cdot\nabla\phi = 0, \tag{5.1}$$

with $\Box\phi\not\equiv 0$. For a time-transported solution, without loss of generality of form $\phi = t - \tau(x,y,z)$, where $(t,x,y,z)\equiv(x^0,x^1,x^2,x^3)$, we have from (5.1)

$$\nabla\tau\cdot\nabla(\nabla^2\tau) + K(\nabla^2\tau)^2 = 0,$$

$$\nabla\tau\cdot\nabla\tau = 1$$

with $\nabla^2\tau\not\equiv 0$. If $K \neq 0$, the only (real) solutions to this are[5]

(a) $K = 1$, and τ is the a-eliminant of

$$\tau = l(a)x + m(a)y + n(a)z + A(a)$$

and

$$0 = l'(a)x + m'(a)y + n'(a)z + A'(a),$$

where $l^2(a) + m^2(a) + n^2(a) = 1$ and at most one of $l'(a)$, $m'(a)$, and $n'(a)$ is identically zero;

(b) $K = \frac{1}{2}$, and

$$\tau = \pm[(x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2]^{1/2} + A,$$

where x_1, y_1, z_1, and A are arbitrary real constants.

If $K = 0$, then it follows either from the discussion in Ref. 5 or from the proof in Sec. III that the equipotentials $\{\tau = \text{const}\}$ are planes, and that $\nabla^2\tau\equiv 0$, a contradiction. Therefore, the general solution of the system

$$D(\Box\phi) + K(\Box\phi)^2 = 0,$$

$$\nabla\phi\cdot\nabla\phi = 0$$

with $\Box\phi\not\equiv 0$ is given (when $n = 4$) by:

(a) $K = 1$, and ϕ is specified implicitly as the a-eliminant

of

$$t = l(a,\phi)x + m(a,\phi)y + n(a,\phi)z + A(a,\phi)$$

and

$$0 = \frac{\partial l}{\partial a}x + \frac{\partial m}{\partial a}y + \frac{\partial n}{\partial a}z + \frac{\partial A}{\partial a},$$

where $l^2(a,\phi) + m^2(a,\phi) + n^2(a,\phi) = 1$ and at most one of $\partial l/\partial a$, $\partial m/\partial a$, and $\partial n/\partial a$ is identically zero;

(b) $K = \frac{1}{2}$, and ϕ is specified implicitly by

$$t = \pm\{[x - x_1(\phi)]^2 + [y - y_1(\phi)]^2 + [z - z_1(\phi)]^2\}^{1/2} + A(\phi),$$

where $x_1(\phi)$, $y_1(\phi)$, $z_1(\phi)$, and $A(\phi)$ are arbitrary (real) functions.

The case (*b*) corresponds to the "expanding" solution (*ii*) of (4.2).

(*iv*) $n = 2$: We first show that $\Box\phi \equiv 0$. For a time-transported solution without loss of generality of form $\phi = t - \tau(x)$, where $(t,x) = (x^0,x^1)$, we have

$$\nabla\phi\cdot\nabla(\Box\phi) + K(\Box\phi)^2 = 0 \Rightarrow \frac{d\tau}{dx}\frac{d^3\tau}{dx^3} + K\left(\frac{d^2\tau}{dx^2}\right)^2 = 0 \qquad (5.2a)$$

and

$$\nabla\phi\cdot\nabla\phi = 0 \Rightarrow \left(\frac{d\tau}{dx}\right)^2 = 1. \qquad (5.2b)$$

Clearly (5.2a) is a consequence of (5.2b), and so $\tau = \pm x + x_0$, where x_0 is a constant; thus $\nabla^2\tau = d^2\tau/dx^2 = 0$ implies $\Box\phi \equiv 0$. Using Theorem 1, we have that the general solution in the case $n = 2$ is given by solving $t = \pm x + x_0(\phi)$ for ϕ, i.e., that $\phi = \phi(t \pm x)$, and we recover the well-known d'Alembert solution. This result can also be obtained by the coordinate substitution $u = t + x$, $v = t - x$ in the general two-dimensional problem.

(*v*) $n > 2$; $K \leqslant 1/(n-2)$ *or* $D(\Box\phi) = 0$; *kinematic quantities*: Our aim here is to show that under certain circumstances, viz., when $K \leqslant 1/(n-2)$ or when $D(\Box\phi) = 0$, it necessarily follows that $\Box\phi = 0$ and that the general null field solution is given by solving

$$x^0 = l_1(\phi)x^1 + l_2(\phi)x^2 + \cdots + l_{n-1}(\phi)x^{n-1} + A(\phi)$$

for ϕ, where $l_\alpha(\phi)$ $(\alpha = 1,2,\ldots,n-1)$ and $A(\phi)$ are functions satisfying $\Sigma_{\alpha=1}^{n-1} l_\alpha^2(\phi) = 1$. This generalizes the result of Sec. II.

We first employ a decomposition of the covariant derivative, analogous to that used in the pseudo-Riemannian manifolds of general relativistic cosmology.[6,16] Let $\tau = \tau(x^1,x^2,\ldots,x^{n-1})$ with $\nabla\tau\cdot\nabla\tau = 1$, so

$$\tau_{|\alpha|\beta} = \theta_{\alpha\beta} = \sigma_{\alpha\beta} + [1/(n-2)]\theta h_{\alpha\beta},$$

where $\theta_{\alpha\beta}$, $\sigma_{\alpha\beta}$, and θ are respectively interpreted as the "expansion tensor," the "shear tensor," and the "volume expansion scalar" of the congruence normal to the $(n-2)$-surfaces $\{x^0,\tau = \text{const}\}$. They satisfy the conditions $\theta_{\alpha\beta}\tau^{|\beta} = 0$, $\theta_{\alpha\beta} = \theta_{\beta\alpha}$, $\theta^\alpha{}_\alpha = \theta$, $\sigma_{\alpha\beta}\tau^{|\beta} = 0$, $\sigma_{\alpha\beta} = \sigma_{\beta\alpha}$, and $\sigma^\alpha{}_\alpha = 0$. The tensor $h_{\alpha\beta}$ is the "projection tensor" into the tangent plane at each point on a $(n-2)$-surface, i.e., $h_{\alpha\beta} = \bar{g}_{\alpha\beta} - \tau_{|\alpha}\tau_{|\beta}$, $h_{\alpha\beta} = h_{\beta\alpha}$, $h^\alpha{}_\alpha = n - 2$, and $h_{\alpha\beta}h^{\alpha\gamma} = \delta^\gamma_\beta$, where $\bar{g}_{\alpha\beta}$ is the metric induced on a hypersurface $\{x^0 = \text{const}\}$. We define also the shear scalar σ by $2\sigma^2 = \sigma_{\alpha\beta}\sigma^{\alpha\beta}$, $\sigma \geqslant 0$, and note that $\sigma_{\alpha\beta} = 0 \Leftrightarrow \sigma = 0$. In analogy with the situation in general relativity,[16] we may derive equations which specify the propagations of θ and $\sigma_{\alpha\beta}$ along the normal congruence. Thus $\dot{\theta}_{\alpha\beta} \equiv \theta_{\alpha\beta|\gamma}\tau^{|\gamma} = \tau_{|\alpha|\beta|\gamma}\tau^{|\gamma}$ $= \tau_{|\gamma|\alpha|\beta}\tau^{|\gamma} = (\tau_{|\gamma|\alpha}\tau^{|\gamma})_{|\beta} - \tau_{|\gamma|\alpha}\tau^{|\gamma}{}_{|\beta}$. Since $\tau_{|\gamma}\tau^{|\gamma} = 1$, it follows that $\dot{\theta}_{\alpha\beta} = -\theta_{\alpha\beta}\theta^\alpha{}_\beta$ and hence

$$\dot{\theta} = -\theta_{\alpha\beta}\theta^{\alpha\beta} = -\{[1/(n-2)]\theta^2 + 2\sigma^2\}. \qquad (5.3a)$$

Also $\dot{h}_{\alpha\beta} = h_{\alpha\beta|\gamma}\tau^{|\gamma} = (\bar{g}_{\alpha\beta} - \tau_{|\alpha}\tau_{|\beta})_{|\gamma}\tau^{|\gamma} = 0$, so it follows that $\dot{\sigma}_{\alpha\beta} = \dot{\theta}_{\alpha\beta} - [1/(n-2)]\dot{\theta}h_{\alpha\beta}$, i.e.,

$$\dot{\sigma}_{\alpha\beta} = -\sigma_{\alpha\gamma}\sigma^\gamma{}_\beta - \frac{2}{n-2}\theta\sigma_{\alpha\beta} + \frac{2\sigma^2}{n-2}h_{\alpha\beta},$$

from which

$$(\sigma^2)^{\cdot} = \dot{\sigma}_{\alpha\beta}\sigma^{\alpha\beta} = -\sigma_{\alpha\gamma}\sigma^\gamma{}_\beta\sigma^{\alpha\beta} - [4/(n-2)]\theta\sigma^2. \qquad (5.3b)$$

If we seek null field solutions of the equation $D(\Box\phi) + K(\Box\phi)^2 = 0$, then, following the procedure of Theorem 1, we substitute $\phi = x^0 - \tau(x^1,x^2,\ldots,x^{n-1})$, to obtain

$$\dot{\theta} + K\theta^2 = 0. \qquad (5.4)$$

Combining this with (5.3a), we obtain

$$[K - 1/(n-2)]\theta^2 - 2\sigma^2 = 0, \qquad (5.5)$$

from which we may conclude that if $K \leqslant 1/(n-2)$, then $\sigma = \theta = 0$. Similarly, it follows from (5.3a) that if $D(\Box\phi) = 0$, then $\sigma = \theta = 0$. Thus, if either $D(\Box\phi) = 0$ or $K \leqslant 1/(n-2)$, we have $\sigma = \theta = 0$, so $\tau_{|\alpha|\beta} = 0$ and hence $\tau = l_1x^1 + l_2x^2 + \cdots + l_{n-1}x^{n-1} + A$, where l_α $(\alpha = 1,2,\ldots,n-1)$ and A are constants satisfying $\Sigma_{\alpha=1}^{n-1} l_\alpha^2 = 1$. Applying the procedure of Theorem 1, it follows that the general solution is given by solving

$$x^0 = l_1(\phi)x^1 + l_2(\phi)x^2 + \cdots + l_{n-1}(\phi)x^{n-1} + A(\phi)$$

for ϕ, where $l_\alpha(\phi)$ $(\alpha = 1,2,\ldots,n-1)$ and $A(\phi)$ are arbitrary functions satisfying $\Sigma_{\alpha=1}^{n-1} l_\alpha^2(\phi) = 1$.

It is of interest to note also that in the special case $n = 4$ [without the restrictions $K \leqslant 1/(n-2)$ or $D(\Box\phi) = 0$ in force], $\sigma_{\alpha\gamma}\sigma^\gamma{}_\beta\sigma^{\alpha\beta}$ is identically zero, so Eq. (5.3b) simplifies, and the propagation of (5.5) requires either $\theta \equiv 0$ or $\theta \neq 0$ and

$$(K - \tfrac{1}{2})K\theta^2 - 2\sigma^2 = 0,$$

where use is made of (5.3b) and (5.4). Using (5.5), it follows that either $\theta \equiv \sigma \equiv 0$ or $\theta \neq 0$, $\sigma \equiv 0$, and $K = \frac{1}{2}$, or $\sigma \equiv \frac{1}{4}|\theta| \neq 0$ and $K = 1$. These situations were discussed in Sec. II and in case (iii) above.

(*vi*) *Generalization to include a covariantly constant vector field*: Suppose that the function f in Theorem 1 is allowed to depend on some covariantly constant vector field A^i. The simplest such dependence would involve f being a function not only of ϕ, $\phi_{,i}{}^{,i}$, and $\phi_{,i;j}\phi^{,i;j}$, but also of $\phi_{,i}A^i$ and A_iA^i. However, since $A_{i;j} = 0$, $(A_iA^i)_{;j} = 0$; in other words, A_iA^i is constant. Hence we shall suppose that f is a function of ϕ, $\phi_{,i}{}^{,i}$, $\phi_{,i;j}\phi^{,i;j}$, and $\phi_{,i}A^i$. Let $\mathbf{A} = a^i\partial/\partial X^i = b^i\partial/\partial x^i$, where

X^i and x^i are coordinates as in Sec. I. Then $a^i = b^{\ j}\partial X^i/\partial x^j$, $b^{\ i} = a^j\,\partial x^i/\partial X^{\ j}$, and $\phi_{;i}A^{\ i} = a^0$. Thus, if $f = f(\phi,\phi_{;i}{}^{;i}$, $\phi_{;i;j}\phi^{\ ;i;j}$, $\phi_{;i}A^{\ i})$ with $\partial f/\partial(\phi_{;i}A^{\ i}) \neq 0$, then the technique of coordinate transformations discussed in Sec. I will be valid provided

$$f\left(X^0, -\frac{\nabla^2\tau}{\tau_0}, -\frac{\nabla\tau\cdot\nabla(\nabla^2\tau)}{\tau_0^2}, a^0\right) = 0$$

$$\Leftrightarrow f(x^0 - \tau, -\nabla^2\tau, -\nabla\tau\cdot\nabla(\nabla^2\tau), a^0\tau_0) = 0 \qquad (5.6)$$

for arbitrary τ_0 and for the expressions in (5.6) to be explicitly independent of X^0 (here we have used the fact that if $\phi = t - \tau$,

$$\phi_{;i}A^{\ i} = b^{\ 0} - b^{\ \alpha}\partial\tau/\partial x^\alpha$$

$$= a^j\partial\tau/\partial X^{\ j} - a^\alpha\partial\tau/\partial X^{\ \alpha} = a^0\tau_0).$$

Arguing as in the proof of Theorem 1, it follows that either $\Box\phi \equiv 0$, or that f is independent of its third argument, or that (5.6) gives

$$\frac{\nabla\tau\cdot\nabla(\nabla^2\tau)}{(\nabla^2\tau)^2} = l\left(\frac{-\nabla^2\tau}{\tau_0}, a^0\right) = l(-\nabla^2\tau, a^0\tau_0)$$

for all τ_0. Writing $w = u/v$ for $v \neq 0$, we define $m(w,v) = l(u,v)$, so $m(-\nabla^2\tau/a^0\tau_0, a^0) = m(-\nabla^2\tau/a^0\tau_0, a^0\tau_0)$ for all τ_0 (provided $a^0 \neq 0$), and hence $m = m(-\nabla^2\tau/a^0\tau_0)$. In this case the original partial differential equation is of form

$$\frac{\nabla\phi\cdot\nabla(\Box\phi)}{(\Box\phi)^2} = m\left(\frac{\Box\phi}{\phi_{;i}A^{\ i}}\right),$$

where $m = m(w)$ is an arbitrary function. If, however, $a^0 = 0$, then

$$\phi_{;i}A^{\ i} = 0$$

(which has null field solutions only if $A^{\ i}A_i \geqslant 0$, i.e., $A^{\ i}$ is not timelike). Finally, if f is independent of its third argument, we must have $\nabla^2\tau/a^0\tau_0 = \text{const}$, and the original partial differential equation is of form

$$\Box\phi + \mathbf{A}\cdot\nabla\phi = 0.$$

This last case is of considerable interest, since it belongs to a particular class of scalar wave equations considered by Friedlander[2] in his book.

ACKNOWLEDGMENTS

This work was partially supported by an operating grant from the Natural Sciences and Engineering Research Council of Canada. I am grateful to the referee for several pertinent comments, which helped improve the exposition of this article, and to Helen Warren for her careful typing of the manuscript.

[1]F. G. Friedlander, Proc. Cambridge Phil. Soc. **43**, 360 (1946).
[2]F. G. Friedlander, *The Wave Equation on a Curved Space–Time* (Cambridge U.P., Cambridge, 1975).
[3]See, e.g., D. Lovelock and H. Rund, *Tensors, Differential Forms, and Variational Principles* (Wiley, New York, 1975).
[4]C. B. Collins, Math. Proc. Cambridge Phil. Soc. **80**, 165 (1976).
[5]C. B. Collins, J. Math. Phys. **21**, 240 (1980).
[6]C. B. Collins, J. Math. Phys. **21**, 249 (1980).
[7]See Ref. 3. A Killing vector generates a 1-parameter group of isometries of space–time. In Minkowski space–time the most general Killing vector generates a combination of translations, rotations, and Lorentz boosts.
[8]D. Cox, J. Math. Phys. **18**, 1188 (1977).
[9]J. Ehlers and W. Kundt, "Exact Solutions of the Gravitational Field Equations," in *Gravitation: An Introduction to Current Research*, edited by L.Witten (Wiley, New York, 1962).
[10]E. T. Newman and J. Winicour, J. Math. Phys. **15**, 426 (1974).
[11]R. Penrose, J. Math. Phys. **8**, 345 (1967).
[12]R. Penrose, Int. J. Theor. Phys. **1**, 61 (1968).
[13]G. C. Debney, R. P. Kerr, and A. Schild, J. Math. Phys. **10**, 1842 (1969).
[14]R. O. Hansen and E. T. Newman, Gen. Relativ. Grav. **6**, 361 (1975).
[15]D. Cox and E. J. Flaherty, Jr., Commun. Math. Phys. **47**, 75 (1976).
[16]G. F. R. Ellis, "Relativistic Cosmology," in *General Relativity and Cosmology*, Proceedings of the International School of Physics "Enrico Fermi" Course XLVII, 1969, edited by R. K. Sachs (Academic, London and New York, 1971), p. 104.

Deformation of Lie group representations and linear filters, Part I[a)]

Robert Hermann
Association for Physical and Systems Mathematics, 53 Jordan Road, Brookline, Massachusetts 02146

(Received 3 June 1982; accepted for publication 20 August 1982)

"Deformation theory" is a branch of mathematics which studies the geometry of dependence on *parameters* of geometric and physical systems. Material arising from Lie group theory and mathematical physics (e.g., in the study of asymptotic behavior of angular momentum) is applied to study the asymptotic behavior of certain linear filters depending on parameters. A mathematical machine which unifies many of these problems will be developed in this series.

PACS numbers: 02.30. + g, 02.10.Sp, 84.20.Ma, 84.30.Vn

1. INTRODUCTION

This paper and those to follow are a sequel to Ref. 1. There I showed that the theory of linear input-output systems and filters is closely linked to certain aspects of harmonic analysis and the Lie group theoretical explanation of the properties of the Special Functions of mathematical physics.[2,3] Now, these Special Functions often come with parameters naturally attached. It is known from earlier work[4,5] that certain phenomena involving the parameters involves what is called in the mathematical literature the theory of *deformations* of Lie groups and their linear representations. The theory of linear input-output systems depending on parameters has also been developed in the last ten years.[6-8] It is the purpose of this paper to bring these two streams together, and apply them to some relatively concrete problems and formulas involving the asymptotic formulas for the Special Functions of mathematical physics.

Much of the work in this paper will be motivated by *one* example, the following formula in Whittaker and Watson (Ref. 9, p. 367):

$$J_0(t) = \lim_{n\to\infty} P_n\left(\frac{t}{n}\right), \tag{1.1}$$

where $t\to J_0(t)$, $x\to P_n(x)$ are the usual Bessel functions and Legendre polynomials. On the mathematical physics-Lie group theory side, it is known that the right-hand side of (1.1) (for finite n) is the matrix element of a one-parameter subgroup of the rotation group SO(3,R) in the spin n-representation, while the left-hand side is the matrix element of a one-parameter subgroup of a semidirect product of SO(2,R) and a two-dimensional abelian subgroup, a solvable Lie group which is isomorphic to the group of rigid motions of R^2. Thus (1.1) represents in a concrete formula the whole geometric process of deformation of SO(3,R) and its representations over to the group of rigid motions in R^2. [This asymptotic formula is also a key example in the Inonu–Wigner theory[10] of "contractions" of Lie groups. The relation between the Inonu–Wigner theory and the theory of deformation of Lie groups and algebra is discussed in Ref. 4.]

System theoretically, this involves a family of linear time-invariant scalar input-output systems, parameterized by the integer n, $n = 0,1,\cdots$. The left-hand side of (1.1) is the *impulse response* (or *kernel of the linear filter determining the input-output relations*) of a system with infinite dimensional state-space.[1] The right-hand side, for finite n, is the impulse response of such a system with minimal state dimension $2n + 1$, i.e., the dimension of the spin n representation of SO(3,R). We shall see that there are important issues here in the theory of linear input-output systems, particularly the question of limit of sequences of finite and increasing state dimension systems, and that of approximation of systems with "large" (possibly infinite) state dimensions by systems with "small" state dimensions.

An approach of M. Hazewinkel[6] gives us a useful mathematical framework to think about this area of approximation and limits of input-output systems. It is also useful to take the Laplace transform of both sides of (1.1), the result is that in the "frequency domain" the "transfer functions" (i.e., Laplace transform of the impulse response, or "symbol" in the appropriate pseudodifferential operator sense) will not converge *pointwise*, but will have some suitable "asymptotic" relation. It seems appropriate to look for the *geometric* nature of these limiting relations in the work done by Martin and myself[11] on the geometric interpretation of the "transfer function" as a complex-analytic curve in a Grassman manifold.

The formulas for Laplace transform of both sides of (1.1) are as follows[12]:

$$\int_0^\infty P_{2n}(\cos t)\, e^{-st}\, dt = \frac{N_n(s)}{D_n(s)}, \tag{1.2}$$

with numerator and denominator polynomials as follows:

$$N_n(s) = (s^2 + 1)(s^2 + 9)\cdots(s^2 + (2n-1)^2), \tag{1.3}$$

$$D_n(s) = s(s^2 + 4)(s^2 + 16)\cdots(s^2 + (2n)^2), \tag{1.4}$$

$$\int_0^\infty e^{-st} J_0(t)\, dt = \frac{1}{(s^2+1)^{1/2}}. \tag{1.5}$$

The question arises of the relation between formulas (1.2)–(1.5) as $n\to\infty$. It is seen that these are *Padé approximations*. Thus, we see new and unexpected relations arise between different parts of mathematics, motivated by certain areas of applications.

My aim in this paper is to develop a broader explanation in terms of what geometers call *deformation theory* for this type of asymptotic formulas. There is also an innovative mathematical feature involved in the work here. In the classical literature, limit formulas of type (1.1) are proved by

[a)] Supported by a grant from the Ames Research Center (NASA), #NSG2402, from the Army Research Office, #ILIG1102RHN7-05MATH, and from the National Science Foundation, Grant No. MCS-8201779.

 0022-2488/83/010029-04$02.50

residue calculus or estimates of terms of power series. The techniques used here (and in previous work[4,5]) involves the Lebesgue dominated convergence theory applied to one-parameter groups of diffeomorphisms (often even gradient flows) acting on manifolds. On the applied front, I believe that techniques will be useful in broader areas of applications of systems/filters.

I would like to thank M. Hazewinkel and G. Zames for many conversations about this cross-disciplinary material. In particular, I note that recent work by Zames[13,14] is closely related to this material, but uses a different mathematical formalism, namely, the theory of the Hardy H_p spaces. I expect that investigation of the relation between the Lie group deformation theory and approximation in the Hardy sense will be a fruitful field of mathematical investigation.

2. THE DEFORMATION OF THE LEGENDRE INTO THE BESSEL FUNCTIONS

As preparation for a more general setting, let us examine what is involved geometrically in formula (1.1). Consider the classical Laplacian integral formula: for nth degree Legendre polynomials with n an integer

$$P_{nx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} (x + i(1-x^2)^{1/2}\cos\theta)^n\, d\theta$$
$$x \in \mathbb{C}, \quad n = 0,1,2,\cdots. \tag{2.1}$$

Let us convert this explicitly into an integral over the unit circle S^1 in R^2, that we will parameterize by $z = e^{i\theta} \in \mathbb{C}$. Let σ denote a point of S^1 as an abstract real-analytic manifold (so that "θ" is a real-analytic function on S^1), and let "$d\sigma$" be the volume-element differential form-measure on S^1 of total volume 1, which is invariant under the action of rotations. Thus

$$\int_{S^1} f\, d\sigma = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\, d\theta, \tag{2.2}$$

when $\sigma \to f(\sigma)$ is a measurable function on S^1. We have

$$P_n(x) = \int_{S^1} (x + \tfrac{1}{2}(x^2-1)^{1/2}(z+z^{-1})^n)\, d\sigma, \tag{2.3}$$

where "z" denotes the complex valued function $\theta \to e^{i\theta} = z(\theta)$ on S^1.

Notice that we can also write (2.3) as the integral over S^1 of a meromorphic one-differential form in $\mathbb{C}^2$, the space of complex variables (λ, z). Set

$$x = \tfrac{1}{2}(\lambda + \lambda^{-1}), \tag{2.4}$$
$$x^2 - = \tfrac{1}{4}(\lambda^2 + 2 + \lambda^{-2}) - 1$$
$$= \tfrac{1}{4}(\lambda^2 - 2 + \lambda^{-2})$$
$$= \tfrac{1}{4}(\lambda - \lambda^{-1})^2$$

or

$$(x^2-1)^{1/2} = \tfrac{1}{2}(\lambda - \lambda^{-1}).$$

Set

$$\omega = (i/4i)\,[(\lambda + \lambda^{-1}) + (\lambda - \lambda^{-1})(z + z^{-1})]^n\, z^{-1}\, dz. \tag{2.5}$$

Theorem 2.1: $P_n(x)$, the value of the nth Legendre polynomial at $x \in \mathbb{C}$, is equal to the integral over the curve $z = e^{i\theta}$ of the meromorphic one-form ω on $\mathbb{C}\times\mathbb{C}$ given by formula (2.5), with λ and x related via the linear fractional transformation (2.4). The polynomial dependence on x results (geometrically) from the fact that ω is a *rational* differential form, and that the integral can be evaluated as a residue at the singularity $z = 0$, which is rationally related to x via (2.4). We also have the following formula in terms of natural "intrinsic" geometry on S^1:

$$P_n(x) = \int_{S^1} \tfrac{1}{2}\,[(\lambda + \lambda^{-1}) + \tfrac{1}{2}(\lambda - \lambda^{-1})(z + z^{-1})]^n\, d\sigma \tag{2.6}$$

with λ and x again related via (2.4).

Now we are prepared to discuss, geometrically, the limit (1.1).

$$P_n\left(\cos\frac{t}{n}\right)$$
$$= \int_{S^1}\left[\cos\frac{t}{n} + \frac{i}{2}\sin\frac{t}{n}(z+z^{-1})\right]^n d\sigma$$
$$= \int_{S^1}\left[1 + \frac{i}{2}\left(\frac{\sin(t/n)}{t/n}\right)t\left(\frac{z+z^{-1}}{n}\right)\right.$$
$$\left.\times\left(\cos\left(\frac{t}{n}\right)\right)^{-1}\right]^n\left[\cos\frac{t}{n}\right]^n d\sigma. \tag{2.7}$$

Abstractly, we have a sequence of C^∞ functions,

$$f_n(t, z),$$

defined on $\mathbb{C}\times S^1$, such that

$$\lim_{n\to\infty} f_n(t,z) = \exp\left(z\frac{i(z+z^{-1})}{2}\right)$$
$$\text{for each} \quad t\in\mathbb{C},\ z\in S^1. \tag{2.8}$$

The right-hand side of (2.6) is an integrable function on S^1. The Lebesgue dominated convergence theorem[15] then implies the following result:

Theorem 2.2: For each $t \in \mathbb{C}$,

$$\lim_{n\to\infty} P_n\left(\cos\left(\frac{t}{n}\right)\right) = \int_{S^1}\exp\left(\frac{it}{2}\cos\theta\right)d\sigma. \tag{2.9}$$

The right-hand side of (2.9) is the Laplace formula for the Bessel function, hence this formula is equivalent to (1.1).

3. HAZEWINKEL'S DEFINITION OF "LIMIT" OF A SEQUENCE OF LINEAR FILTERS

Let R_+ denote the additive semigroup of nonnegative real numbers parameterized by $t \in R$, $t \geqslant 0$. Let $C(R_+)$ or $C(0,\infty)$ denote the space of complex-valued, continuous functions on R_+. A (scalar, input-output, time-invariant) *linear filter* is a linear map: $F: C(R_+) \to C(R_+)$ of the following form:

$$F(\underline{u})(t) = \int_0^t f(t-\tau)\,\underline{u}(\tau)\, d\tau \quad \text{for } \underline{u}\in C(R_+), \tag{3.1}$$

where the Lebesgue measurable function $f: R^+ \to \mathbb{C}$ satisfies the following condition:

$$\int_a^b |f(t)|\, dt < \infty \quad \text{for all } a,b \in R^+. \tag{3.2}$$

f is the *impulse response* of the filter. This can be written algebraically as

$$F(\underline{u}) = f * u, \tag{3.3}$$

where $*$ is the *causal convolution*[1] on $C(R_+)$, studied by Titchmarsh and Mikusinski.[15,16]

In linear system theory, one often encounters families of such filters depending on parameters and looks for natural ways of defining limits of such systems as the parameters vary. The standard functional analysis[15] methods for defining topologies on linear operators do not seem completely satisfactory, for reasons I will not go into here. Instead, we will use an approach which is a hybrid of the classical and modern techniques[9,15] suggested by Hazewinkel.[16] One first of all provides a linear subspace U of inputs, i.e., a linear subspace of $C(R_+)$. Then, one imposes a family of exponentially weighted sup-norms on $C(R_+)$, and requires that a sequence $F_0, F_1, F_2, \cdots$ of filters converges to F if, for each $\underline{u} \in U$, the outputs

$$y_n = F_n(\underline{u}), \quad n = 0,1,2,$$

converges in *all* the family of norms. For details, refer to Ref. 6.

4. LINEAR SYSTEMS DEFINED BY INTEGRATION

There is another feature of Sec. 2 that is worthwhile defining in general—the way the kernels of the linear filters $t \to P_n(\cos(t/n))$ and $t \to J_0(t)$ are defined by *integration* over a measure space Z. (In the case of Sec. 2, the Z is the unit circle in R^2, with the measure just *Lebesgue measure*, or, if it is identified with the Lie group SO(2,R), just the *Haar measure*.

Let (Z, dz) be a space with a countably additive field of measurable sets[15] and a countably additive measure dz defined on this field.[15] Impose the usual Lebesgue measure on R_+, defined on the Borel sets. Let

$$k: R_+ \times Z \to \mathbb{C},$$

be a map which is measurable with respect to the product measure on $R_+ \times Z$ and Lebesgue measure on Z such that the following condition is satisfied:

$$\iint_{[a,b]\times Z} |k(t,z)|\, dt\, dz < \infty$$

$$\text{for each finite interval } [a,b] \subset R_+ . \tag{4.1}$$

The Fubini theorem on product measures[15] then guarantees that the following formula

$$f(t) = \int_Z k(t,z)\, dz, \tag{4.2}$$

defines a map: $R_+ \to \mathbb{C}$, which is defined for all but a set of measure zero in R_+. Further, f is *locally integrable*, in the sense that

$$\int_a^b |f(t)|\, dt < \infty \quad \text{for all } a,b \in R_+ . \tag{4.3}$$

Let us use the function defined by formula (4.2) to define a linear filter map $C(R^+) \to C(R^+)$

$$F(\underline{u})(t) = \int_0^t u(t-\tau) f(\tau)\, d\tau$$

$$\text{for } \underline{u}: t \to u(t) \text{ an element of } C(R^+). \tag{4.4}$$

Use (4.2) and the Fubini theorem again to write the filter as follows:

$$\begin{aligned} F(\underline{u})(t) &= \int_0^t u(t-\tau) \int_Z k(\tau, z)\, dz\, d\tau \\ &= \int_0^t \int_Z u(t-\tau)\, k(\tau, z)\, d\tau\, dz \\ &= \int_0^t \int_Z u(\tau)\, k(t-\tau, z)\, d\tau\, dz . \end{aligned} \tag{4.5}$$

We can also estimate the exponentially weighted sup-norms used by Hazewinkel[6]:

$$e^{-bt} |\underline{u}(t)| \leq \int_0^t \int_Z e^{-bt} |u(\tau)\, k(t-\tau)|\, d\tau\, dz . \tag{4.6}$$

5. CONVERGENCE, IN THE HAZEWINKEL TOPOLOGY, OF LINEAR FILTERS DEFINED BY INTEGRATION

Now, let the kernel functions k, and the linear filters they determine, as described in Sec. 4, depend on parameter. For simplicity, in this paper the only parameter we will consider will be the integers $n = 0,1,2, \cdots$.

Z is a measure space with measure dz. Let

$$k_n: R_+ \times Z \to \mathbb{C}$$

be a sequence of measurable kernel functions which satisfy the condition (4.1), hence define, for each n a linear filter

$$F_n(\underline{u}) = \int_0^t \int_Z u(t-\tau)\, k_n(\tau, z)\, d\tau\, dz . \tag{5.1}$$

Let

$$F_\infty(\underline{u}) = \int_0^t \int_Z u(t-\tau)\, k_\infty(\tau, z)\, d\tau\, dz \tag{5.2}$$

be another linear filter with similar properties.

To apply Hazewinkel's ideas[6] and discuss when it may be considered that

$$\lim_{n\to\infty} F_n = F_\infty , \tag{5.3}$$

we are interested in sufficient conditions for the integral on the right-hand side of (5.1) to converge, as $n \to \infty$, for fixed $t \to u(t)$ to (5.2). This can be done, given our hypotheses, by the Lebesgue dominated convergence theorem.

Theorem 5.1: If the following conditions are satisfied:

$$\lim_{n\to\infty} k_n(t,k) = k_\infty(t,z) \tag{5.4}$$

$$\text{for almost all } (t,z) \in R_+ \times Z,$$

$$\int_a^b |k_\infty(t,z)|\, dt\, dz < \infty , \tag{5.5}$$

$$\int_a^b |u(t)|\, dt < \infty \quad \text{for } a,b \in R_+ , \tag{5.6}$$

then,

$$\lim_{n\to\infty} |F_n(\underline{u})(t) - F_\infty(\underline{u})(t)| = 0 \quad \text{for all } t \in R.$$

Our strategy now is to specialize Z to be a manifold with the measures "dz" defined by smooth differential forms, and the kernels $k(\ ,\ )$ associated with Lie algebras of differential operators on Z. In this paper, we will only consider the case

$$Z = S^1, \quad \text{the unit circle in } R^2, \tag{5.7}$$

and the Lie algebra that the representation (depending on a parameter) of the Lie algebra of SO(3,R). As the values of the parameters go to infinity, we shall be able to apply the limit theorems sketched in this section to obtain "degeneration" of these linear filters to those associated with a "contraction" (in the Inonu–Wigner sense[10]) of SO(3,R) to the group of rigid motions in R^2.

6. LINEAR SYSTEMS DEFINED BY REPRESENTATIONS OF THE LIE ALGEBRA OF SL(2,C) BY ONE-VARIABLE DIFFERENTIAL OPERATORS

As explained in the end of Sec. 5, we are motivated to choose

$$Z = S^1$$

and the linear filters whose kernel is of the form

$$f(t) = \int_Z h(z) \exp(tD)(f)(z)\, dz, \tag{6.1}$$

where f, h are functions: $Z \to \mathbb{C}$, and D is a first order differential operator associated with the representation of the Lie algebra of SO(3,C). The formulas for this situation have been worked out in Ref. 5.

Let S^1 be the unit circle in $\mathbb{C}$ with parameter θ, i.e.,

$$\theta \to z = e^{i\theta}$$

is the embedding map from $S^1 \to \mathbb{C}$. Consider the following differential operator:

$$A_1 = \frac{d}{d\theta} = iz^{-1}\frac{d}{dz}, \tag{6.2}$$

$$A = i\sin\theta\frac{d}{d\theta} + \alpha i\cos\theta \tag{6.3}$$

$$= -\frac{i}{2}(z - z^{-1})z^{-1}\frac{d}{dz} + i\frac{\alpha}{2}(z + z^{-1}), \tag{6.4}$$

$$A_2 = [A_1, A] \tag{6.5}$$

$$= i\cos\theta\frac{d}{d\theta} - i\alpha\sin\theta$$

$$= -\frac{i}{2}(z - z^{-1})z^{-1}\frac{d}{dz} + \frac{\alpha}{2}(z - z^{-1}). \tag{6.6}$$

Then, (A, A_1, A_2) satisfy (for fixed α), the commutation relations of the Lie algebra $\mathcal{G}$ of the Lie group $G = \mathrm{SO}(3,R)$.

For each $\alpha \in \mathbb{C}$, these families define a representation of $\mathcal{G}$ by linear maps on the $C^\infty(S^1)$, the C^∞, complex-valued functions on S^1. Let $D^1(S^1)$ be the Lie algebra of first order linear differential operators on S^1, considered as acting on $C^\infty(S^1)$. Formulas (6.2)–(6.6) define, for each $\alpha \in \mathbb{C}$, a Lie algebra homomorphism

$$\rho_\alpha : \mathcal{G} \to D^1(S^1).$$

Thus, for each $\alpha \in \mathbb{C}$, the image $\rho_\alpha(\mathcal{G})$ is a Lie subalgebra of linear differential operators on S^1. We now ask what happens as $\alpha \to \infty$, i.e., as $\beta \equiv 1/\alpha \to 0$.

Theorem 6.1: Consider

$$\beta \to \rho_{1/\beta}(\mathcal{G}) \equiv \mathcal{L}_\beta$$

as a one-parameter family of Lie algebras of differential operators for $\beta \neq 0$. If $\mathcal{L}$ is defined as the Lie algebra generated by the following operators

$$\frac{d}{d\theta}, \quad i\cos\theta, \quad -i\sin\theta, \tag{6.7}$$

then $\beta \to \mathcal{L}_\beta$ is a smooth family of Lie subalgebras of $D^1(S^1)$, including the point $\beta = 0$. $\mathcal{L}_0$ is isomorphic to the Lie algebra of the group of rigid motions in R^2.

Proof: $(\beta A, \beta A_2, A_1)$ form a basis for $\mathcal{L}_\beta$, which goes over, as $\beta \to 0$, to the Lie algebra (6.7). Q.E.D.

We can now, as in Ref. 5, p. 174, easily compute the matrix elements of the one-parameter groups $t \to \exp(t\rho_\alpha(A))$, and verify, using the geometric techniques developed there, that the hypotheses are satisfied that are needed to apply the methods of Sec. 5 for describing the asymptotic behavior of these matrix elements. We see that the special example of formula (1.1) is quite typical of the general matrix element.

[1] R. Hermann, "Some algebraic, geometric and system-theoretic properties of the Special Functions of mathematical physics," J. Math. Phys. **23**, 1282 (1982).

[2] N. J. Vilenkin, *Special Functions and the Theory of Group Representations* (Amer. Math. Soc., Providence, RI, 1968).

[3] W. Miller, Jr., *Symmetry and Separation of Variables* (Addison-Wesley, Reading, MA, 1977).

[4] R. Hermann, "Analytic continuation of group representations," Comm. Math. Phys.; Part II, **3**, 53–74 (1966); Part III, **3**, 75–97 (1966); Part IV, **5**, 131–156 (1967); Part V **5**, 157–190 (1967); Part VI, **6**, 205–225 (1967).

[5] R. Hermann, "Geometric ideas in Lie group harmonic analysis theory," Proc. Washington Symp. on Symmetric Spaces, edited by W. Boothby and G. Weiss (Marcel Dekker, New York, 1972), pp. 157–209.

[6] M. Hazewinkel, "On families of linear systems: Degeneration phenomena," in *Lectures in Applied Mathematics* (Amer. Math. Soc., Providence, RI, 1980), Vol. 18, pp. 157–190.

[7] M. Hazewinkel, "Moduli and canonical forms for linear dynamical systems. II: The topological case," Math. Systems Theory **10**, 363–385 (1977).

[8] M. Hazewinkel, "(Fine) moduli spaces for linear systems. What are they and what are they good for?" in *Proc. NATO–AMS Summer Inst. on Algebraic and Geometric Methods in Linear Systems Theory* (Reidel, 1980).

[9] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U.P., Cambridge, 1927).

[10] E. Inonu and E. P. Wigner, "On the contraction of groups and their representations," Proc. Natl. Acad. Sci. U.S. **39**, 510–524 (1953).

[11] C. Martin and R. Hermann, "Applications of algebraic geometry to systems theory: The McMillan degree and Kronecker indices of transfer functions as topological and holomorphic system invariants," SIAM J. Control Optim. **16**, 743–755 (1978).

[12] F. Oberhettinger and L. Bφdii, *Tables of Laplace Transforms* (Springer–Verlag, New York, 1973).

[13] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," IEEE Trans. Autom. Control **26**(2), 301–320 (1981).

[14] G. Zames and B. Francis, "Feedback, minimax sensitivity, and optimal synthesis," preprint, McGill Univ., 1982.

[15] K. Yosida, *Functional Analysis*, 6th ed. (Springer-Verlag, New York, 1980).

[16] J. Mikusinski, *Operational Calculus* (Pergamon, London, 1959).

The paths of integration for the new generalized Bessel transform

E. Bahar
Electrical Engineering Department, University of Nebraska, Lincoln, Nebraska 68588

(Received 14 April 1981; accepted for publication 16 October 1981)

It is shown that the comments made in a recent paper regarding the previously developed "New generalized Bessel transform and its relationship to the Fourier, Watson, and Kontorowich–Lebedev transforms" are based on erroneous assumptions. The claim that the path of integration parallel to the real axis (below the singularities of the transform function on the real axis) cannot be transformed to a contour around the singularities of the transfer function in the lower half-plane contradicts the very basis of the firmly established Watson transformation and the most advanced theories in radio wave propagation over the Earth's surface and cylindrical structures.

PACS numbers: 02.30.Gy

I. SUMMARY OF NEW GENERALIZED BESSEL TRANSFORM

In the paper "New generalized Bessel transform and its relationship to the Fourier, Watson, and Kontorowich–Lebedev transforms ,[1] " the following transform pair was derived:

$$E_z(\xi,\phi) = \int_L E(\nu,\phi)\psi_\nu(\xi)d\nu, \tag{1}$$

$$E(\nu,\phi) = \tfrac{1}{4}\int_\xi^\infty E_z(\xi,\phi)H_\nu^{(2)}(\xi)\frac{\nu}{\xi}d\xi, \tag{2}$$

in which

$$\psi_\nu(\xi) = H_\nu^{(1)}(\xi) + R_\nu H_\nu^{(2)}(\xi), \tag{3}$$

where $H_\nu^{(1)}(\xi)$ and $H_\nu^{(2)}(\xi)$ are the Hankel functions of the first and second kind, respectively, ν is the order and $\xi = kr$ is the argument (k is the wave number and r is the distance from the z axis in the cylindrical coordinate system,r,ϕ,z). The path of integration L lies parallel to the real axis such that all the singularities of $E(\nu,\phi)$ on the real axis lie above the path L (see Fig. 1 reproduced from Ref. 1). The coefficient R_ν in (3) depends upon the boundary condition at $\xi = \xi_R$. Thus for the Dirichlet condition

$$E_z(\xi_R,\phi) = 0, \tag{4}$$

$$R_\nu = -H_\nu^{(1)}(\xi_R)/H_\nu^{(2)}(\xi_R). \tag{5}$$

The more general expression for R_ν for the impedance boundary condition is given in Ref. 1. Using (1) and (2) the following expression for the Dirac delta function $\delta(\xi - \xi_0)$ was obtained:

$$\xi\delta(\xi - \xi_0) = \tfrac{1}{4}\int_L \psi_\mu(\xi)H_\mu^{(2)}(\xi_0)\mu\, d\mu, \quad \xi < \xi_0. \tag{6}$$

On applying the transforms (1) and (2) to the problem of radiation by a line source (at $\xi = \xi_0$, $\phi = \phi_0$) parallel to a perfectly-conducting cylinder of radius $\xi_R < \xi_0$, the following expression for the vector potential was obtained:

$$A(\xi,\phi) = \int_L a(\mu,\phi)\psi_\mu(\xi)d\mu, \tag{7}$$

in which the solution for the transform $a(\mu,\phi)$ was given by

$$a(\mu,\phi) = -\tfrac{1}{8}\mu_0 I H_\mu^{(2)}(\xi_0) \times [\cot\mu\pi\cos\mu(\phi - \phi_0) + \sin\mu(\phi - \phi_0)], \tag{8}$$

where μ_0 is the permeability of free space and I is the intensity of the current filament. Since $H_\mu^{(2)}(\xi_0)\psi_\mu(\xi)\sin\mu(\phi - \phi_0)$ is an odd function of μ and since it is analytic on the real axis, the second term in (8) contributes nothing to $A(\xi,\phi)$ and it can therefore be surpressed. Thus (7) was shown to reduce to

$$A(\xi,\phi) = -\tfrac{1}{8}\mu_0 I\int_L H_\mu^{(2)}(\xi_0)\psi_\mu(\xi)\cot\mu\pi\cos\mu(\phi - \phi_0)d\mu. \tag{9}$$

On deforming the path of integration L to the contour $C_1' + C_2$ (see Fig. 1) and employing Cauchy's integral theorem to account for the contributions from the residues at the poles of $a(\mu,\phi)$ on the real axis ($\nu = 0,1,2,3,\ldots$), it was shown that

$$A(\xi,\phi) = i\tfrac{1}{8}\mu_0 I\sum_n \epsilon_n H_n^{(2)}(\xi_0)\psi_n(\xi)\cos n(\phi - \phi_0), \tag{10}$$

where

$$\epsilon_n = \begin{cases} 1, & n = 0 \\ 2, & n = 1,2,3,\cdots. \end{cases} \tag{11}$$

The solution for $\xi > \xi_0$ is obtained by interchanging ξ with ξ_0 in (10).[1]

The corresponding Watson transform is obtained by closing the path of integration L by an infinite semicircle in the negative half-plane.[2,3] "The contribution from this por-

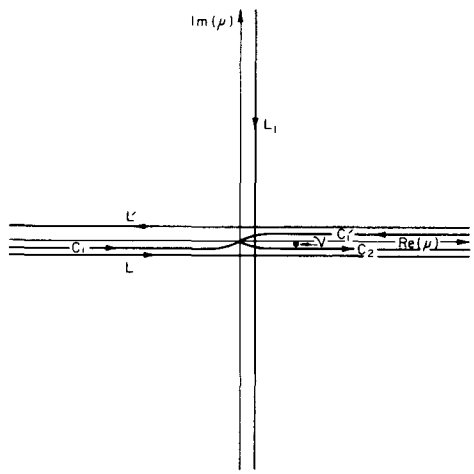


FIG. 1. Integration paths in the complex μ plane.

 0022-2488/83/010033-03$02.50

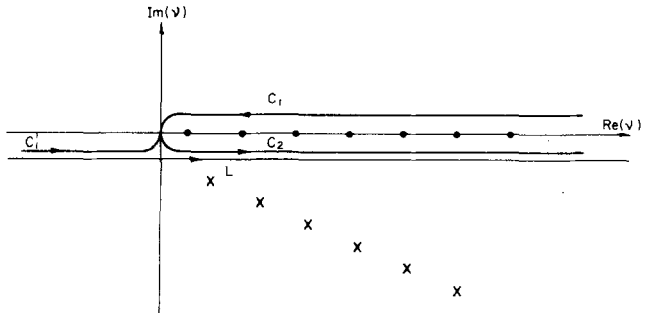


FIG. 2. The contour in the complex ν plane showing the location of the real and complex poles.

tion of the contour vanishes as the radius of the semicircle approaches infinity" (Ref. 3). Thus on noting that the poles in the lower half-plane are at the values of ν_n that satisfy the modal equation $1/R_\nu = 0$ (see Fig. 2 reproduced from Ref. 3), the following Watson expansion for $A(\xi,\phi)$ was derived from (10):

$$A(\xi,\phi) = -\frac{1}{4} i\pi\mu_r I \sum_n \left[H_\mu^{(2)}(\xi_0) \frac{H_\mu^{(1)}(\xi_R)}{\partial [H_\mu^{(2)}(\xi_R)]/\partial\mu} \times H_\mu^{(2)}(\xi) \cot \mu\pi \cos\mu(\phi - \phi_0) \right]_{\mu=\mu_n}. \quad (12)$$

The relationship between (10) and (12) is at the core of the Watson transformation.[2,3]

For $\xi_R \to \infty$ and for $\xi_R \to 0$ the above Watson expression cannot be used and it was shown that in these limits the solution for $A(\xi,\phi)$, Eq.(7), reduces to the Fourier transform and the Kontorowich–Lebedev transform, respectively.[1]

As indicated in Ref. 1, the motivation for the derivation of the transform pair (1) and (2) was to obtain complete expansions for the electromagnetic fields in the vicinity of cylindrical structures characterized by variable radius of curvature $\rho(0 < \rho < \infty)$ and surface impedance.[4] Fields transforms have also been derived for cylindrical or spherical structures with n concentric layers.[5]

II. EXAMINATION OF SAMADDAR'S CLAIMS

In a recent paper,[6] Samaddar claims that on using an entirely different method to derive the transform pair (1) and (2), he has shown that instead of the contour L (see Fig. 1), the path of integration should have been the closed contour around the singularities of R_ν(5) [namely the zeros of $H_\nu^{(2)}(\xi_R)$]. He further states that it is not permissible to deform the closed contour around the singularities of R_ν to the path L. Such a statement contradicts the very foundation of the Watson transformation,[2] which has been securely established for over 60 years and is the basis for the most advanced theories in radio wave propagation around the surface of the Earth.[3] According to Samaddar,[6] the contour L in (6) and therefore the transform pair (1) and (2) "cannot be used for any arbitrary function $E_z(\xi,\phi)$, which has a strong singularity like a delta function $\delta(\xi - \xi_0)$." To justify his statement he maintains that the condition

$$\tau\pi/2\phi < 2|\nu|/e\xi_R, \quad \tau > 1 \quad (13)$$

must be met as $|\nu| \to \infty$ in order to deform the closed contour around the poles of R_ν to the path L. Thus he goes on to state "It may be noted that since $\phi = 0$ in (2.5) condition (2.13) cannot be fulfilled and consequently the contour C in (2.5) cannot be deformed onto the portion of L lying to the left of the lowest-order zero of $H_\nu^{(2)}(\xi_r)$." (Equation numbers of the form (n) are in the notation of the present paper; those of the form (m,n) are in the notation of Samaddar's paper.)

Samaddar also states that "in the development of the pair (1) and (2) it is assumed implicitly that $E(\nu,\phi)$ is analytic in a horizontal strip bounded by two lines parallel to the real axis of the complex ν plane (one below and the other above)." Nowhere in the development of the transform pair (1) and (2) was such an implicit or explicit assumption made. On the contrary, the transform $a(\mu,\phi)$(8)[1] contains not only an analytic part $H_\mu^{(2)}\xi_0\psi_\mu(\xi)\sin\mu(\phi - \phi_0)$ (that was suppressed) but also a term proportional to $\cot\mu\pi$. Were it not for the poles at $\mu = 0,1,2,\ldots$, it would be impossible to obtain the correct results (10). Based on his assumption, this author then states "the path L in (1) can be shifted onto the entire real axis of the ν plane." It is "evident" according to him from the properties of the Hankel functions and $E(\nu,\phi)$ that the integrand in (1) is an odd function of ν and, therefore, the integral (1) vanishes identically. If, as this author claims, the integral (1) "vanishes identically" the transform pair (1) and (2) cannot be used for any function $E_z(\xi,\phi)$ whether it has a strong or weak singularity or no singularity at all. Yet the transform pair (1) and (2) from which the spectral representation for the Dirac delta function $\delta(\xi - \xi_0)$ was derived (6), yields the correct result for the fields due to a line source in the vicinity of a perfectly conducting cylinder (10),(12).[1]

Again in his concluding remarks this author assumes that in going from Eqs. (2.6b) to (2.7b) in Ref. 1 "the transform function $E(\nu,\phi)$ was implicitly treated as analytic inside a horizontal strip containing the real axis in the ν plane." However, in going from (2.6b) to (2.7b) it is only necessary to note that the analytic terms can be suppressed without affecting the final results. Stated more formally, we partition the set of transforms into a set of "equivalence class of functions" by means of the following equivalence relation. Any two functions in the set of transforms $E(\nu,\phi)$ are considered equal if they differ by a function $F(\nu,\phi)$ such that $F(\nu,\phi)\psi_\nu(\xi)$ is odd and regular on the real axis.

It is interesting to note that the "entirely different method"[7] Samaddar refers to in his paper[6] was also applied to the more general problem of propagation in cylindrical or spherical structures with n concentric layers.[5] Line source excitations were considered involving Dirac delta functions. A precise criterion for the deformation of the contour was derived[5] and it was shown that these more general results are consistent with those obtained by Bahar in his earlier work.[1] Regretably Samaddar made several erroneous assumptions and did not realize what is firmly established through the Watson transform,[2,3,5] namely, that the integration along the path L can be deformed to the contour integration around the poles of the function R_ν(5).[2,3,5]

[1]E. Bahar, J. Math. Phys. **12**, 179 (1971).
[2]G. N. Watson, Proc. R. Soc. (London) **A95**, 546 (1919).
[3]J. R. Wait, *Electromagnetic Waves in Stratified Media* (Pergamon, Oxford, 1962), p. 110.
[4]E. Bahar, J. Math. Phys. **12**, 186 (1971).
[5]E. Bahar, Can. J. Phys. **53**, 1078 (1975).
[6]S. N. Samaddar, J. Math. Phys. **22**, 39 (1981).
[7]B. Friedman, *Principles and Techniques of Applied Mathematics* (Wiley, London, 1956).

Thorpe–Hitchin inequality for compact Einstein 4-manifolds of metric signature (+ + − −) and the generalized Hirzebruch index formula

Yasuo Matsushita
Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, Kyoto, Japan

(Received 30 November 1981; accepted for publication 20 August 1982)

It is proved that the Euler characteristic and the Hirzebruch index of a compact oriented Einstein 4-manifold of metric signature (+ + − −) satisfy an inequality which is well known as the Thorpe–Hitchin inequality for the case of a Riemannian metric. To derive the inequality, a generalized Hirzebruch formula relating the index to the first pseudo-Pontrjagin number of the manifold is proved. This formula may be contrasted with Chern's generalized Gauss–Bonnet formula for a pseudo-Riemannian manifold.

PACS numbers: 02.40. + m

I. INTRODUCTION

It was shown by Thorpe[1] and later by Hitchin[2] that for a compact oriented Einstein 4-manifold with a Riemannian metric, the Euler characteristic χ and the Hirzebruch index τ of the manifold satisfy the inequality

$$|\tau| \leqslant \tfrac{2}{3}\chi. \tag{1.1}$$

The purpose of this paper is to show that this inequality also holds for compact oriented Einstein 4-manifolds of metric signature (+ + − −). The author's earlier result[3] is that for a compact oriented 4-manifold of metric signature (+ + − −), the Euler characteristic of the manifold is even and is congruent mod 4 to the Hirzebruch index of the manifold. Moreover, if the manifold admits an Einstein metric of such a signature, then the Euler characteristic is nonnegative.

Every indefinite metric on a two- or three-dimensional manifold is necessarily of Lorentz type. A compact oriented manifold admits a Lorentz metric if and only if the Euler characteristic vanishes.[4] Thus an investigation of 4-manifolds of metric signature (+ + − −) is important from the point of view of the differential topology of pseudo-Riemannian manifolds.

As remarked at the end of the author's earlier paper,[3] it has remained open whether or not the pseudo-Pontrjagin number defined on a pseudo-Riemannian bundle over a manifold coincides with the Pontrjagin number defined on the tangent bundle over the manifold.

In this paper, this is affirmatively proved for the case of metric signature (+ + − −). Such a coincidence enables us to give an analog of the Hirzebruch formula between the Hirzebruch index and the pseudo-Pontrjagin number for 4-manifolds of metric signature (+ + − −). Applying this new generalized Hirzebruch formula to an Einstein 4-manifold of this signature, we have the Thorpe–Hitchin inequality. It should be noted that the generalized Hirzebruch formula may be contrasted to the generalized Gauss–Bonnet formula of Chern[5] for a pseudo-Riemannian manifold.

In comparison of the three types of metric signature (+ + + +), (+ + + −), and (+ + − −), the last type under consideration is important for the following three reasons. (1) The first is that such a metric is the lowest-dimensional example of an indefinite metric that is not a Lorentz metric, as stated before. (2) There is a similarity between Riemannian 4-manifolds and 4-manifolds of metric signature (+ + − −) since the Lie algebras so(4) and so(2,2) have similar isomorphisms

$$\mathrm{so}(4) = \mathrm{so}(3) + \mathrm{so}(3), \tag{1.2a}$$

$$\mathrm{so}(2,2) = \mathrm{so}(1,2) + \mathrm{so}(1,2). \tag{1.2b}$$

(3) There is another similarity between Einstein 4-manifolds of metric signature (+ + − −) and those of Lorentz signature in that there are three types, known as the Petrov types,[6,7] of the normal forms of the curvature tensors.

Thus the signature (+ + − −) becomes the primary concern of the present paper.

In Sec. II, the generalized Hirzebruch formula is proved along a line of thought of Chern.[5] In Sec. III, as preliminaries, the three types of normal forms of the curvature tensor for an Einstein manifold will be given together with the Euler characteristic and the pseudo-Pontrjagin number for each type.[3] In the last section, the Thorpe–Hitchin inequality is proved for each type of Einstein 4-manifold of metric signature (+ + − −), and the characteristic numbers are illustrated.

II. GENERALIZED HIRZEBRUCH INDEX FORMULA

By a manifold we mean a connected, paracompact, C^∞-differentiable manifold. Let ξ be the tangent bundle $\pi: E \to M$ over the four-dimensional manifold M. The bundle ξ is called a pseudo-Riemannian tangent bundle if there exists a nondegenerate symmetric bilinear form $(\ ,\ )$ in each fiber $\pi^{-1}(x)$, which varies in a C^∞ way with $x \in M$.

For the present we only consider the case that the signature of $(\ ,\ )$ is of type (+ + − −) throughout the manifold M. We may impose, in addition, a Riemannian structure $\langle\ ,\ \rangle_R$ on ξ, such that $\langle u,v\rangle_R$, $u, v \in \pi^{-1}(x)$, is a quadratic form of signature (+ + + +), which also varies in a C^∞ way with x. Such a form exists since M is paracompact.

For a fixed $x \in M$, a vector $u_0 \in \pi^{-1}(x)$ is an eigenvector of $(\ ,\ )$ relative to $\langle\ ,\ \rangle_R$, with eigenvalue λ, if

$$(u_0, v) = \lambda \langle u_0, v\rangle_R \tag{2.1}$$

for all $v \in \pi^{-1}(x)$. There are two positive and two negative eigenvalues. The fiber over each point x can be split into two-dimensional subspaces $\pi_+^{-1}(x)$ and $\pi_-^{-1}(x)$ spanned by the

 0022-2488/83/010036-05$02.50

positive and negative eigenvectors, respectively, as follows:

$$\pi^{-1}(x) = \pi_+^{-1}(x) + \pi_-^{-1}(x). \tag{2.2}$$

Thus any vector $u \in \pi^{-1}(x)$ can be decomposed into two parts:

$$u = u_+ + u_-, \tag{2.3}$$

where $u_+ \in \pi_+^{-1}(x)$ and $u_- \in \pi_-^{-1}(x)$. Such a structure implies that the bundle ξ can be written as a Whitney sum

$$\xi = \xi_+ + \xi_- \tag{2.4}$$

of the subbundles ξ_+ and ξ_- with the total spaces $E_+ = U_{x\in M}\, \pi_+^{-1}(x)$ and $E_- = U_{x\in M}\, \pi_-^{-1}(x)$, respectively.

We define in terms of $(\ ,\ )$, two quadratic forms $\langle\ ,\ \rangle_+$, $\langle\ ,\ \rangle_-$ as follows for any two vectors $u = u_+ + u_-$ and $v = v_+ + v_-$:

$$\langle u,v\rangle_+ = (u_+,v_+), \tag{2.5a}$$

$$\langle u,v\rangle_- = -(u_-,v_-). \tag{2.5b}$$

Both quadratic forms are positive definite, and accordingly define the Riemannian structures on ξ_+ and ξ_-, respectively. The quadratic form (u,v) can therefore be written in terms of these forms as

$$(u,v) = \langle u,v\rangle_+ - \langle u,v\rangle_-, \tag{2.6a}$$

and the expression

$$\langle u,v\rangle = \langle u,v\rangle_+ + \langle u,v\rangle_- \tag{2.6b}$$

defines a Riemannian structure on the bundle ξ.

In ξ_+ and ξ_-, take connections ω_+ and ω_- admissible to $\langle\ ,\ \rangle_+$ and $\langle\ ,\ \rangle_-$, respectively. Then the direct sum

$$\omega = \omega_+ + \omega_- \tag{2.7}$$

is a connection in ξ, which is admissible to both structures $(\ ,\ )$ and $\langle\ ,\ \rangle$. Denote by $\Omega_{(+)}$ and $\Omega_{(-)}$ the curvature 2-forms on ξ_+ and ξ_-, expressed by 2×2 matrices, as derived from the connections ω_+ and ω_-, respectively. Then the 4×4 matrix

$$\Omega = \begin{bmatrix} \Omega_{(+)} & 0 \\ 0 & \Omega_{(-)} \end{bmatrix} \tag{2.8}$$

is a curvature on ξ.

The first Pontrjagin class p_1 of the bundle ξ is represented by the ad(SO(4))-invariant closed 4-form β_1 on M given by the formula

$$\det[I_4 - (1/2\pi)\Omega\,] = \pi^*(1+\beta_1), \tag{2.9}$$

where $I_4 = \mathrm{diag}[+1,+1,+1,+1]$. Analogously, we give:

Definition 1: The class represented by the ad(SO(2,2))-invariant 4-form $\bar{\beta}_1$ on M in the formula

$$\det[I_{2,2} - (1/2\pi)\Omega\,] = \pi^*(1+\bar{\beta}_1) \tag{2.10}$$

is called the first pseudo-Pontrjagin class and is denoted by $\bar{p}_1$, where $I_{2,2} = \mathrm{diag}[+1,+1,-1,-1]$.

For such classes we have:

Proposition 2: $p_1 = \bar{p}_1$.

Proof: This is shown by a simple calculation as follows. Using (2.8), we have

$$\begin{aligned}
&\det[I_{2,2} - (1/2\pi)\Omega\,] \\
&= \det[I_2 - (1/2\pi)\Omega_{(+)}\,] \wedge \det[-I_2 - (1/2\pi)\Omega_{(-)}\,] \\
&= \det[I_2 - (1/2\pi)\Omega_{(+)}\,] \wedge \det{}^t[-I_2 - (1/2\pi)\Omega_{(-)}\,] \\
&= \det[I_2 - (1/2\pi)\Omega_{(+)}\,] \wedge \det[-I_2 + (1/2\pi)\Omega_{(-)}\,] \\
&= \det[I_2 - (1/2\pi)\Omega_{(+)}\,] \wedge \det[I_2 - (1/2\pi)\Omega_{(-)}\,],
\end{aligned}$$

which coincides with (2.9). In the above, the last equality holds since $\det A = \det(-A)$ for any 2×2 matrix A. □

Now let us state the generalized Hirzebruch index formula.

Theorem 3: Let M be a compact oriented 4-manifold of metric signature $(+\ +\ -\ -)$, and $\bar{p}_1[M]$ be the first pseudo-Pontrjagin number of M, a numerical analog of the first Pontrjagin number $p_1[M]$. Then the Hirzebruch index $\tau[M]$ of M is given in terms of $\bar{p}_1[M]$ by the formula

$$\tau[M] = \tfrac{1}{3}\bar{p}_1[M]. \tag{2.11}$$

Proof: This is clear from the Hirzebruch formula $\tau[M] = \frac{1}{3}p_1[M]$ together with Proposition 2. □

III. CURVATURES OF EINSTEIN 4-MANIFOLDS OF METRIC SIGNATURE $(+\ +\ -\ -)$

It is important to compare the Lie algebras so(4), so(2,2), and so(3,1), which are the Lie algebras of the structure groups for a Riemannian 4-manifold, a 4-manifold of metric signature $(+\ +\ -\ -)$, and a Lorentz 4-manifold, respectively. For the first two Lie algebras, there are similar isomorphisms

$$\mathrm{so}(4) = \mathrm{so}(3) + \mathrm{so}(3), \tag{3.1a}$$

$$\mathrm{so}(2,2) = \mathrm{so}(1,2) + \mathrm{so}(1,2). \tag{3.1b}$$

On the other hand, there is no such decomposition for so(3,1). These isomorphisms imply that the space Λ^2 of 2-forms at each point is decomposed into two parts

$$\Lambda^2 = \Lambda^2_+ + \Lambda^2_-, \tag{3.2}$$

where $\Lambda^2_\pm$ are the ± 1 eigenspace of the Hodge star operator $*$, with

$$*^2 = 1. \tag{3.3}$$

On the Lorentz 4-manifold, however, the star operator for Λ^2 satisfies

$$*^2 = -1, \tag{3.4}$$

and Λ^2 cannot be decomposed in a similar way.

Now consider Λ^2 on M of metric signature $(+\ +\ -\ -)$. Corresponding to the decomposition (3.2), we introduce a basis $\{E^i_+, E^j_-\}$ $(i,j = 1,2,3)$, the duality basis, for $\Lambda^2 = \Lambda^2_+ + \Lambda^2_-$, with the following properties: for the star operator

$$*E^i_+ = E^i_+, \quad *E^i_- = -E^i_-; \tag{3.5}$$

for the inner product

$$(E^i_\sigma, E^j_\rho) = \hat{\epsilon}^i \delta_{\sigma\rho}\delta^{ij} \quad (\sigma,\rho = +,-); \tag{3.6}$$

and for the wedge product

$$\begin{aligned}
&E^i_+ \wedge E^j_+ = -E^i_- \wedge E^j_- = \hat{\epsilon}^i \delta^{ij} w, \\
&E^i_+ \wedge E^j_- = 0, \quad E^i_- \wedge E^j_+ = 0,
\end{aligned} \tag{3.7}$$

where the symbols used above are as follows: δ^{ij} is the Kronecker delta; $\hat{\epsilon}^1 = -\hat{\epsilon}^2 = -\hat{\epsilon}^3 = +1$, $\delta_{++} = \delta_{--} = 1$, $\delta_{+-} = \delta_{-+} = 0$, and w is the volume element.

The curvature tensor R is a linear transformation in Λ^2, and is expressed by a 6×6 matrix. Relative to the duality basis, R is decomposed into four disjoint parts

$$R = R_+ + R_- + R_{+-} + R_{-+}, \tag{3.8}$$

where $R_\pm \in \mathrm{End}(\Lambda^2_\pm)$, $R_{+-} \in \mathrm{Hom}(\Lambda^2_-, \Lambda^2_+)$, $R_{-+} \in \mathrm{Hom}(\Lambda^2_+, \Lambda^2_-)$.

If M admits an Einstein metric of signature $(+\ +\ -\ -)$, then the curvature tensor takes the general form

$$R_+ = \begin{bmatrix} P_+ & 0 \\ 0 & 0 \end{bmatrix}, \quad R_- = \begin{bmatrix} 0 & 0 \\ 0 & P_- \end{bmatrix}, \quad R_{+-} = R_{-+} = 0, \tag{3.9}$$

where

$$P_\pm = \begin{bmatrix} a_1 \pm \alpha_1 & b \mp \beta & c \mp \gamma \\ -b \mp \beta & -a_2 \pm \alpha_2 & d \pm \delta \\ -c \mp \gamma & d \pm \delta & -a_3 \pm \alpha_3 \end{bmatrix} \tag{3.10}$$

with constraints

$$\tfrac{1}{2}\mathrm{tr}R = \mathrm{tr}P_+ = \mathrm{tr}P_- = \sum_{i=1}^{3} \hat{\epsilon}^i a_i = \tfrac{1}{4}S, \tag{3.11a}$$

S = scalar curvature,

and

$$\sum_{i=1}^{3} \alpha_i = 0. \tag{3.11b}$$

The normal forms of the curvature tensor are given as follows.

Lemma 4[3]: For an Einstein 4-manifold of metric signature $(+\ +\ -\ -)$, the curvature tensor at each point takes one of the following three forms:

$$(1)\quad P_\pm = \begin{bmatrix} \mu_1 \pm \nu_1 & 0 & 0 \\ 0 & -\mu_2 \pm \nu_2 & 0 \\ 0 & 0 & -\mu_3 \pm \nu_3 \end{bmatrix}, \tag{3.12a}$$

with constraints

$$\sum_{i=1}^{3} \hat{\epsilon}^i \mu_i = \frac{S}{4}, \quad \sum_{i=1}^{3} \nu_i = 0; \tag{$*$}$$

$$(2)\quad P_\pm = \begin{bmatrix} \mu_2 \pm (\nu_1/2 + \nu_2) & \mp \nu_1/2 & 0 & 0 \\ \mp \nu_1/2 & -\mu_2 & \pm(\nu_1/2 - \nu_2) & 0 \\ 0 & 0 & \tfrac{1}{4}S \mp \nu_1 \end{bmatrix}, \tag{3.12b}$$

with $\nu_1 \neq 0$;

$$(3)\quad P_\pm = \begin{bmatrix} -\tfrac{1}{4}S & \mp\kappa & 0 \\ \mp\kappa & \tfrac{1}{4}S & \pm\kappa \\ 0 & \pm\kappa & \tfrac{1}{4}S \end{bmatrix}, \tag{3.12c}$$

with $\kappa \neq 0$.

Hereafter we consider a compact oriented Einstein 4-manifold of metric signature $(+\ +\ -\ -)$, also denoted by M. We need the Euler forms and the pseudo-Pontrjagin forms for the main theorem stated in the next section.

At each point of M, corresponding to each normal form of the curvature tensor, the Euler form χ is of one of the following three types:

$$(1)\quad \chi = (1/4\pi^2) \sum_{i=1}^{3} (\mu_i^2 + \nu_i^2) w, \tag{3.13a}$$

with constraints $(*)$ in Lemma 4;

$$(2)\quad \chi = (1/4\pi^2)[(\tfrac{1}{4}S)^2 + \nu_1^2 + 2(\mu_2^2 + \nu_2^2)]w, \tag{3.13b}$$

with $\nu_1 \neq 0$;

$$(3)\quad \chi = (3S^2/2^6\pi^2)w. \tag{3.13c}$$

Corresponding to each of the above expressions, the pseudo-Pontrjagin form $\bar{p}_1$ is of one of the following types:

$$(1)\quad \bar{p}_1 = (1/\pi^2) \sum_{i=1}^{3} (\hat{\epsilon}^i \mu_i \nu_i) w, \tag{3.14a}$$

with constraints $(*)$;

$$(2)\quad \bar{p}_1 = (1/\pi^2)(-\tfrac{1}{4}S\nu_1 + 2\mu_2\nu_2)w, \tag{3.14b}$$

with $\nu_1 \neq 0$;

$$(3)\quad \bar{p}_1 = 0. \tag{3.14c}$$

The Euler characteristic and the pseudo-Pontrjagin number are obtained by integrating above forms over M.

IV. THORPE–HITCHIN INEQUALITY

By Theorem 3, the Hirzebruch index form is given in terms of the pseudo-Pontrjagin form. Thus we have:

Proposition 5: Let M be a compact oriented Einstein 4-manifold of metric signature $(+\ +\ -\ -)$. At each point of M, corresponding to each normal form of the curvature tensor given in Lemma 4, the Hirzebruch index form τ is of one of the following three types:

$$(1)\quad \tau = (1/3\pi^2) \sum_{i=1}^{3} (\hat{\epsilon}^i \mu_i \nu_i) w, \tag{4.1a}$$

with constraints $(*)$;

$$(2)\quad \tau = (1/3\pi^2)(-\tfrac{1}{4}S\nu_1 + 2\mu_2\nu_2)w, \tag{4.1b}$$

with $\nu_1 \neq 0$;

$$(3)\quad \tau = 0. \tag{4.1c}$$

The Hirzebruch index of M is obtained by integrating the above form over M.

Now let us state the main theorem.

Theorem 6: Let M be a compact oriented Einstein 4-manifold of metric signature $(+\ +\ -\ -)$. Then the Euler characteristic $\chi[M]$ and the Hirzebruch index $\tau[M]$ of M satisfy the inequality

$$|\tau[M]| \leqslant \tfrac{2}{3}\chi[M]. \tag{4.2}$$

Proof: Consider a point of M. If the curvature tensor at the point is of type 1, then from (3.13a) and (4.1a) we have

$$\tfrac{2}{3}\chi \mp \tau = \frac{1}{6\pi^2}\sum_{i=1}^{3}[(\mu_i^2+\nu_i^2)\mp 2(\hat{\epsilon}^i\mu_i\nu_i)]w,$$
$$= \frac{1}{6\pi^2}\sum_{i=1}^{3}(\hat{\epsilon}^i\mu_i \mp \nu_i)^2 w, \tag{4.3a}$$

where the rhs vanishes iff $\hat{\epsilon}^i\mu_i = \pm\nu_i$ $(i=1,2,3)$.

Similarly if it is of type 2, then from (3.13b) and (4.1b) we have

$$\tfrac{2}{3}\chi \mp \tau = (1/6\pi^2)[(\tfrac{1}{4}S \pm \nu_1)^2 + 2(\mu_2 \mp \nu_2)^2]w, \tag{4.3b}$$

where the rhs vanishes iff $\nu_1 = \mp\frac{1}{4}S$ and $\mu_2 = \pm\nu_2$.

If it is of type 3, then from (4.1c), $\tau = 0$, and hence we have

$$\tfrac{2}{3}\chi \mp \tau = \tfrac{2}{3}\chi = (S^2/2^5\pi^2)w, \tag{4.3c}$$

where the rhs vanishes iff $S=0$. Therefore, integrating the above forms over M, we can conclude that the inequality

$$\tfrac{2}{3}\chi[M] \mp \tau[M] \geqslant 0 \tag{4.4}$$

holds. This completes the proof. □

Next we consider special cases for M, where the curvature tensor has the same type of normal form at every point of the manifold.

Theorem 7: Let M_i be a compact oriented Einstein 4-manifold of metric signature $(+\ +\ -\ -)$, whose curvature tensor is of type i at every point of the manifold. Then for $i=1,2$, the Euler characteristic $\chi[M_i]$ and the Hirzebruch index $\tau[M_i]$ of M_i satisfies the inequality

$$|\tau[M_i]| \leqslant \tfrac{2}{3}\chi[M_i], \tag{4.5}$$

where the equality holds if and only if the curvature elements at each point of M_i satisfy the following relations:

for M_1: $\hat{\epsilon}^j\mu_j = \nu_j$ or $\hat{\epsilon}^j\mu_j = -\nu_j$ $(j=1,2,3)$;

for M_2: $\nu_1 = -\frac{1}{4}S$, $\mu_2 = \nu_2$ or $\nu_1 = \frac{1}{4}S$, $\mu_2 = -\nu_2$.

For M_3, we have

$$\tau[M_3] = 0, \tag{4.6}$$

and hence

$$\chi[M_3] \geqslant 0 \tag{4.7}$$

with equality iff $S=0$.

Proof: This is clear from the proof of Theorem 6. □

The following facts are fundamental concerning the existence of pseudo-Riemannian metrics on manifolds.[4]

Proposition 8: Let X be a compact oriented manifold. Then X admits a Lorentz metric if and only if the Euler characteristic $\chi[X]$ of X vanishes.

Proposition 9: Let X be a $4m$-dimensional compact oriented manifold. Assume that X admits a pseudo-Riemannian metric of signature $(\underbrace{+\cdots+}_{p}\ \underbrace{-\cdots-}_{q})$, with $q \equiv 2 \bmod 4$, $p+q=4m$. Then the Euler characteristic $\chi[X]$ of X is even and is congruent mod 4 to the Hirzebruch index $\tau[X]$ of X.

Proof: It is known[4] that a compact manifold admits an everywhere defined, continuous, nonsingular, quadratic form of signature q if and only if it admits a continuous field of tangent q planes. The assertion of this proposition is directly derived from a result[8] of Atiyah that for a compact oriented $4m$-manifold admitting a field of tangent q planes, with $q \equiv 2 \bmod 4$, the Euler characteristic is even and congruent mod 4 to the Hirzebruch index of the manifold. □

Corollary 10: Let M be a compact oriented Einstein 4-manifold of metric signature $(+\ +\ -\ -)$. Then there is a nonnegative integer n such that

$$\chi[M] = \tau[M] + 4n, \tag{4.8}$$

with $n=0$ iff $\chi[M] = \tau[M] = 0$.

Proof: The previous proposition asserts that there is an integer n' such that $\chi[M] = \tau[M] + 4n'$. It follows from the main theorem that n' is nonnegative since

$$n' = \tfrac{1}{4}(\chi[M] - \tau[M]) \geqslant \tfrac{1}{4}(\tfrac{3}{2}|\tau[M]| - \tau[M]) \geqslant 0,$$

with equalities iff $\chi[M] = \tau[M] = 0$. □

Now let us look at the situations for each M_i in some detail, where the equalities of (4.5) and (4.7) hold.

Type 1, M_1: The curvature $R = R_+ + R_-$ takes a diagonal form, which is quite similar to the normal form, obtained by Singer and Thorpe,[9] of the curvature tensor of an Einstein 4-manifold with a Riemannian metric. When the conditions $\hat{\epsilon}^j\mu_j = \nu_j$ $(j=1,2,3)$ are satisfied, the part R_- of R vanishes, that is,

$$R = R_+, \quad R_- = R_{+-} = R_{-+} = 0. \tag{4.9}$$

This implies that R satisfies the so-called self-duality condition

$$*R = R. \tag{4.10}$$

In this case we have

$$\chi[M_1] = \tfrac{1}{2}\tau[M_1]. \tag{4.11}$$

On the other hand, R_+ vanishes iff $\hat{\epsilon}^j\mu_j = -\nu_j$ $(j=1,2,3)$. In this case the curvature satisfies the anti-self-duality condition

$$*R = -R. \tag{4.12}$$

For the characteristic numbers we have

$$\chi[M_1] = -\tfrac{1}{2}\tau[M_1]. \tag{4.13}$$

Concerning these manifolds, we have

Proposition 11: For M_1, if the curvature tensor at each point is self-dual in the above sense, then there is a nonnegative integer k such that

$$\chi[M_1] = 12k, \quad \tau[M_1] = 8k. \tag{4.14}$$

On the other hand, if it is anti-self-dual at each point, then there is a nonnegative integer k such that

$$\chi[M_1] = 12k, \quad \tau[M_1] = -8k. \tag{4.15}$$

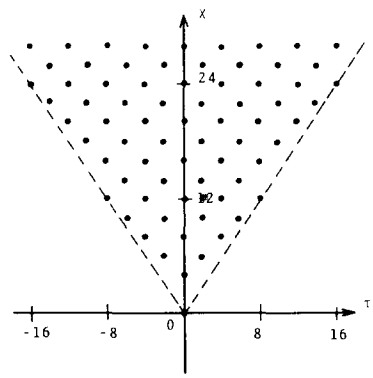


FIG. 1. (τ,χ) for M_1.

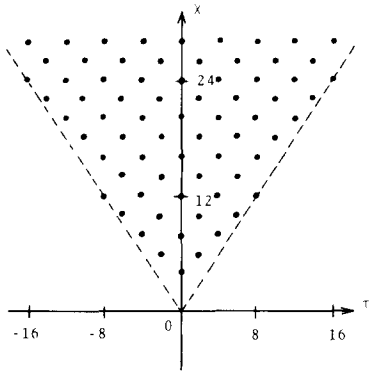


FIG. 2. (τ,χ) for M_2.

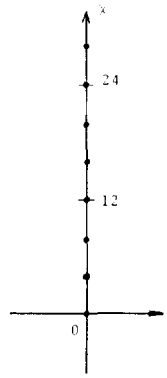


FIG. 3. (τ,χ) for M_3.

In both cases, $k = 0$ iff M_1 is flat.

Proof: Combining (4.8) with (4.11) and (4.13), we can easily show the desired relations. □

For these situations, see Fig. 1.

Type 2, M_2: Since the parts R_+ and R_- of type 2 contain a nonzero element ν_1, the curvature can never satisfy any duality condition. The Euler characteristic is always positive and its least value is 4. When the conditions $\nu_1 = -\frac{1}{4}S$, $\mu_2 = \nu_2$ (or $\nu_1 = \frac{1}{4}S$, $\mu_2 = -\nu_2$) are satisfied, we have

$$\begin{aligned}\chi[M_2] &= \tfrac{3}{2}\tau[M_2] \quad (\text{or} \quad -\tfrac{3}{2}\tau[M_2]) \\ &= \frac{S^2}{2^5\pi^2}\,\mathrm{vol}(M_2) + \frac{2}{\pi^2}\int_{M_2}\mu_2^2 w \\ &= 12k, \quad k \geqslant 1. \end{aligned} \tag{4.16}$$

See Fig. 2.

Type 3, M_3: For M_3, the Hirzebruch index always vanishes. Therefore, the Euler characteristic is a multiple of 4 as

$$\chi[M_3] = \frac{3S^2}{2^6\pi^2}\,\mathrm{vol}(M_3) = 4k \geqslant 0, \tag{4.17}$$

with $k = 0$ iff $S = 0$. No duality condition can be satisfied for M_3. See Fig. 3.

Remark: For a compact oriented Lorentz manifold the Euler characteristic vanishes. From a similar argument, it is easily seen that for a compact oriented Lorentz 4-manifold the Hirzebruch index is also zero. Thus the Thorpe-Hitchin inequality holds as a special case, where both sides vanish. Therefore, we may say as follows.

Theorem 12: Let X be a compact oriented 4-manifold. Assume that X admits an Einstein metric. Then the Euler characteristic $\chi[X]$ and the Hirzebruch index $\tau[X]$ of X satisfy the Thorpe–Hitchin inequality

$$|\tau[X]| \leqslant \tfrac{2}{3}\chi[X], \tag{4.18}$$

irrespective of type of signature of the metric.

ACKNOWLEDGMENTS

The author is grateful to Professor Chiaki Ihara and Professor Mineo Ikeda for discussions and encouragements. He also wishes to thank the referee for valuable comments on the first version of the manuscript and for telling him about the work of Thorpe.

[1] J. A. Thorpe, J. Math. Mech. **18**, 779 (1969).
[2] N. J. Hitchin, J. Diff. Geom. **9**, 435 (1974).
[3] Y. Matsushita, J. Math. Phys. **22**, 979 (1981).
[4] N. Steenrod, *The Topology of Fibre Bundles* (Princeton U.P., Princeton, N.J., 1951), p. 207.
[5] S. S. Chern, Acad. Brasileira Ciencias **35**, 17 (1963).
[6] A. Z. Petrov, Sci. Notes Kazan State Univ. **114**, 55 (1954).
[7] A. Z. Petrov, *Einstein Spaces* (Pergamon, Oxford, 1969), Chap. 18.
[8] M. F. Atiyah, *Vector Fields on Manifolds* (Westdeutschen-Verlag, Cologne and Opladen, 1970).
[9] I. M. Singer and J. A. Thorpe, in "The curvature of 4-dimensional Einstein Spaces," in *Global Analysis*, paper in honor of K. Kodaira (Princeton U.P., Princeton, N.J., 1969), pp. 355–65.

41,45

Critical behavior of the two-state doubling algorithm

D. Isaacson[a)]
Rensselaer Polytechnic Institute, Troy, New York 12181

E. L. Isaacson[b)]
Rockefeller University, New York, New York 10021

D. Marchesin[c)] and P. J. Paes-Leme[d)]
Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil

(Received 26 June 1981; accepted for publication 16 September 1981)

We describe an algorithm which produces $K^2 \times K^2$ matrix approximations to the low energy part of the Schrödinger operator for N coupled oscillators. We carry out the algorithm analytically in the case $K = 2$, for arbitrary N. In particular we show explicitly in this case how the $N \to \infty$ limit exhibits critical behavior.

PACS numbers: 02.60. + y, 02.70. + d, 03.65.Ge

1. INTRODUCTION

We describe an algorithm (called the K-state doubling algorithm) which produces $K^2 \times K^2$ matrix approximations to the low energy part of the Schrödinger operator for N coupled oscillators. These operators have the form

$$H^{(N)} = H^{(N)}(q_1,\ldots,q_N) = \frac{1}{2}\sum_{j=1}^{N}(-\partial^2/\partial q_j^2 + V(q_j)) + \frac{1}{2}\sum_{j=0}^{N}\epsilon(q_{j+1}-q_j)^2,$$

where $q_0 = q_{N+1} = 0$ and $N = 2^L, L = 0,1,2,3,\cdots$. The algorithm is a variant of one we have used[1-4] to obtain numerical approximations of the lowest few eigenvalues and eigenfunctions of $H^{(N)}$. For the potential $V(\phi) = g{:}\phi^4{:}$ and for each K the method yields (in the limit $N \to \infty$) an approximation $\nu(K)$ to the mass critical exponent ν of $H^{(\infty)}$. The sequence $\{\nu(K)\}$ appears to converge rapidly as K increases, for example $\nu(2) \simeq 0.817, \nu(4) \simeq 0.969, \nu(8) \simeq 1.00$.

Our main purpose here is to carry out all the steps of the algorithm analytically for $K = 2$. In particular we calculate explicitly the spectra of the resulting $2^2 \times 2^2$ matrix approximations and their critical behavior in the limit $N \to \infty$.

The K-state procedure is motivated by two observations. First,

$$H^{(2N)} = H_1^{(N)} + H_2^{(N)} + B^{(N)},$$

where

$$H_1^{(N)} = H^{(N)}(q_1,\ldots,q_n),$$

$$H_2^{(N)} = H^{(N)}(q_{N+1},\ldots,q_{2N}),$$

and

$$B^{(N)} = -\epsilon q_N q_{N+1}.$$

Second, for many potentials including $g{:}\phi^4{:}$, $B^{(N)}$ is Kato bounded with respect to $H_1^{(N)} + H_2^{(N)}$. This suggests that the eigenfunctions of $H_1^{(N)} + H_2^{(N)}$ are good approximations to the eigenfunctions of $H^{(2N)}$ and leads to the four-step algorithm:

(i) Set $N = 1$ and $\tilde{H}^{(1)} = H^{(1)}$.

(ii) Let $P_K^{(2N)}$ be the projection onto the K^2-dimensional space spanned by the tensor products of the eigenfunctions of $\tilde{H}^{(N)}$ corresponding to its lowest K eigenvalues. Then define[5]

$$\tilde{H}^{(2N)} \equiv P_K^{(2N)}(\tilde{H}_1^{(N)} + \tilde{H}_2^{(N)} + B^{(N)})P_K^{(2N)}.$$

(iii) Compute the K lowest eigenvalues $e_1^{(2N)} \leqslant \cdots \leqslant e_K^{(2N)}$ and corresponding eigenfunctions $\psi_1^{(2N)},\ldots,\psi_K^{(2N)}$ of $\tilde{H}^{(2N)}$.

(iv) Double N and go to (ii).

We carry out these steps for $K = 2$ when $V(\phi)$ is even and the lowest two eigenvalues of $H^{(1)}$ are nondegenerate. In this case the approximate mass gaps $m_N = e_2^{(N)} - e_1^{(N)}$ depend on $V(q)$ through m_1 and $x_1 \equiv \langle\psi^{(1)}|q|\psi^{(1)}\rangle$.

In the remaining sections we show that

(1) The mass $m_\infty \equiv \lim m_N$ exists and the dimensionless ratio $m \equiv m_\infty / m_1$ depends only on the dimensionless parameter $Z_1 = m_1/\epsilon x_1^2$.

(2) The dimensionless ratio $m(Z)$ satisfies the recursion relation

$$m(Z) = f(Z)m(h(Z))$$

for two elementary functions $f(Z)$ and $h(Z)$.

(3) There is a unique positive value Z_c for which

$$Z_c = h(Z_c) \simeq 2.553\,484\,559\,6885\cdots$$

(4) $m(Z)$ exhibits critical behavior, i.e.,

$$m(Z) > 0 \quad \text{for} \quad Z > Z_c,$$

$$m(Z) \downarrow 0 \quad \text{as} \quad Z \downarrow Z_c,$$

$$m(Z) = 0 \quad \text{for} \quad Z < Z_c.$$

(5) $m(Z)$ is analytic on a complex neighborhood of $(Z_c, \infty]$.

(6) There are positive constants D_1, D_2 for which

$$D_1(Z - Z_c)^{\nu(2)} \leqslant m(Z) \leqslant D_2(Z - Z_c)^{\nu(2)},$$

[a)]Supported in part by the NSF grants MCS-80-02938 and INT-7920728.
[b)]Supported in part by the NSF grant PHY-80-01979 and an NSF Postdoctoral Fellowship.
[c)]Supported in part by the NSF grant PHY-80-01979.
[d)]Supported in part by the CNPq grant CNPq/NSF 0310.1465/80.

 0022-2488/83/010041-05$02.50

when $Z \geqslant Z_c$, where

$$\nu(2) = -\ln f(Z_c)/\ln h'(Z_c) \simeq 0.817\,266\,06\cdots.$$

We point out that the mass m_∞ produced by the two-state doubling algorithm may be regarded as an approximation to the mass gap of the exactly solvable one-dimensional spin chain solved by Stoeckly and Scalapino in Ref. 6. For that model $Z_c = 2$ and $\nu = 1$. Hence the two-state doubling approximation yields the critical point to within 30% and the mass exponent to within 20%. Therefore, we do not advocate the use of the two-state doubling algorithm to obtain accurate approximation of the critical behavior of $H(\infty)$. Rather our purpose is to illustrate the critical behavior in an exactly solvable case. Also, we show in the case $K = 2$ that the doubling algorithm can be viewed as a mapping Γ_K from a space of $K^2 \times K^2$ self-adjoint matrices into itself with the following properties: Let

$$M^{(2N)} = \Gamma_K[M^{(N)}] = \Gamma_K \circ \Gamma_K \circ \cdots \circ \Gamma_K[M^{(1)}],$$

where $M^{(N)}$ denotes a $K^2 \times K^2$ matrix representation of $\tilde{H}^{(N)} - e_1^{(N)} I$. Then $M^{(N)}$ converges as N tends to infinity to a fixed point of the mapping Γ_K. In the $g{:}\phi^4$: case there is a one-parameter family of matrices which get mapped into 0 after infinitely many iterations. All other matrices get mapped ultimately into nonzero fixed points of Γ_K with positive or zero mass gap.

2. TWO-STATE DOUBLING ALGORITHM

We carry out the algorithm for $K = 2$ and calculate explicitly the quantities of interest. When $N = 1$, $\tilde{H}^{(1)} = H^{(1)}$, and

$$\tilde{H}_i^{(1)} \psi_j^{(1)}(q_i) = e_j^{(1)} \psi_j^{(1)}(q_i),$$

for $i, j = 1,2$. Also,

$$P_2^{(2)} = \sum_{i,j=1}^{2} |\Psi_{i,j}\rangle\langle\Psi_{i,j}|,$$

where

$$\Psi_{i,j} \equiv \psi_i^{(1)}(q_1)\psi_j^{(1)}(q_2)$$

are the eigenfunctions of $\tilde{H}_1^{(1)} + \tilde{H}_2^{(1)}$. Thus

$$\begin{aligned}\tilde{H}^{(2)} &= P_2^{(2)}[\tilde{H}_1^{(1)} + \tilde{H}_2^{(1)} + B^{(1)}]P_2^{(2)} \\ &= \sum_{\substack{i,i'=1 \\ j,j'=1}}^{2} [(e_i^{(1)} + e_j^{(1)})\delta_{i,i'}\delta_{j,j'} \\ &\quad - \epsilon X_{i,i'} X_{j,j'}]|\Psi_{i,j}\rangle\langle\Psi_{i',j'}|,\end{aligned}$$

where

$$X_{i,j} \equiv \langle\psi_i^{(1)}(q)|q|\psi_j^{(1)}(q)\rangle.$$

The lowest two eigenvalues and eigenfunctions of $\tilde{H}^{(2)}$ are (ignoring the nullspace of $P_2^{(2)}$)

$$\begin{aligned}e_1^{(2)} &= e_1^{(1)} + e_2^{(1)} - ((e_2^{(1)} - e_1^{(1)})^2 + \epsilon^2 x_1^4)^{1/2}, \\ e_2^{(2)} &= e_1^{(1)} + e_2^{(1)} - \epsilon x_1^2, \\ \psi_1^{(2)} &= \psi_1^{(2)}(q_1,q_2) = a\Psi_{1,1} + b\Psi_{2,2}, \\ \psi_2^{(2)} &= \psi_2^{(2)}(q_1,q_2) = (\Psi_{1,2} + \Psi_{2,1})/\sqrt{2},\end{aligned}$$

where

$$\begin{aligned}a &= [2(1 + Z_1^2 - Z_1(1 + Z_1^2)^{1/2})]^{-1/2}, \\ b &= (-Z_1 + (1 + Z_1^2)^{1/2})/a,\end{aligned}$$

and

$$Z_1 = (e_2^{(1)} - e_1^{(1)})/\epsilon x_1^2.$$

For general N we obtain the matrix representation

$$E^{(N)} \otimes I + I \otimes E^{(N)} - \epsilon X^{(N)} \otimes X^{(N)}$$

of $\tilde{H}^{(2N)}$ in the subspace spanned by $\{\psi_i^{(N)}(q_1,\ldots,q_N)\psi_j^{(N)}(q_{N+1},\ldots,q_{2N})\}_{i,j=1}^2$.

Here

$$E^{(N)} = \begin{pmatrix} e_1^{(N)} & 0 \\ 0 & e_2^{(N)} \end{pmatrix},$$

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$X^{(N)} = \begin{pmatrix} 0 & X_N \\ X_N & 0 \end{pmatrix},$$

and

$$X_N = \langle\psi_1^{(N)}(q_1,\ldots,q_N)|q_1|\psi_2^{(N)}(q_1,\ldots,q_N)\rangle.$$

The diagonal elements of $X^{(N)}$ are zero because $V(q)$ is assumed to be even. The eigenvalues of $\tilde{H}^{(2N)}$ satisfy the recursion relations

$$\begin{aligned}e_1^{(2N)} &= e_1^{(N)} + e_2^{(N)} - ((e_1^{(N)} - e_2^{(N)})^2 + \epsilon^2 x_N^4)^{1/2}, \\ e_2^{(2N)} &= e_1^{(N)} + e_2^{(N)} - \epsilon x_N^2, \\ e_3^{(2N)} &= e_1^{(N)} + e_2^{(N)} + \epsilon x_N^2, \\ e_4^{(2N)} &= e_1^{(N)} + e_2^{(N)} + ((e_1^{(N)} - e_2^{(N)})^2 + \epsilon^2 x_N^4)^{1/2}.\end{aligned} \tag{1}$$

The main quantities of interest are the masses m_N defined by

$$m_N = e_2^{(N)} - e_1^{(N)}.$$

From (1) it follows that

$$m_{2N} = (m_N^2 + \epsilon^2 x_N^4)^{1/2} - \epsilon x_N^2, \tag{2}$$

and from the definition of x_N it follows that

$$\epsilon x_{2N}^2 = \frac{1}{2} \frac{[\epsilon x_N^2 + (-m_N + (m_N^2 + \epsilon^2 x_N^4)^{1/2})]^2}{\epsilon^2 x_N^4 + (-m_N + (m_N^2 + \epsilon^2 x_N^4)^{1/2})^2} \epsilon x_N^2. \tag{3}$$

Defining the dimensionless parameter

$$Z_N = m_N/\epsilon x_N^2, \tag{4}$$

we get

$$\begin{aligned}Z_{2N} = h(Z_N) &\equiv 2(Z_N^2 + 1)^{1/2} \\ &\quad \times ((Z_N^2 + 1)^{1/2} - 1)/((Z_N^2 + 1)^{1/2} + 1).\end{aligned} \tag{5}$$

Equations (2) and (3) can be written in terms of Z_N as follows:

$$\begin{aligned}m_{2N} &= m_N\{(1 + 1/Z_N^2)^{1/2} - 1/Z_N\} \\ &\equiv m_N f(Z_N),\end{aligned} \tag{6}$$

$$\begin{aligned}\epsilon x_{2N}^2 &= \frac{\epsilon x_N^2}{2} \frac{[1 + ((Z_N^2 + 1)^{1/2} - Z_N)]^2}{1 + ((Z_N^2 + 1)^{1/2} - Z_N)^2} \\ &\equiv \epsilon x_N^2 k(Z_N).\end{aligned} \tag{7}$$

Iterating these formulas yields for $N = 2^L$,

$$m_{2^L} = m_1 \prod_{k=0}^{L-1} f(Z_{2^k}), \tag{8}$$

$$\epsilon x_{2^L}^2 = \epsilon x_1^2 \prod_{k=0}^{L-1} k(Z_{2^k}). \tag{9}$$

3. CRITICAL BEHAVIOR

We establish certain properties of the elementary functions h, f, and k from which the critical behavior of the model follows.

Properties of $h(Z)$:

(1) $h(Z)$ is analytic in the Z-plane slit from $-i$ to i.

(2) $h(Z)$ is strictly increasing and strictly convex in $[0,\infty)$; $h'(0) = 0$, $h'(\infty) = 2$, and $h(0) = 0$.

(3) There is a unique value Z_c in $(0,\infty)$ for which $h(Z_c) = Z_c$. (We call Z_c the critical value of $Z, Z_c = 2.553\,484\,559\,688\,537\cdots$.)

It follows easily that as $N \to \infty$,

$$\begin{aligned} &Z_N \downarrow 0 \quad \text{if } Z_1 < Z_C, \\ &Z_N \equiv Z_C \quad \text{if } Z_1 = Z_C, \\ &Z_N \uparrow \infty \quad \text{if } Z_1 > Z_C. \end{aligned} \tag{10}$$

The convergence above is at least geometric.

Properties of $f(Z)$ and $k(Z)$:

(1) $f(Z)$ and $k(Z)$ are analytic in the extended Z-plane slit from $-i$ to i.

(2) $f(Z)$ is strictly increasing on $[0,\infty]$,
$f(0) = 0$, and $f(\infty) = 1$;
$k(Z)$ is strictly decreasing on $[0,\infty]$,
$k(0) = 1$, and $k(\infty) = \frac{1}{2}$.

(3) $f(Z)/k(Z) = h(Z)/Z$ as follows from (4), (5), (6), and (7). Therefore $f(Z_c) = k(Z_c)$ by Property (3) of $h(Z)$. Formulae (6) and (7) and the properties above imply that the sequences $\{m_N\}$ and $\{\epsilon x_N^2\}$ decrease monotonically as $N = 2^L \to \infty$. Therefore, they converge for all positive values of Z_1. We denote their limits by m_∞ and ϵx_∞^2, respectively.

It follows from (10), Eqs. (8) and (9), and the properties of h, f, and k that as a function of Z_1 (with m_1 fixed),

$$\begin{aligned} &m_\infty = 0 \text{ and } \epsilon x_\infty^2 > 0 \text{ for } Z_1 < Z_c, \\ &m_\infty = 0 \text{ and } \epsilon x_\infty^2 = 0 \text{ for } Z_1 = Z_c, \\ &m_\infty > 0 \text{ and } \epsilon x_\infty^2 = 0 \text{ for } Z_1 > Z_c. \end{aligned} \tag{11}$$

(See Fig. 1.)

Remark: Consider the mapping Γ from the first quadrant to itself defined by

$$\Gamma(m_1, \epsilon x_1^2) = (m_2, \epsilon x_2^2).$$

Each point on the nonnegative axes is a fixed point for Γ. Actually, these are the only ones since Γ is strictly monotone in each of its variables. The curve $m/\epsilon x^2 = Z_c$ is mapped into itself. A point on this curve gets mapped into another point on this curve closer to the origin under the action of Γ. A point lying below this curve gets mapped into a point closer to the m axis. A point lying above this curve gets mapped into a point closer to the ϵx^2 axis (see Fig. 2). The mapping Γ on the pair $(m, \epsilon x^2)$ induces a mapping Γ_2 on the renormalized $2^2 \times 2^2$ matrices $M^{(N)}$ described in the introduction.

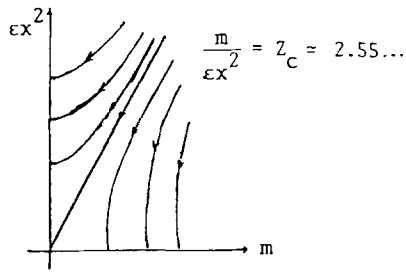


FIG. 2. Action of Γ.

4. THE HIGHER SPECTRUM

Using Eqs. (1) and (2) we obtain

$$\begin{aligned} e_3^{(2N)} - e_1^{(2N)} &= (m_N^2 + \epsilon^2 x_N^4)^{1/2} + \epsilon x_N^2 \\ &= m_{2N} + 2\epsilon x_N^2 \end{aligned}$$

and

$$\begin{aligned} e_4^{(2N)} - e_1^{(2N)} &= 2(m_N^2 + \epsilon^2 x_N^4)^{1/2} \\ &= 2m_{2N} + 2\epsilon x_N^2. \end{aligned}$$

Denoting the limits (as $N \to \infty$) of these two equations by m'_∞ and m''_∞, respectively, it follows from (11) that (see Fig. 3)

$$m'_\infty = \begin{cases} m_\infty & \text{if } Z_1 \geqslant Z_c, \\ 2\epsilon x_\infty^2 & \text{if } Z_1 \leqslant Z_c, \end{cases}$$

$$m''_\infty = \begin{cases} 2m_\infty & \text{if } Z_1 \geqslant Z_c, \\ 2\epsilon x_\infty^2 & \text{if } Z_1 \leqslant Z_c. \end{cases}$$

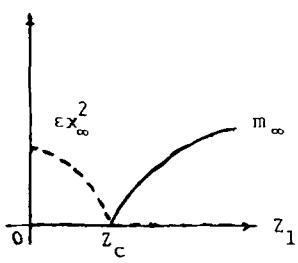


FIG. 1. Critical behavior of the mass m_∞ and ϵx_∞^2.

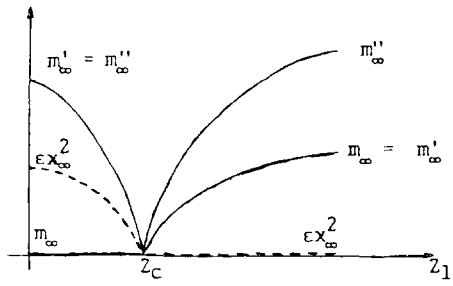


FIG. 3. The full spectrum.

5. ANALYTICITY OF THE MASS m

We define the dimensionless mass m by

$$m(Z_1) \equiv m_\infty / m_1 = \prod_{k=0}^{\infty} f(Z_{2^k}).$$

This may be rewritten as

$$m(Z) = \prod_{k=0}^{\infty} f(h^k(Z)), \tag{12}$$

where $h^k(Z) = h \circ h \circ \cdots \circ h(Z)$. The functional equation

$$m(Z) = f(Z) m(h(Z)) \tag{13}$$

follows easily.

Formula (12) can be extended to include large complex values of Z. In fact for $|Z|$ sufficiently large we have

$$f(Z) = 1 - 1/Z + O(1/Z^2),$$

so that

$$m(Z) = \prod_{k=0}^{\infty} \left\{ 1 - \frac{1}{h^k(Z)} + O\left[\left(\frac{1}{h^k(Z)} \right)^2 \right] \right\}.$$

To prove the analyticity of this product for $|Z|$ sufficiently large it suffices to establish the uniform convergence of the series

$$\sum_{k=0}^{\infty} \left| \frac{1}{h^k(Z)} + O\left[\frac{1}{h^k(Z)} \right]^2 \right|.$$

For $|Z|$ large, $|h(Z)| > (1+\delta)|Z|$ for some $\delta > 0$ so that $|h^k(Z)| > (1+\delta)^k |Z|$. Therefore the series above is bounded by

$$\sum_{k=0}^{\infty} \left| \frac{\text{const}}{h^k(Z)} \right| \leqslant \frac{1}{|Z|} \sum_{k=0}^{\infty} \frac{1}{(1+\delta)^k} < \infty.$$

Now the region of analyticity can be extended by using the functional equation (13). First m can be continued analytically from a complex neighborhood of $(Z, \infty]$, $Z > Z_c$ to a complex neighborhood of $(h^{-1}(Z), \infty]$ because f and h are analytic. Also, from (10) it follows that $h^{-k}(Z)$ converges to Z_c. Therefore m can be continued analytically to a complex neighborhood of (Z_c, ∞) by repeated applications of the functional equation.

We remark that for the $g:\phi^4$: case m_1 and ϵx_1^2 are analytic functions of the parameter g in a complex neighborhood of $(0, \infty)$. It follows that the mass $m_\infty(g) = m_1(g) m(Z_1(g))$ is analytic in a complex neighborhood of $(0, g_c)$.

6. THE EXPONENT ν

In this section we establish the following behavior of the mass near the critical point: there are positive constants D_1, D_2, and ν for which

$$m(Z) = D(Z)(Z - Z_c)^\nu \quad (Z > Z_c), \tag{14}$$

where $0 < D_1 \leqslant D(Z) \leqslant D_2 < \infty$.

The critical exponent ν can be calculated explicitly in the following heuristic way. Assume $D(Z)$ is constant. Evaluate (14) at two points Z_1 and Z_2, take logarithms, and subtract one equation from the other to obtain

$$\nu = \log \frac{m(Z_1)}{m(Z_2)} \Big/ \log \left(\frac{Z_1 - Z_c}{Z_2 - Z_c} \right).$$

By choosing $Z_2 = h(Z_1)$ and using (13) we get

$$\nu = \log \frac{m(Z_1)}{m(h(Z_1))} \Big/ \log \left(\frac{Z_1 - Z_c}{h(Z_1) - Z_c} \right)$$
$$= \log f(Z_1) \Big/ \log \left(\frac{Z_1 - Z_c}{h(Z_1) - h(Z_c)} \right).$$

The value of ν is then obtained by taking the limit of the expression above as $Z_1 \downarrow Z_c$:

$$\nu = -\log f(Z_c) / \log h'(Z_c) \tag{15}$$

(explicitly $\nu = 0.817\,266\,06\cdots$).

Given the value of ν in (15) we define $D(Z)$ by (14). By (13), $D(Z)$ satisfies the functional equation

$$D(Z) = l(Z) D(h(Z)), \tag{16}$$

where

$$l(Z) = f(Z)[(h(Z) - Z_c)/(Z - Z_c)]^\nu.$$

Properties of $l(Z)$:

(1) $l(Z) = 1 + O(Z - Z_c)$ by (15) since l is differentiable.

(2) $l(Z)$ is strictly increasing on $[0, \infty]$ because $f(Z)$ is strictly increasing and $h(Z)$ is strictly convex.

$l(0) = 0$ and $l(\infty) = 2^\nu$.

We show that

$$0 < D_1 \equiv \lim_{Z \downarrow Z_c} \inf m(Z)/(Z - Z_c)^\nu$$
$$\leqslant \lim_{Z \downarrow Z_c} \sup m(z)/(Z - Z_c)^\nu \equiv D_2 < \infty.$$

In fact we obtain numerically that $D(Z_c) \simeq 0.48\cdots$. Iterating (16) yields

$$D(h^{-N}(Z)) = D(Z) \prod_{k=1}^{N} l(h^{-k}(Z)). \tag{17}$$

Fix $Z_0 > Z_c$. Then the product in (17) converges uniformly as $N \to \infty$ for $Z \in [Z_0, h(Z_0)]$. Therefore

$$D_2 \leqslant \max\{ D(Z) : Z \in [Z_0, h(Z_0)] \} \prod_{k=1}^{\infty} l(h^{-k}(h(Z_0))),$$

$$D_1 \geqslant \min\{ D(Z) : Z \in [Z_0, h(Z_0)] \} \prod_{k=1}^{\infty} l(h^{-k}(Z_0)).$$

Remark: In an analogous way we can show that there are positive constants D_1^-, D_2^- and ν' for which

$$m'(Z) = D^-(Z)(Z_c - Z)^{\nu'} \quad (Z < Z_c),$$

where $0 < D_1^- \leqslant D^-(Z) \leqslant D_2^- < \infty$ and

$$\nu' = -\log k(Z_c)/\log h'(Z_c).$$

From Property (3) of f and k it follows that

$$\nu' = \nu.$$

[1] D. Isaacson, D. Marchesin, and P. J. Paes-Leme, Int. J. Eng. **18**, 341–349 (1980).

[2] D. Isaacson, E. L. Isaacson, D. Marchesin, and P. J. Paes-Leme, "Numerical Analysis of Spectral Properties of Coupled Oscillator Schrödinger Operators I. Single and Double Well Anharmonic Oscillators," Math. Comp. (to appear).

[3] D. Isaacson, E. L. Isaacson, D. Marchesin, and P. J. Paes-Leme, "Numeri-

cal Analysis of Spectral Properties of Coupled Oscillator Schrödinger Operators II. Two Coupled Anharmonic Oscillators," SIAM J. Numer. Anal. (to appear).
[4]D. Isaacson, E. L. Isaacson, D. Marchesin, and P. J. Paes-Leme, "Numerical Analysis of Spectral Properties of Coupled Oscillator Schrödinger Operators III. The Doubling Algorithm," in preparation.
[5]We do not consider the null space of $P_K^{(2N)}$ to be an eigenspace of $\tilde{H}^{(2N)}$.
[6]B. Stoeckly and D. J. Scalapino, Phys. Rev. B **11**, 205 (1975).

Direct approach to the periodic solutions of the multidimensional sine–Gordon equation

J. Zagrodziński
Instytut Fizyki P.A.N., 02-668 Warsaw, Poland

(Received 22 July 1981; accepted for publication 4 September 1981)

A number of identities for multidimensional theta functions and their derivatives are derived. Application to the nonlinear partial differential equations is exemplified for the sine–Gordon equation. In consequence, the multidimensional sine–Gordon equation can be reduced to a functional equation, and then to a set of algebraic equations. Several particular cases are also discussed.

PACS numbers: 03.40.Kf

1. INTRODUCTION

The problems associated with nonlinear partial differential equations (NL-PDE) have been in the center of interest of theoretical physics for more than fifteen years. This interest is still growing. More and more problems arise and await solution, mainly in nonlinear optics, plasma physics, superconductivity, quantum field theory, and the like.

There are, however, some problems which seem to be particularly important for the further development of the theory and its application in physics. In our opinion, the multidimensional periodic solutions of NL-PLDE's particularly of the sine–Gordon (sG) or Korteweg–de Vries (KdV) equation, can well be considered among these problems.

There is a vast literature on the subject of periodic solutions of NL-PDE's,[1–7] mainly inspired by the inverse scattering method.

As is known, in the case of KdV or sG equations with the imposed requirement of periodic solutions, the application of the inverse scattering formalism leads to expressions involving an abstract theta function (Θ-f). However, as yet only the periodic solutions in $(1+1)$-dimensional space have been thoroughly investigated.

In the present paper, we intend to give an insight into the question of multidimensional solutions of the sG equation, generalizing to some extent the well-known results for the $(1+1)$-dimensional case.

The outline of the paper is as follows. In the first part we derive some fundamental identities for the abstract multidimensional Θ-f and its derivatives. Next, using the previously-derived relations, we formulate a few theorems that reduce the question of the solution of a multidimensional sG equation to a purely algebraic problem. Some preliminary results concerning these problems were already announced briefly in Ref. 8.

We want to emphasize that our approach is considerably different from that presented in papers on abelian integrals and their application to the Riemann Θ-f in soliton theory. We are interested here in algebraic identities, recursion formulas for the Θ-f and, most of all, for its second derivatives, which would be useful directly in the analysis of the multidimensional sG equation.

2. IDENTITIES FOR THE THETA FUNCTION

We provide the following definitions:

i. C^g—the g-dimensional complex vector space;

ii. Z^g—the g-dimensional lattice, i.e., the set of g-dimensional vectors $\mathbf{k}$ with integer (real) components k_i ($k_i = 0, \pm 1, \pm 2, \ldots, i = 1,2,\ldots g$); and

iii. D^g—the g-dimensional "die" (cube), i.e., the set of g-dimensional vectors $\boldsymbol{\epsilon}$ with components ϵ_i taking only two values, 0 or 1 ($\epsilon_i = 0, 1, i = 1,2,\ldots,g$).

We adopt here the following definition of an abstract or multidimensional Θ-f, of argument $\mathbf{z}\in C^g$ [1,3–7,9]:

$$\Theta(\mathbf{z}|B) = \sum_{\mathbf{k}\in Z^g} \exp[i\pi(2\mathbf{z}\cdot\mathbf{k} + \mathbf{k}\cdot B\,\mathbf{k})], \tag{1}$$

where B is the g-dimensional symmetric complex matrix, $\mathbf{z}\cdot\mathbf{k}$ denotes the scalar product

$$\mathbf{z}\cdot\mathbf{k} = \sum_{i=1}^{g} z_i k_i,$$

and the sum is over the lattice Z^g:

$$\sum_{k\in Z^g} = \sum_{k_1=-\infty}^{\infty} \cdots \sum_{kg=-\infty}^{\infty}.$$

The series (1) will be convergent if there exists $C>0$, such that

$$\mathrm{Im}(\mathbf{k}\cdot B\,\mathbf{k}) \geqslant C(\mathbf{k}\cdot\mathbf{k}). \tag{2}$$

Matrix B is known as the period's matrix, if the Θ-f is defined by means of abelian integrals. Although we do not proceed in this way, our results can supply some information to the analysis of Θ-f from the abelian-integral point of view. An exhaustive analysis of the Θ-f in terms of abelian integrals and differentials on a Riemann surface can be found in the previously cited papers. Some algebraic properties of the Θ-f 's are discussed in Krazer's monograph[9] devoted to the Θ-fs with characteristics, which form a broader class than the Θ-fs considered here (cf. also Ref. 4).

The function defined by (1) is called by several authors an abstract Θ-f, while others prefer to call it multi-dimensional Θ-f. Since, in fact, it is a scalar function of g-independent arguments, the first designation seems to be less legitimate. However, we use both terms equally.

 0022-2488/83/010046-07$02.50

The Θ-f given by (1) is quasiperiodic; it has the following properties[1,5,7,9]:

$$\Theta(\mathbf{z}+\mathbf{q}|B)=\Theta(\mathbf{z}|B),\quad \mathbf{q}\in Z^g, \tag{3}$$

$$\Theta(\mathbf{z}+\mathbf{B}^j|B)=\{\exp[-i\pi(2z_j+B_{jj})]\}\Theta(\mathbf{z}|B), \tag{4}$$

$$\Theta(-\mathbf{z}|B)=\Theta(\mathbf{z}|B), \tag{5}$$

where $\mathbf{B}^j$ is the jth column of matrix B and B_{jj} is ist j,j element.

Applying (3) repeatedly we have for each $\mathbf{q}\in Z^g$ and $\mathbf{z}\in C^g$

$$\Theta(\mathbf{z}+B\mathbf{q}|B)=\exp[-i\pi(2\mathbf{z}\cdot\mathbf{q}+\mathbf{q}\cdot B\mathbf{q})]\Theta(\mathbf{z}|B). \tag{6}$$

One can prove also that

$$[\Theta(z|B)]^2$$

$$=\sum_{\epsilon\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]\Theta(B\boldsymbol{\epsilon}|2B)\Theta(2\mathbf{z}+B\boldsymbol{\epsilon}|2B). \tag{7}$$

Proof:

$$\Theta^2(\mathbf{z}|B)$$

$$=\sum_{\mathbf{m}\in Z^g}\sum_{\mathbf{n}\in Z^g}\exp\{i\pi[2(\mathbf{m}+\mathbf{n})\cdot\mathbf{z}+\mathbf{m}\cdot B\mathbf{m}+\mathbf{n}\cdot B\mathbf{n}]\}$$

$$=\sum_{\mathbf{q}\in Z^g}\sum_{\mathbf{n}\in Z^g}\exp\{i2\pi[\mathbf{q}\cdot(\mathbf{z}+B\mathbf{q})+(\mathbf{n}-\mathbf{q})\cdot B\mathbf{n}]\}$$

$$=\sum_{\mathbf{q}\in Z^g}\exp[i\pi(2\mathbf{z}\cdot\mathbf{q}+\mathbf{q}\cdot B\mathbf{q})]\Theta(B\mathbf{q}|2B), \tag{8}$$

where the sum is over the lattice Z^g.

Substituting $\mathbf{q}=2\mathbf{p}+\boldsymbol{\epsilon}$, where $\mathbf{p}\in Z^g$, $\boldsymbol{\epsilon}\in D^g$, Eq. (8) takes the form

$$\sum_{\epsilon\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})\Theta(B\boldsymbol{\epsilon}|2B)]$$

$$\times\sum_{p\in Z^g}\exp\{i\pi[2\mathbf{p}\cdot(2\mathbf{z}+B\boldsymbol{\epsilon})+\mathbf{p}\cdot B\mathbf{p}]\}$$

$$=\sum_{\epsilon\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]\Theta(B\boldsymbol{\epsilon}|2B)$$

$$\times\Theta(2\mathbf{z}+B\boldsymbol{\epsilon}|2B). \tag{9}$$

Observe that the sum is now over the "die" D^g and thus it contains only the finite number of elements (2^g).

For the case $\mathbf{z}=\mathbf{m}\in Z^g$ we have

$$\Theta^2(\tfrac{1}{2}\mathbf{m}|B)=\sum_{\epsilon\in D^g}(-1)^{(\mathbf{m}\cdot\boldsymbol{\epsilon})}\exp(i\pi\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})\Theta^2(B\boldsymbol{\epsilon}|2B). \tag{10}$$

Relations (7), and also (10), can be inverted. If we change the variables, $\mathbf{z}\to\mathbf{z}+\frac{1}{2}\boldsymbol{\delta}$, $(\boldsymbol{\delta}\in D^g)$ in (7), multiply by $(-1)^{(\boldsymbol{\epsilon}'\cdot\boldsymbol{\delta})}$, and then sum over $\boldsymbol{\delta}$, since

$$\sum_{\delta\in D^g}(-1)^{\boldsymbol{\epsilon}\cdot\boldsymbol{\delta}}(-1)^{\boldsymbol{\epsilon}'\cdot\boldsymbol{\delta}}=2^g\delta_{\boldsymbol{\epsilon},\boldsymbol{\epsilon}'}, \tag{11}$$

(where $\delta_{\boldsymbol{\epsilon},\boldsymbol{\epsilon}'}$ is the Kronecker delta), we obtain

$$\Theta(2\mathbf{z}+B\boldsymbol{\epsilon}|2B)=\frac{\exp[-i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]}{2^g\Theta(B\boldsymbol{\epsilon}|2B)}$$

$$\times\sum_{\delta\in D^g}(-1)^{\boldsymbol{\epsilon}\cdot\boldsymbol{\delta}}\Theta^2(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\delta}|B), \tag{12}$$

or

$$\Theta(2\mathbf{z}|2B)=2^{-g}\Theta^{-1}(0|2B)\sum_{\delta\in D^g}\Theta^2(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\delta}|B). \tag{13}$$

Here the sum again is over the "die" D^g.

Substituting $\mathbf{z}=0$ or $\mathbf{z}=\boldsymbol{\epsilon}=0$, (12) yields, respectively,

$$\Theta(B\boldsymbol{\epsilon}|2B)=2^{-g/2}\left[\sum_{\delta\in D^g}(-1)^{\boldsymbol{\epsilon}\cdot\boldsymbol{\delta}}\Theta^2(\tfrac{1}{2}\boldsymbol{\delta}|B)\right]^{1/2}$$

$$\times\exp[-i\pi\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon}/2], \tag{14}$$

or

$$\Theta(0|2B)=2^{-g/2}\left[\sum_{\delta\in D^g}\Theta^2(\tfrac{1}{2}\boldsymbol{\delta}|B)\right]^{1/2}. \tag{15}$$

For the case of only one variable all the formulas (7)–(15) simplify to already-known results.[10–12] For example, relation (13) represents the multivariable variant of the Landen transformation.[10]

One can prove[9,12] that

$$\sum_{\mathbf{n}\in Z^g}\exp[i\pi(2\mathbf{n}\cdot\mathbf{z}+\mathbf{n}\cdot B\mathbf{n})]$$

$$=(-i)^{-g/2}(\det B)^{-1/2}\sum_{\mathbf{n}\in Z^g}\exp[-i\pi(\mathbf{z}\cdot\mathbf{n})\cdot B^{-1}(\mathbf{z}+\mathbf{n})], \tag{16}$$

which in Θ-f language is

$$\Theta(\mathbf{z}|B)=(-i)^{-g/2}(\det B)^{-1/2}\exp[i\pi\mathbf{z}\cdot B^{-1}\mathbf{z}]$$

$$\times\Theta(B^{-1}\mathbf{z}|-B^{-1}), \tag{17}$$

and is the multivariable version of the modular transformation.[8–10,12]

Combining (17) and (13), we obtain another useful representation:

$$\Theta(\mathbf{z}|B)\Theta(0|B)=\sum_{\epsilon\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]\Theta^2(\mathbf{z}+B\boldsymbol{\epsilon}|2B), \tag{18}$$

or

$$\Theta(\mathbf{z}|B)=\left[\sum_{\epsilon\in D^g}\exp(i\pi\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})\Theta^2(B\boldsymbol{\epsilon}|2B)\right]^{-1/2}$$

$$\times\sum_{\epsilon\in D^g}\{\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]\Theta^2(\mathbf{z}+B\boldsymbol{\epsilon}|2B)\}. \tag{19}$$

Similarly, one can derive the relation between Θ-f 's of matrices B and $4B$. Indeed, substituting $\mathbf{k}=2\mathbf{n}+\boldsymbol{\epsilon}$, $\mathbf{n}\in Z^g$, $\boldsymbol{\epsilon}\in D^g$, by definition (1) we have

$$\Theta(\mathbf{z}|B)$$

$$=\sum_{\mathbf{n}\in Z^g}\sum_{\epsilon\in D^g}\exp\{i\pi[2(2\mathbf{n}+\boldsymbol{\epsilon})\cdot\mathbf{z}+(2\mathbf{n}+\boldsymbol{\epsilon})\cdot B(2\mathbf{n}+\boldsymbol{\epsilon})]\}$$

$$=\sum_{\epsilon\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]$$

$$\times\sum_{\mathbf{n}\in Z^g}\exp\{i4\pi[(\mathbf{z}+B\boldsymbol{\epsilon})\cdot\mathbf{z}+\mathbf{n}\cdot B\mathbf{n}]\}. \tag{20}$$

Thus finally

$$\Theta(\mathbf{z}|B)=\sum_{\epsilon\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]\Theta(2(\mathbf{z}+B\boldsymbol{\epsilon})|4B). \tag{21}$$

Shifting the argument $\mathbf{z}\to\mathbf{z}+\frac{1}{2}\boldsymbol{\delta}$, multiplying by $(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\epsilon}'}$, and making use of (11), relation (21) can be inverted:

$$2^g\Theta(2(\mathbf{z}+B\boldsymbol{\epsilon}')|4B)\exp[i\pi(2\boldsymbol{\epsilon}'\cdot\mathbf{z}+\boldsymbol{\epsilon}'\cdot B\boldsymbol{\epsilon}')]$$
$$=\sum_{\boldsymbol{\delta}\in D^g}(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\epsilon}'}\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\delta}|B). \tag{22}$$

Thus for $\boldsymbol{\epsilon}'=0$,

$$\Theta(2\mathbf{z}|4B)=2^{-g}\sum_{\boldsymbol{\delta}\in D^g}\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\delta}|B), \tag{23}$$

and

$$\Theta(\mathbf{z}|B)=2^{-g}\sum_{\boldsymbol{\delta}\in D^g}\Theta(\tfrac{1}{2}(\mathbf{z}+\boldsymbol{\delta})|\tfrac{1}{4}B). \tag{24}$$

Similarly, from (6), after an appropriate manipulation of indices, the product of the Θ-f 's takes the form

$$\Theta(\mathbf{z}_1|B)\Theta(\mathbf{z}_2|B)$$
$$=\sum_{\mathbf{n}\in Z^g}\exp[i\pi(2\mathbf{n}\cdot\mathbf{z}_2+\mathbf{n}\cdot B\mathbf{n})]\Theta(\mathbf{z}_1+\mathbf{z}_2+B\mathbf{n}|2B)$$
$$=\sum_{\boldsymbol{\epsilon}\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}_2+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]\Theta(\mathbf{z}_2-\mathbf{z}_1+B\boldsymbol{\epsilon}|2B)$$
$$\times\Theta(\mathbf{z}_1+\mathbf{z}_2+B\boldsymbol{\epsilon}|2B), \tag{25}$$

or

$$\Theta(\mathbf{z}_1|B)\Theta(\mathbf{z}_2|B)$$
$$=\sum_{\boldsymbol{\epsilon}\in D^g}\exp[i\pi(2\boldsymbol{\epsilon}\cdot\mathbf{z}_1+\boldsymbol{\epsilon}\cdot B\boldsymbol{\epsilon})]$$
$$\times\Theta(\mathbf{z}_1-\mathbf{z}_2+B\boldsymbol{\epsilon}|2B)\Theta(\mathbf{z}_1+\mathbf{z}_2+B\boldsymbol{\epsilon}|2B). \tag{26}$$

The first sum in (25) is over the lattice, while the second one is over the "die"; thus it is finite. If $\mathbf{z}_1=\mathbf{z}_2$, relation (26) reduces to (7).

Then comparing relations (25) and (8), it follows that

$$\sum_{\mathbf{n}\in Z^g}[\Theta(2\mathbf{z}+B\mathbf{n}|2B)-\Theta(B\mathbf{n}|2B)]$$
$$\times\exp[i\pi(2\mathbf{n}\cdot\mathbf{z}+\mathbf{n}\cdot B\mathbf{n})]\equiv 0, \tag{27}$$

and next

$$\sum_{\mathbf{n}\in Z^g}\exp(i\pi\mathbf{n}\cdot B\mathbf{n})[\partial_{z_i}\Theta(\mathbf{z}|2B)]_{\mathbf{z}=B\mathbf{n}}\equiv 0, \tag{28}$$

where here and henceforth the derivatives with respect to z_i ($i=1,2,\ldots,g$) are designated by ∂_{z_i}.

Next we derive a few indentities relating the Θ-fs of parameters B and $2B$ in the spirit of the Landen transformation.[10]

First, substituting $\mathbf{z}_i\to\mathbf{z}_i+\frac{1}{2}\boldsymbol{\delta}$ in (25) and summing over $\boldsymbol{\delta}$, by (11) we have the identity

$$\Theta(\mathbf{z}_2-\mathbf{z}_1|2B)\Theta(\mathbf{z}_2+\mathbf{z}_1|2B)$$
$$=2^{-g}\sum_{\boldsymbol{\delta}\in D^g}\Theta(\mathbf{z}_1+\tfrac{1}{2}\boldsymbol{\delta}|B)\Theta(\mathbf{z}_2+\tfrac{1}{2}\boldsymbol{\delta}|B), \tag{29}$$

for arbitrary $\mathbf{z}_1$ and $\mathbf{z}_2$.

In particular, putting $\mathbf{z}_2=\mathbf{z}_1+\frac{1}{2}\boldsymbol{\mu}$, $\boldsymbol{\mu}\in D^g$, we obtain the relation

$$\Theta(2\mathbf{z}+\tfrac{1}{2}\boldsymbol{\mu}|2B)\Theta(\tfrac{1}{2}\boldsymbol{\mu}|2B)$$
$$=2^{-g}\sum_{\boldsymbol{\nu}\in D^g}\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}|B)\Theta(\mathbf{z}+\tfrac{1}{2}(\boldsymbol{\nu}+\boldsymbol{\mu})|B), \tag{30}$$

which for $\boldsymbol{\mu}=0$ reduces to (13) or (15). Now substituting $\mathbf{z}\to\mathbf{z}+\frac{1}{4}\boldsymbol{\mu}$ in (13), and comparing with (30) we get the identity

$$\Theta(\tfrac{1}{2}\boldsymbol{\mu}|2B)\sum_{\boldsymbol{\delta}\in D^g}\Theta^2(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\delta}+\tfrac{1}{4}\boldsymbol{\mu}|B)$$
$$=\Theta(0|2B)\sum_{\boldsymbol{\nu}\in D^g}\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}|B)\Theta(\mathbf{z}+\tfrac{1}{2}(\boldsymbol{\mu}+\boldsymbol{\nu})|B). \tag{31}$$

On the other hand, substituting $z_1=0$ in (29), we have

$$\Theta^2(\mathbf{z}|2B)=2^{-g}\sum_{\boldsymbol{\delta}\in D^g}\Theta(\tfrac{1}{2}\boldsymbol{\delta}|B)\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\delta}|B), \tag{32}$$

which is the representation of the square of the Θ-f in terms of the sum of linearly independent functions $\Theta(\mathbf{z}+\frac{1}{2}\boldsymbol{\delta}|B)$, (cf. Appendix B).

Similarly, relation (30) allows us to determine uniquely $\Theta(2\mathbf{z}|2B)$ by $\Theta(\mathbf{z}+\frac{1}{2}\boldsymbol{\nu}|B)$. The inverse relation is not unique, however. Indeed, multiplying (30) by $(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\mu}}$ and summing over $\boldsymbol{\mu}\in D^g$, we obtain

$$\sum_{\boldsymbol{\mu}\in D^g}\sum_{\boldsymbol{\nu}\in D^g}(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\mu}}\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}|B)\Theta(\mathbf{z}+\tfrac{1}{2}(\boldsymbol{\mu}+\boldsymbol{\nu})|B)$$
$$=2^g\sum_{\boldsymbol{\mu}\in D^g}(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\mu}}\Theta(\tfrac{1}{2}\boldsymbol{\mu}|2B)\Theta(2\mathbf{z}+\tfrac{1}{2}\boldsymbol{\mu}|2B). \tag{33}$$

Substituting $\boldsymbol{\mu}\to\boldsymbol{\mu}+\boldsymbol{\nu}$, the left-hand side of (33) reduces to a perfect square and (33) becomes

$$\left[\sum_{\boldsymbol{\nu}\in D^g}(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\nu}}\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}|B)\right]^2$$
$$=2^g\sum_{\boldsymbol{\mu}\in D^g}(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\mu}}\Theta(\tfrac{1}{2}\boldsymbol{\mu}|2B)\Theta(2\mathbf{z}+\tfrac{1}{2}\boldsymbol{\mu}|2B). \tag{34}$$

Taking the square root, by the standard technique we find (for: $|\mathrm{Re}(z_i)|<\frac{1}{2}$)

$$\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}|B)$$
$$=2^{-g/2}\sum_{\boldsymbol{\delta}\in D^g}\tau_{\boldsymbol{\delta}}(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\nu}}\left[\sum_{\boldsymbol{\mu}\in D^g}(-1)^{\boldsymbol{\delta}\cdot\boldsymbol{\mu}}\right.$$
$$\left.\times\Theta(\tfrac{1}{2}\boldsymbol{\mu}|2B)\Theta(2\mathbf{z}+\tfrac{1}{2}\boldsymbol{\mu}|2B)\right]^{1/2}, \tag{35}$$

where $\tau_{\boldsymbol{\delta}}$ are the square roots of unity (arbitrarily chosen), i.e., $\tau_{\boldsymbol{\delta}}$ can take the values ∓ 1.

The choice of the set of $\tau_{\boldsymbol{\delta}}$ determines the variety of solutions $\Theta(\mathbf{z}|B)$ and the number of different solutions is 2^g, where g is the dimension of the die D^g. Thus putting $\tau_{\boldsymbol{\delta}}=(-1)^{\tau\cdot\boldsymbol{\delta}}$, we can denote these different solutions in (35) by $\Theta_\tau(\mathbf{z}+\frac{1}{2}\boldsymbol{\nu}|B)$.

It is also possible to change the shift of an argument of Θ-f to the shift in the B domain. Let us consider $\Theta(\mathbf{z}|B\mp Q)$, where Q denotes a symmetric matrix whose elements q_{ij} are (real) integers. We have

$$\Theta(\mathbf{z}|B+Q)$$
$$=\sum_{\mathbf{n}\in Z^g}(-1)^{\mathbf{n}\cdot Q\mathbf{n}}\exp[i\pi(2\mathbf{z}\cdot\mathbf{n}+\mathbf{n}\cdot B\mathbf{n})]$$
$$=\sum_{\mathbf{n}\in Z^g}(-1)^{\mathbf{q}\cdot\mathbf{n}}\exp[i\pi(2\mathbf{z}\cdot\mathbf{n}+\mathbf{n}\cdot B\mathbf{n})]$$
$$=\Theta(\mathbf{z}+\tfrac{1}{2}\mathbf{q}|B), \tag{36}$$

where $\mathbf{q}=(q_{11},g_{22},\ldots q_{gg})$.

These formulas allow us to confine the variety of B matrices to a set of matrices whose elements have real parts that

satisfy

$$|\mathrm{Re}\, b_{ij}| \leqslant 1/2, \tag{37}$$

because one can always subtract an integer.

3. APPLICATION TO MULTIDIMENSIONAL sG EQUATION

In the application to NL-PDE we require the relations involving partial derivatives of the Θ-f. We report here a few formulas involving the first and the second derivatives.

Theorem I: For any δ (considered here as an arbitrary parameter, but in further application usually choosen as $\delta \in D^g$),

$$\partial_{z_i} \ln[\Theta(\mathbf{z} - \tfrac{1}{4}\delta|B)\Theta^{-1}(\mathbf{z} + \tfrac{1}{4}\delta|B)] = \sum_{\mu \in D^g} \Psi^i_\mu \frac{\Theta^2(\mathbf{z} + \frac{1}{2}\mu|B)}{\Theta^2(\mathbf{z}|B)}, \tag{38}$$

and the coefficients $\Psi^i_\mu(\delta)$ are independent of $\mathbf{z}$ and given by

$$\Psi^i_\mu(\delta) = \sum_{\epsilon \in D^g} (-1)^{\epsilon\cdot\mu} \times \frac{\partial_{w_i}[\Theta(2\mathbf{w} + B\epsilon|2B)\exp(i2\pi\mathbf{w}\cdot\epsilon)]|_{\mathbf{w} = -\frac{1}{4}\delta}}{2^g\Theta(B\epsilon|2B)}, \tag{39}$$

which means that $\Psi^i_\mu(\delta)$ are expressed as the finite sum of the derivatives of the Θ-f 's, but now at fixed points determined by the set $B\epsilon$.

Proof: The numerator of the left-hand side of (38), taking into account (25), can be written as

$$\begin{aligned}
&\partial_{w_i}[\Theta(\mathbf{z} - \mathbf{w}|B)\Theta(\mathbf{z} + \mathbf{w}|B)]_{\mathbf{w} = -\frac{1}{4}\delta} \\
&\quad = \partial_{w_i}(\Theta(2\mathbf{w} + B\epsilon|2B)\Theta(2\mathbf{z} + B\epsilon|2B) \\
&\quad \times \exp\{i\pi[2\epsilon\cdot(\mathbf{z} + \mathbf{w}) + \epsilon\cdot B\epsilon]\})_{\mathbf{w} = -\frac{1}{4}\delta} \\
&\quad = \sum_{\epsilon \in D^g} \{\partial_{w_i}[\Theta(2\mathbf{w} + B\epsilon|2B)\exp(i2\pi\epsilon\cdot\mathbf{w})]\}_{\mathbf{w} = -\frac{1}{4}\delta} \\
&\quad \times \Theta(2\mathbf{z} + B\epsilon|2B)\exp[i\pi(2\epsilon\cdot\mathbf{z} + \epsilon\cdot B\epsilon)].
\end{aligned} \tag{40}$$

Next, using formula (12), relation (40) becomes

$$\sum_{\mu \in D^g} \sum_{\epsilon \in D^g} 2^{-g}(-1)^{\epsilon\cdot\mu}\Theta^{-1}(B\epsilon|2B)\Theta^2(\mathbf{z} + \tfrac{1}{2}\mu|B) \times \partial_{w_i}[\Theta(2\mathbf{w} + B\epsilon|2B)\exp(i2\pi\epsilon\cdot\mathbf{w})]_{\mathbf{w} = -\frac{1}{4}\delta}. \tag{41}$$

Relation (38) is useful, for example, if one wants to discuss periodic solutions of the KdV equation. Obviously, the form of coefficients $\Psi^i_\mu(\delta)$ is not determined uniquely and using the previously given identities, one can easily find other equivalent forms.

Let us now derive the formula for the second (mixed) derivative of the Θ-f.

Theorem II:

$$\partial_{z_i}\partial_{z_j} \ln[\Theta(\mathbf{z}|B)] = \sum_{\delta \in D^g} \Omega^{ij}_\delta \frac{\Theta^2(\mathbf{z} + \frac{1}{2}\delta|B)}{\Theta^2(\mathbf{z}|B)}, \tag{42}$$

where the coefficients Ω^{ij}_δ are independent of the variable $\mathbf{z}$ and are given by

$$\Omega^{ij}_\delta = 2^{-(g+1)} \sum_{\epsilon \in D^g} (-1)^{\epsilon\cdot\delta} \times \frac{\partial_{w_i}\partial_{w_j}[\Theta(2\mathbf{w} + B\epsilon|2B)\exp(i2\pi\epsilon\cdot\mathbf{w})]|_{\mathbf{w}=0}}{\Theta(B\epsilon|2B)}. \tag{43}$$

Thus, similarly, as in the previous formula, the Ω^{ij}_δ are given in the form of a finite sum (over die D^g) of the Θ-f derivatives, again at fixed points. Another, equivalent form of the coefficient Ω^{ij}_δ can be found in Ref. 8.

Proof: Using the definition (1) let us consider the form

$$\Theta(\mathbf{z}|B)\partial_{z_i}\partial_{z_j}\Theta(\mathbf{z}|B) - \partial_{z_i}\Theta(\mathbf{z}|B)\partial_{z_j}\Theta(\mathbf{z}|B) = \tfrac{1}{2}\{\partial_{w_i}\partial_{w_j}[\Theta(\mathbf{z} + \mathbf{w}|B)\Theta(\mathbf{z} - \mathbf{w}|B)]\}_{\mathbf{w}=0}. \tag{44a}$$

By the relation (25), as in Theorem I, the right-hand side of (44a) can be written

$$\begin{aligned}
&\partial_{w_i}\partial_{w_j}[\Theta(\mathbf{z} - \mathbf{w}|B)\Theta(\mathbf{z} + \mathbf{w}|B)] \\
&\quad = \sum_{\epsilon \in D^g} \partial_{w_i}\partial_{w_j}[\Theta(2\mathbf{w} + B\epsilon|2B)\exp(i2\pi\mathbf{w}\cdot\epsilon)] \\
&\quad \times \Theta(2\mathbf{z} + B\epsilon|2B)\exp[i\pi(2\mathbf{z}\cdot\epsilon + \epsilon\cdot B\epsilon)].
\end{aligned} \tag{44b}$$

Thus, applying (12) we finally find

$$\Theta(\mathbf{z}|B)\partial_{z_i}\partial_{z_j}\Theta(\mathbf{z}|B) - \partial_{z_i}\Theta(\mathbf{z}|B)\partial_{z_j}\Theta(\mathbf{z}|B) = \sum_{\delta \in D^g} \sum_{\epsilon \in D^g} \left\{(-1)^{\delta\cdot\epsilon} \frac{\partial_{w_i}\partial_{w_j}[\Theta(2\mathbf{w} + B\epsilon|2B)\exp(i2\pi\epsilon\cdot\mathbf{w})]|_{\mathbf{w}=0}}{2^{g+1}\Theta(B\epsilon|2B)}\right\} \Theta^2(\mathbf{z} + \tfrac{1}{2}\delta|B). \tag{45}$$

Since the left-hand side of (45) represents the numerator of

$$\partial_{z_i}\partial_{z_j} \ln[\Theta(\mathbf{z}|B)],$$

relation (45) proves the theorem (42), yielding the coefficients in form (43). In this manner the second derivative of $\ln \Theta$ reduces to the sum of 2^g elements.

The collection of identities for the Θ-f 's can be supplemented by another one, which will become important later on. The following identity holds:

Theorem III:

$$\begin{aligned}
&[\Theta(\mathbf{z} + \tfrac{1}{2}\epsilon|B)\Theta(\mathbf{z}|B)]^2 \\
&\quad = 2^{-4g} \sum_{\delta,\nu,\mu,\eta \in D^g} \Theta(\tfrac{1}{4}(\epsilon \pm \delta) + \tfrac{1}{2}\eta|\tfrac{1}{4}B) \\
&\quad \times \Theta(\tfrac{1}{4}(\epsilon \pm \delta) + \tfrac{1}{2}(\nu + \mu + \eta)|\tfrac{1}{4}B) \\
&\quad \times \Theta(\tfrac{1}{4}\delta + \tfrac{1}{2}\mu|\tfrac{1}{4}B)\Theta(\mathbf{z} + \tfrac{1}{4}\delta + \tfrac{1}{2}\nu|\tfrac{1}{4}B),
\end{aligned} \tag{46}$$

where $\epsilon \in D^g$ and is fixed. The signs in the first two terms on the right-hand side can be either both positive or both nega-

tive. Relation (46) can be proved by applying the relation (29) three times. Indeed,

$$\text{r.h.s. of (46)} = 2^{-3g} \sum_{\delta,\nu,\mu\in D^g} \Theta(\tfrac{1}{2}(\epsilon \pm \delta + \mu + \nu)|\tfrac{1}{2}B) \times \Theta(\tfrac{1}{2}(\nu+\mu)|\tfrac{1}{2}B)\Theta(\tfrac{1}{4}\delta + \tfrac{1}{2}\mu|\tfrac{1}{4}B) \times \Theta(z + \tfrac{1}{4}\delta + \tfrac{1}{2}\nu|\tfrac{1}{4}B). \quad (47a)$$

Changing the indices $\mu + \nu \to \mu$,

$$\begin{aligned}\text{r.h.s. (47a)} &= 2^{-3g} \sum_{\delta,\nu,\mu\in D^g} \Theta(\tfrac{1}{2}(\epsilon+\delta+\mu)|\tfrac{1}{2}B) \times \Theta(\tfrac{1}{2}\mu|\tfrac{1}{2}B)\Theta(\tfrac{1}{4}\delta + \tfrac{1}{2}(\mu+\nu)|\tfrac{1}{4}B) \times \Theta(z + \tfrac{1}{4}\delta + \tfrac{1}{2}\nu|\tfrac{1}{4}B) \\ &= 2^{-2g} \sum_{\delta,\mu\in D^g} \Theta(\tfrac{1}{2}(\epsilon+\delta+\mu)|\tfrac{1}{2}B)\Theta(\tfrac{1}{2}\mu|\tfrac{1}{2}B) \times \Theta(z + \tfrac{1}{2}(\delta+\mu)|\tfrac{1}{2}B)\Theta(z + \tfrac{1}{2}\mu|\tfrac{1}{2}B) \\ &= 2^{-g} \sum_{\mu\in D^g} \Theta(z + \tfrac{1}{2}\epsilon|B)\Theta(z - \tfrac{1}{2}\epsilon|B) \times \Theta(\tfrac{1}{2}\mu|\tfrac{1}{2}B)\Theta(z + \tfrac{1}{2}\mu|\tfrac{1}{2}B) \\ &= \Theta^2(z + \tfrac{1}{2}\epsilon|B)\Theta^2(z|B), \end{aligned} \quad (47b)$$

which concludes the proof.

In Appendix B we show that the functions $\Theta(z + \frac{1}{4}\delta + \frac{1}{2}\mu|\frac{1}{4}B)$ for $\delta, \mu \in D^g$ form a set of linearly independent functions (for fixed B). Thus, relation (46) represents the expansion of $[\Theta(z + \frac{1}{2}\epsilon|B)\Theta(z|B)]^2$ in terms of the finite sum of the linearly-independent functions $\Theta(z + \frac{1}{4}\delta + \frac{1}{2}\mu|\frac{1}{4}B)$. Let us try to apply the formula (42) to the multidimensional sG equation:

$$\sum_{k=1}^{N} \partial_{x_k}^2 \Psi = \sin\Psi. \quad (48)$$

It is convenient to write the sG equation as above and eventually introduce the time t by the substitution $x_n = it$. By $\mathbf{d}$ ($\mathbf{d}\in D^g$) we denote the g-dimensional vector

$$\mathbf{d} = (1,1,\ldots,1). \quad (49)$$

Koziel and Kotlarov[2] (see also Refs. 1 and 7) found that the g-periodic solution of a $(1+1)$-sG equation has the form

$$\Psi = 2i \ln[\Theta(z + \tfrac{1}{2}\mathbf{d}|B)\Theta^{-1}(z|B)] + C, \quad (50)$$

where $z = \alpha x + \beta t + \gamma$ is a g-dimensional vector; α, β, γ and C are constants.

We shall try to determine these constants, particularly α and β, in relation to B, and also to generalize our results in order to describe more than only the $(1+1)$-dimensional case, e.g., the $(N+1)$- or better $[(N-1)+1]$-dimensional case.

We assume the solution Ψ is in the form

$$\Psi = 2i \ln[\Theta(z + \tfrac{1}{2}\mathbf{d}|B)\Theta^{-1}(z|B)] + (1 \mp 1)\pi/2, \quad (51)$$

where $z = (z_1, z_2, \ldots, z_g)$ and

$$z_p = \sum_{j=1}^{N} a_{pj} x_j + z_{p0}, \quad p = 1,2,\ldots,g. \quad (52)$$

Here a_{pj} and z_{p0} are constants.

The set a_{pj} forms the matrix $g \times N$ (in general rectangular!) We denote

$$A_{pq} = \sum_{j=1}^{N} a_{pj} a_{qj}. \quad (53)$$

Since

$$\partial_{x_k}^2 = \sum_{p,q=1}^{g} (\partial_{x_k} z_p \partial_{x_k} z_q)\partial_{z_p}\partial_{z_q}, \quad (54)$$

the sG equation (48) becomes

$$\sum_{p,q=1}^{g} A_{pq}\partial_{z_p}\partial_{z_q}\Psi = \sin\Psi, \quad (55)$$

and by (51) and (42) takes the form of the functional equation

$$\sum_{\epsilon\in D^g} \left[\sum_{p,q=1}^{g} A_{pq}\Omega_{\epsilon}^{pq} \pm \tfrac{1}{4}\delta_{\epsilon,\mathbf{d}}\right] F(z;\epsilon|B) = 0, \quad (56)$$

where

$$F(z;\epsilon|B) = \Theta^2(z + \tfrac{1}{2}\mathbf{d} + \tfrac{1}{2}\epsilon|B)\Theta^2(z|B) - \Theta^2(z + \tfrac{1}{2}\epsilon|B)\Theta^2(z + \tfrac{1}{2}\mathbf{d}|B). \quad (57)$$

$\delta_{\epsilon,\mathbf{d}}$ is the Kronecker symbol (1 only if $\epsilon = \mathbf{d}$, otherwise 0). A_{pq} and Ω_{ϵ}^{pq} are given by (53) and (43), respectively. Equation (55) will be satisfied, for example, if for each $\epsilon \neq 0$ ($\epsilon \in D^g$)

$$\sum_{p,q=1}^{g} A_{pq}\Omega_{\epsilon}^{pq} \pm \tfrac{1}{4}\delta_{\epsilon,\mathbf{d}} = 0. \quad (58)$$

This system of $2^g - 1$ simultaneous equations determines the set of $g(g+1)/2$ unknown "scalar produces" A^{pq} (since $A_{pq} = A_{qp}$ and $\Omega_{\epsilon}^{pq} = \Omega_{\epsilon}^{qp}$). The equation for $\epsilon = 0$ in (58) drops out, since $F(z;0|B) = 0$.

For $g = 0$ or 1 the system (58) always has a solution for each symmetrical matrix B [satisfying (3)] if

$$\det\Omega_{\epsilon}^{pq} \neq 0, \quad (59)$$

since then $2^g - 1 = g(g+1)/2$. In the case of $g = 1$, we obtain the usual pendulum solution. For $g = 2$ we have the two-periodic solution, which is more general than the commonly known solutions expressed by 4 arctan $[f(u)g(v)] + C$. This is due to the fact that f and g are the elliptic functions, so that the above solution can always be transformed to a form involving one-dimensional ϑ-functions, whereas the two-dimensional Θ-f allows the representation in terms of one-dimensional ϑ-functions only in exceptional cases.[9] An equivalence holds if $B_{11} = B_{22}$.[13]

In the case $g > 2$, the system of equations (58) is overdetermined. This indicates the existence of additional conditions on elements of matrix B, and/or constraints imposed on A_{pq}.

We intend to discuss these problems in a future paper.

If the functions $F(z;\epsilon|B)$ (indexed by ϵ) are linearly independent, then condition (58) is sufficient for (51) to be the solution of the sG equation (48). Otherwise, it is necessary to find an appropriate set of linearly-independent functions and to express $F(z;\epsilon|B)$ in terms of these functions.

In Appendix B we show that for fixed B, the set of functions

$$\{\Theta(z + \tfrac{1}{2}\mu + \tfrac{1}{4}\nu|B)\}_{\mu,\nu}, \quad (60)$$

doubly indexed by the pair $\mu, \nu \in D^g$, form the set of linearly-independent functions.

Writing (57) in a more suitable form,

$$F(z;\epsilon|B) = \Theta^2(z + \tfrac{1}{2}(\mathbf{d}-\epsilon)|B)\Theta^2(z|B) - \Theta^2(z + \tfrac{1}{2}\mathbf{d} + \tfrac{1}{2}(\mathbf{d}-\epsilon)|B)\Theta^2(z + \tfrac{1}{2}\mathbf{d}|B), \quad (61)$$

one can now apply the formula (46). Observe that the second term on the right-hand side of (61) is simply the first one, with the shifted argument $\mathbf{z}\to\mathbf{z}+\frac{1}{2}\mathbf{d}$. Thus, after some rearrangement (61) becomes

$$F(\mathbf{z};\boldsymbol{\epsilon}|B)=2^{-4g}\sum_{\boldsymbol{\delta},\boldsymbol{\nu},\boldsymbol{\mu},\boldsymbol{\eta}\in D^g}\Theta(\tfrac{1}{4}(\boldsymbol{\epsilon}'\pm\boldsymbol{\delta})+\tfrac{1}{2}\boldsymbol{\eta}|\tfrac{1}{4}B)\Theta(\tfrac{1}{4}(\boldsymbol{\epsilon}'\pm\boldsymbol{\delta})+\tfrac{1}{2}(\boldsymbol{\nu}+\boldsymbol{\mu}+\boldsymbol{\eta})|\tfrac{1}{4}B)$$
$$\times[\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}\boldsymbol{\mu}|\tfrac{1}{4}B)-\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}(\mathbf{d}-\boldsymbol{\mu})|\tfrac{1}{4}B)]\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}+\tfrac{1}{4}\boldsymbol{\delta}|\tfrac{1}{4}B), \tag{62}$$

or, if we apply (29),

$$F(\mathbf{z};\boldsymbol{\epsilon}|B)=2^{-3g}\sum_{\boldsymbol{\delta},\boldsymbol{\nu},\boldsymbol{\mu}\in D^g}\Theta(\tfrac{1}{2}(\boldsymbol{\epsilon}'+\boldsymbol{\delta}+\boldsymbol{\nu}+\boldsymbol{\mu})|\tfrac{1}{2}B)\Theta(\tfrac{1}{2}(\boldsymbol{\nu}+\boldsymbol{\mu})|\tfrac{1}{2}B)$$
$$\times[\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}\boldsymbol{\mu}|\tfrac{1}{4}B)-\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}(\mathbf{d}-\boldsymbol{\mu})|\tfrac{1}{4}B)]\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}+\tfrac{1}{4}\boldsymbol{\delta}|\tfrac{1}{4}B), \tag{63}$$

where $\boldsymbol{\epsilon}'=\mathbf{d}-\boldsymbol{\epsilon}$. Since the $\Theta(\mathbf{z}+\frac{1}{2}\boldsymbol{\nu}+\frac{1}{4}\boldsymbol{\delta}|\frac{1}{4}B)$ are now linearly independent, the functional equation (56) takes the form of the algebraic system of 2^{2g} simultaneous nonhomogeneous equations (with respect to A_{pq})

$$\sum_{\substack{\boldsymbol{\epsilon}'\in D^g\\ \boldsymbol{\epsilon}\neq\mathbf{d}}}\Big[\sum_{p,q=1}^{g}A_{pq}\Omega^{pq}_{\mathbf{d}-\boldsymbol{\epsilon}'}\pm\tfrac{1}{4}\delta_{\boldsymbol{\epsilon}',0}\Big]$$
$$\times\sum_{\boldsymbol{\mu}\in D^g}\Theta(\tfrac{1}{2}(\boldsymbol{\epsilon}'+\boldsymbol{\delta}+\boldsymbol{\mu}+\boldsymbol{\nu})|\tfrac{1}{2}B)\Theta(\tfrac{1}{2}(\boldsymbol{\mu}+\boldsymbol{\nu})|\tfrac{1}{2}B)$$
$$\times[\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}\boldsymbol{\mu}|\tfrac{1}{4}B)-\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}(\mathbf{d}+\boldsymbol{\mu})|\tfrac{1}{4}B)]$$
$$=0, \tag{64}$$

for each $\boldsymbol{\delta},\boldsymbol{\nu}\in D^g$.

Since the last sum in (64) can be written also as

$$\sum_{\boldsymbol{\mu}\in D^g}\Theta(\tfrac{1}{2}(\boldsymbol{\epsilon}'+\boldsymbol{\mu})|\tfrac{1}{2}B)\Theta(\tfrac{1}{2}(\boldsymbol{\mu}+\boldsymbol{\delta})|\tfrac{1}{2}B)$$
$$\times[\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}(\boldsymbol{\mu}+\boldsymbol{\nu})|\tfrac{1}{4}B)-\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}(\mathbf{d}+\boldsymbol{\nu}+\boldsymbol{\mu})|\tfrac{1}{4}B)], \tag{65}$$

performing the summation over $\boldsymbol{\epsilon}'$, we finally obtain

$$\sum_{\boldsymbol{\mu}\in D^g}\Big[\sum_{p,q=1}^{g}A_{pq}\Xi^{pq}_{\boldsymbol{\mu}}\pm\tfrac{1}{4}\Theta(\tfrac{1}{2}\boldsymbol{\mu}|\tfrac{1}{2}B)\Big]\Theta(\tfrac{1}{2}(\boldsymbol{\delta}+\boldsymbol{\mu})|\tfrac{1}{2}B)$$
$$\times[\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}(\boldsymbol{\mu}+\boldsymbol{\nu})|\tfrac{1}{4}B)-\Theta(\tfrac{1}{4}\boldsymbol{\delta}+\tfrac{1}{2}(\mathbf{d}+\boldsymbol{\nu}+\boldsymbol{\mu})|\tfrac{1}{4}B)]$$
$$=0, \tag{66}$$

for each pair pair $\boldsymbol{\delta},\boldsymbol{\nu}\in D^g$. We have denoted here

$$\Xi^{pq}_{\boldsymbol{\mu}}=\sum_{\substack{\boldsymbol{\epsilon}'\in D^g\\ \boldsymbol{\epsilon}'\neq\mathbf{d}}}\Omega^{p,q}_{\mathbf{d}-\boldsymbol{\epsilon}'}\Theta(\tfrac{1}{2}(\boldsymbol{\epsilon}'+\boldsymbol{\mu})|\tfrac{1}{2}B). \tag{67}$$

[In principle, the summation is over all $\boldsymbol{\epsilon}'$, but the element for $\boldsymbol{\epsilon}'=\mathbf{d}$ vanishes, since $F(\mathbf{z};0|B)=0$.]

We have thus far proved that the multidimensional sG equation in the form (48) has the solution (51) if and only if the system of simultaneous algebraic equations (66) is satisfied (for each $\boldsymbol{\delta},\boldsymbol{\nu}\in D^g$). Relation (58) is the particular case of (66) if the appropriate determinants in this equation do not vanish. Similarly, one can analyze the other particular cases.

In general, the system of equations (66) is overdetermined; it contains 2^{2g} equations for each pair $\boldsymbol{\nu}$, $\boldsymbol{\delta}$ and there are only $g(g+1)/2$ variables A_{pq} to be determined. We should, however, bear in mind that $g(g+1)/2$ complex elements of the symmetric matrix B must also be determined.

We have discussed the solution in the form (51), i.e., choosing $C=0$ or π in the relation (50). These are the unique values of C. The proof is given in Appendix A.

If there exists a solution of the system (66), we have a g-periodical solution of the sG equation. The dimension of the physical space, e.g., $1+1$, $2+1$, etc., is hidden under the scalar products A_{pq}. Having A_{pq} as the solution of (66), it is necessary to find a_{pj} or, better yet, the class a_{pj} that fulfills (53). This is an elementary algebraic problem, but the result certainly depends on the assumed dimension of the physical space: $N+1$.

In consequence, we have obtained a g-periodic solution in $(N+1)$-dimensional space, which is then reminiscent of the multidimensional g-soliton solution.

4. SUMMARY AND CONCLUSIONS

By deriving a variety of algebraic identities for the multidimensional theta function and for its derivatives, we arrive at the conclusion that the sG equation can be reduced to a functional equation and even to a system of simultaneous algebraic equations.

Thus, the existence of a multidimensional solution of the sG equation in form (51) involving theta functions depends on the solvability of the system of algebraic equations.

The discussed solution represents a natural generalization of the $(1+1)$-dimensional case. Moreover, the derived relations give useful formulas for the determination of the constants appearing also in the simplest $(1+1)$-dimensional case.

A rather fascinating and deep resemblance occurs between the multisoliton and multiperiodic solutions of the sG equation. This resemblance is of practical importance: the methods of soliton theory are better developed whereas the multiperiodic solutions concern the more typical physical situations that arise in a bounded region.

Yet one also senses the essential differences, particularly in the more-than-$(1+1)$-dimensional world.

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. J. Konopka and Dr. S. Lewandowski for critical reading of the manuscript and Mrs. J. Pelka for her help during the preparation of this paper.

APPENDIX A

Below we prove that the constant C appearing in (50) takes only two different values: 0 or π. Substituting (50) into

(55) we find

$$\sum_{\epsilon\in D^{\kappa}}\left\{\left[\sum_{p,q=1}^{g}A_{pq}\Omega_{\epsilon}^{pq}+\tfrac{1}{4}\delta_{\epsilon,\mathbf{d}}\cos C\right]F_{-}(\mathbf{z};\boldsymbol{\epsilon}|B)+\frac{i}{4}\delta_{\epsilon,\mathbf{d}}\sin CF_{+}(\mathbf{z};\boldsymbol{\epsilon}|B)\right\}=0, \tag{A1}$$

where

$$F_{\pm}(\mathbf{z};\boldsymbol{\epsilon}|B)=\Theta^{2}(\mathbf{z}+\tfrac{1}{2}\mathbf{d}+\tfrac{1}{2}\boldsymbol{\epsilon}|B)\Theta^{2}(\mathbf{z}|B)\pm\Theta^{2}(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\epsilon}|B)\Theta^{2}(\mathbf{z}+\tfrac{1}{2}\mathbf{d}|B), \tag{A2}$$

and the remaining symbols were defined previously. [A_{pq} is given by (53); Ω_{ϵ}^{pq} by (43); $\mathbf{d}$ by (49)], F_{-} is equal, of course, to the F in relation (57). If (A1) holds for any $\mathbf{z}$, it holds also for $\mathbf{z}+\frac{1}{2}\mathbf{d}$. But

$$F_{\pm}(\mathbf{z}+\tfrac{1}{2}\mathbf{d};\boldsymbol{\epsilon}|B)=\pm F_{\pm}(\mathbf{z};\boldsymbol{\epsilon}|B). \tag{A3}$$

Thus we have

$$\sum_{\epsilon\in D^{\kappa}}\left\{\left[\sum_{p,q=1}^{g}A_{pq}\Omega_{\epsilon}^{pq}+\tfrac{1}{4}\delta_{\epsilon,\mathbf{d}}\cos C\right]F_{-}(\mathbf{z};\boldsymbol{\epsilon}|B)-\frac{i}{4}\delta_{\epsilon,\mathbf{d}}F_{+}(\mathbf{z};\boldsymbol{\epsilon}|B)\right\}=0. \tag{A4}$$

Adding (A1) and (A4) one obtains

$$(\sin C)F_{+}(\mathbf{z};\mathbf{d}|B)=0. \tag{A5}$$

If $\sin C\neq 0$, (A5) by (A2) yields

$$\Theta^{2}(\mathbf{z}|B)=\pm i\Theta^{2}(\mathbf{z}+\tfrac{1}{2}\mathbf{d}|B). \tag{A6}$$

(A6) substituted into (50) reduces the whole solution to constant. Therefore $\sin C=0$, and of course our choice of $C=(1\pm 1)\pi/2$ is justified.

APPENDIX B

According to A. I. Markuschevich,[4] a Θ-f of range n is defined as follows:

$$\Theta_{n}[\boldsymbol{\chi}](z;A)=\sum_{\mathbf{m}\in Z^{\kappa}}\exp\left\{i\pi(n\mathbf{m}+\boldsymbol{\chi})\cdot\left[2\mathbf{z}+\frac{A}{n}(n\mathbf{m}+\boldsymbol{\chi})\right]\right\}, \tag{B1}$$

where A is symmetrical matrix satisfying the condition (3), n is a positive integer and $\boldsymbol{\chi}$ is a vector with integer components χ_i and

$$0\leqslant\chi_i<n,\quad 1\leqslant i\leqslant g. \tag{B2}$$

By analogy to the previously introduced nomenclature, one can write

$$\boldsymbol{\chi}\in D_{n}^{g} \tag{B3}$$

which means $\boldsymbol{\chi}$ belongs to g-dimensional die of magnitude n. (In this sense the die D^{g} used previously would be D_{2}^{g}.)

Markushevich states that for fixed g,A, and n, the set

$$\{\Theta_{n}[\boldsymbol{\chi}](\mathbf{z};A)\}_{\boldsymbol{\chi}},\quad \boldsymbol{\chi}\in D_{n}^{g}, \tag{B4}$$

in $\boldsymbol{\chi}$ forms the set of linearly-independent functions and the total number of these functions is n^{g}.

Let us consider $\Theta(\mathbf{z}+(1/n)\boldsymbol{\epsilon}_{n}|B)$, where $\boldsymbol{\epsilon}_{n}\in D_{n}^{g}$. We get

$$\Theta(\mathbf{z}+\frac{1}{n}\boldsymbol{\epsilon}_{n}|B)=\sum_{s\in Z^{\kappa}}\exp\left\{i\pi\left[2\mathbf{s}\cdot(\mathbf{z}+\frac{1}{n}\boldsymbol{\epsilon}_{n})+\mathbf{s}\cdot B\mathbf{s}\right]\right\}$$

$$=\sum_{\boldsymbol{\chi}_{n}\in D_{n}^{g}}\sum_{\mathbf{m}\in Z^{g}}\exp\left\{\left[2\left(\mathbf{z}+\frac{1}{n}\boldsymbol{\epsilon}_{n}\right)(\mathbf{m}n+\boldsymbol{\chi}_{n})+(\mathbf{m}n+\boldsymbol{\chi}_{n})B(\mathbf{m}n+\boldsymbol{\chi}_{n})\right]\right\}, \tag{B5}$$

where instead of the sum over $s\in Z^{g}$ we put $\mathbf{s}=n\mathbf{m}+\boldsymbol{\chi}_{n}$, $\mathbf{m}\in Z^{g}$, $\boldsymbol{\chi}_{n}\in D_{n}^{g}$. After simple rearrangements, (B5) becomes

$$\sum_{\boldsymbol{\chi}_{n}\in D_{n}^{\kappa}}\exp(i2\pi\boldsymbol{\epsilon}_{n}\cdot\boldsymbol{\chi}_{n}/n)\times\sum_{\mathbf{m}\in Z^{\kappa}}\exp\{i\pi[2(\mathbf{m}n+\boldsymbol{\chi}_{n})\cdot\mathbf{z}+(\mathbf{m}n+\boldsymbol{\chi}_{n})\cdot B(\mathbf{m}n+\boldsymbol{\chi}_{n})]\}$$

$$=\sum_{\boldsymbol{\chi}_{n}\in D_{n}^{g}}\exp(i2\pi\boldsymbol{\epsilon}_{n}\cdot\boldsymbol{\chi}_{n}/n)\Theta_{n}[\boldsymbol{\chi}_{n}](\mathbf{z};nB). \tag{B6}$$

Thus $\Theta(\mathbf{z}+(1/n)\boldsymbol{\epsilon}_{n}|B)$ is expressed by a Θ-f of range n. Since

$$n^{-g}\sum_{\boldsymbol{\chi}_{n}\in D_{n}^{g}}\exp[i2\pi\boldsymbol{\epsilon}\cdot(\boldsymbol{\chi}-\boldsymbol{\chi}')/n]=n^{-g}\prod_{i=1}^{g}\sum_{\epsilon_{i}=0}^{n}\exp[i2\pi\epsilon_{i}(\chi_{i}-\chi_{i}')/n]=\prod_{i=1}^{g}\delta_{\chi_{i}\chi_{i}'}=\delta_{\boldsymbol{\chi}\boldsymbol{\chi}'}, \tag{B7}$$

the inverse matrix exists and hence the determinant of the matrix $\exp(i2\pi\boldsymbol{\epsilon}\cdot\boldsymbol{\chi}_{n}/n)$ in (B6) does not vanish. Therefore for fixed n and B, the functions $\Theta(\mathbf{z}+(1/n)\boldsymbol{\epsilon}_{n}|B)$ indexed by $\boldsymbol{\epsilon}_{n}\in D_{n}^{g}$ are also linearly independent.

For example, if $n=2$,

$$\Theta(z+\tfrac{1}{2}\boldsymbol{\mu}|B),\quad \boldsymbol{\mu}\in D_{2}^{g};$$

if $n=4$,

$$\Theta(z+\tfrac{1}{4}\boldsymbol{\mu}|B),\quad \boldsymbol{\mu}\in D_{4}^{g}, \tag{B8}$$

are linearly independent, respectively.

But if $\boldsymbol{\mu}\in D_{4}^{g}$, putting $\boldsymbol{\mu}=2\boldsymbol{\nu}+\boldsymbol{\eta}$ and $\boldsymbol{\nu},\boldsymbol{\eta}\in D_{2}^{g}$, we obtain

$$\Theta(\mathbf{z}+\tfrac{1}{4}\boldsymbol{\mu}|B)=\Theta(\mathbf{z}+\tfrac{1}{2}\boldsymbol{\nu}+\tfrac{1}{4}\boldsymbol{\eta}|B), \tag{B9}$$

and in conclusion the set $\Theta(\mathbf{z}+\frac{1}{2}\boldsymbol{\nu}+\frac{1}{4}\boldsymbol{\eta}|B)$, $\boldsymbol{\nu},\boldsymbol{\eta}\in D_{2}^{g}$ also forms a set of linearly independent functions.

[1] V. B. Matveev, "Abelian Functions and Solitons," Preprint No. 373 of Institute of Theoretical Physics, University of Wroclaw (1976).
[2] V. O. Kozel and V. P. Kotlyarov, Doklady A. N. Ukr. SSR Ser. A **10**, 878 (1976).
[3] A. Nakamura, J. Phys. Soc. Japan **47**, 1701 (1979); **48**, 1365 (1980).
[4] A. I. Markuschevich, *Introduction to Classical Theory of Abelian Functions* (Nauka, Moscow, 1979), p. 187 (in Russian).
[5] A. P. Its and V. B. Matveev, Teor. Mat. Fiz. **23**, 1, 51 (1975).
[6] B. A. Dubrovin, V. B. Matveev, and C. P. Novikov, Usp. Mat. Nauk **XXXI**, 1, 55 (1976).
[7] V. E. Zacharov, S. V. Manakov, S. P. Novikov, and L. P. Pitajevskij, *Theory of Solitons*, edited by S. P. Novikov (Nauka, Moscow, 1980) (in Russian).
[8] J. Zagrodziński, Lett. Nuovo Cimento **30**, 266 (1981).
[9] A. Krazer, *Lehrbuch der Thetafunktionen* (Teubner, Leipzig 1903).
[10] H. Bateman and A. Erdélyi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1955).
[11] W. Magnus and F. Oberhettinger, *Formeln und Sätze für die Speziellen Funktionen der Mathematischen Physik* (Springer, Berlin, 1948).
[12] R. Bellman, *A Brief Introduction to Theta Function* (Holt, Reinhart and Winston, New York, 1961).
[13] J. Zagrodziński, Workshop on NL. Evol. Eqns, Solitons and Inv. Spectral Methods, Trieste (1981).

Existence and asymptotic behavior of Padé approximants to the Korteweg-de-Vries multisoliton solutions

C. Liverani
Istituto di Fisica, Via Irnerio 46, Bologna, Italy

G. Turchetti
Istituto di Fisica, Via Irnerio 46, Bologna, Italy and Istituto Nazionale di Fisica Nucleare, sezione di Bologna, Bologna, Italy

(Received 9 November 1981; accepted for publication 11 June 1982)

The summation procedure of the Padé type is applied to the perturbation expansion of the solution of the potential Korteweg-de-Vries equation (K.d.V.), introduced by Rosales. For the N-soliton solution without background the $[(n-1)/n]$ Padé approximants are shown to exist for $n \leqslant N$. Their asymptotic behavior is investigated and it is found that it corresponds to a system of n solitons with the leading velocity parameters. The analogous results for the K.d.V. then follow in agreement with some previous numerical observations.

PACS numbers: 03.40.Kf, 02.30.Mv, 02.30.Lt

1. INTRODUCTION

Explicit multisoliton solutions are known for several nonlinear partial differential equations. For the Korteweg–de Vries (K.d.V.) and nonlinear cubic Schrödinger equation the classical procedure for obtaining them is based on the associated linear eigenvalue problems according to the pioneering works by Gardner, Green, Kruskal, and Miura[1] and Zakharov and Shabat.[2] More generally, any initial value problem is reduced to a linear problem.

A direct method for finding multisoliton solutions was used by Hirota[3] and a systematic approach based on perturbation expansions was proposed by Rosales.[4] For several classical nonlinear equations explicit sums of the perturbation series were obtained in the multisoliton case, while more generally, the formal sums were shown to satisfy linear integral equations (such as the Marchenko equation for K.d.V.). This method, in spite of some algebraic labor required by the computation of the perturbation series, seems to be quite general and suggests the investigation of systematic summation procedures.

The rational approximations of Padé type (P.A.) proved to exhibit some interesting features for the potential K.d.V. equation. In fact the $[(n-1)/n]$ reproduce exactly the N-soliton solution for $n = N$,[5] and bound it from below for $n < N$.[6] Moreover, numerical evidence was found that for $n < N$ the $[(n-1)/n]$ P.A. behaves asymptotically for $t \to +\infty$ as a system of n solitons having the same parameters as the n leading solitons of the exact solution.[7,8] Even though the numerical examples were restricted to very low values of N and n, it was conjectured that the result would hold for any value of N and n and also in presence of a background. Analogous numerical results were found for the modified K.d.V.

In this paper we rigorously prove the existence of $[(n-1)/n]$ P. A. to a $N \geqslant n$ soliton solution and the above conjecture on their asymptotic behavior for the potential K.d.V. equation. The method we use should allow extensions to other nonlinear equations.

The plan of the work is the following. In Sec. 2 we review the Rosales procedure. In Sec. 3 we prove the existence of the $[(n-1)/n]$ P. A. In Sec. 4 we describe the asymptotic behavior of a multisoliton. In Sec. 5 we quote some preliminary results related to the asymptotic behavior of the $[(n-1)/n]$ P.A. which is finally proved in Sec. 6.

2. KORTEWEG–DE VRIES EQUATION AND PADÉ APPROXIMANTS

The standard form of the Korteweg–de Vries equation is given by

$$u_t + u_{xxx} + 6\lambda u u_x = 0\,, \tag{2.1}$$

and letting $u = -U_x$ the potential K.d.V. equation reads

$$U_t + U_{xxx} - 3\lambda U_x^{\,2} = 0\,. \tag{2.2}$$

The perturbation expansion of (2.2) can be written

$$U = \sum_{n=0}^{\infty} (-\lambda)^n U_n\,, \tag{2.3}$$

where U_0 is a solution of the linear homogeneous equation

$$U_{0,t} + U_{0,xxx} = 0 \tag{2.4}$$

and U_n for $n \geqslant 1$ satisfy the linear inhomogeneous equations

$$U_{n,t} + U_{n,xxx} = 3 \sum_{k=0}^{n-1} U_{k,x}\, U_{n-1-k,x}\,. \tag{2.5}$$

Rosales[4] has shown that choosing the solution of (2.4) as

$$U_0 = \int_c e^{i(kx + k^3 t)} d\mu(k)\,, \tag{2.6}$$

where $d\mu(k)$ is an appropriate measure on the complex plane, then U_n can be written

$$U_n = i^n \int_{c^{n+1}} \frac{\exp\left[i \sum_{j=1}^{n+1} (k_j x + k_j^3 t)\right]}{\prod_{j=1}^{n} (k_j + k_{j+1})} \times d\mu(k_1) \cdots d\mu(k_{n+1}) \quad , \quad n \geqslant 1. \tag{2.7}$$

The usual multisoliton solution corresponds to a discrete positive measure with support on the positive imaginary axis, namely

 0022-2488/83/010053-12$02.50

$$\int_c f(k)d\mu(k) = \sum_{l=1}^{N} a_l^2 f(ik_l). \tag{2.8}$$

The k_j must all be distinct and we order them in a decreasing sequence

$$k_1 > k_2 > \cdots > k_N . \tag{2.9}$$

The coefficients of the perturbation series are then given by

$$U_n = \langle \phi, A^n \phi \rangle, \tag{2.10}$$

where $\langle \cdot,\cdot \rangle$ denotes the scalar product in $\mathbb{R}^N$, ϕ is a vector of $\mathbb{R}^N$ defined by

$$\phi_j = a_j \exp\left[-\tfrac{1}{2} k_j (x - v_j t)\right], \quad v_j = k_j^2, \ j = 1,...,N \tag{2.11}$$

and A is a $N \times N$ matrix defined by

$$A_{ij} = \frac{\phi_i \phi_j}{k_i + k_j}, \quad i,j = 1,...,N. \tag{2.12}$$

The perturbation series can be summed for $|\lambda|$ small enough and analytically continued to give

$$U = \langle \phi, [1 + \lambda A]^{-1} \phi \rangle . \tag{2.13}$$

Indeed, since the k_j are all distinct A is a positive matrix. If $\lambda > 0$, as we shall always assume, then U is bounded and corresponds to a multisoliton solution.

Such a solution is a Stieltjes function of λ for any value of x and t; its poles have positive residues and lie on the real negative axis of the λ plane.[5]

As a consequence, by truncating the associated continued fraction one obtains two sequences of approximations monotonically converging to the exact solution from below and from above.

Such sequences are identical to the $[(n-1)/n]$ and $[n/n]$ Padé approximants (P.A.) and we write

$$[(n-1)/n]_{U(\lambda)} \leqslant U(\lambda) \leqslant [n/n]_{U(\lambda)} , \tag{2.14}$$

where the equal sign in (2.14) holds for $n \geqslant N$. The behavior of the poles of both the exact and the approximate solution was investigated in Ref. 8; however, a rigorous proof of the existence of $[(n-1)/n]$ P.A. and of their asymptonic behavior was missing. The $[(n-1)/n]$ P.A. is defined as the irreducible ratio, if it exists, of two polynomials $P_{n-1}(\lambda)$ and $Q_n(\lambda)$ of degrees $n-1$ and n, respectively,

$$[(n-1)/n]_{U(\lambda)} = \frac{P_{n-1}(\lambda)}{Q_n(\lambda)} \tag{2.15}$$

such that

$$Q_n(\lambda)U(\lambda) - P_{n-1}(\lambda) = O(\lambda^{2n}). \tag{2.16}$$

The $[(n-1)/n]$ P. A. exist according to the above definition if the linear system associated to the normalized coefficients of $Q_n(\lambda)$ [namely $Q_n(0) = 1$] has a unique solution, that is

$$\begin{vmatrix} U_0 & U_1 & \cdots & U_{n-1} \\ U_1 & U_2 & \cdots & U_n \\ \vdots & \vdots & & \vdots \\ U_{n-1} & U_n & \cdots & U_{2n-2} \end{vmatrix} \neq 0. \tag{2.17}$$

An equivalent condition is the linear independence of the vectors

$$\phi, A\phi, ..., A^{n-1}\phi , \tag{2.18}$$

whose Gram determinant is given precisely by the l.h.s. of (2.17). It is also useful to notice that the exact solution can be written as

$$U = \langle \phi, \psi \rangle , \tag{2.19}$$

where

$$\psi = \phi - \lambda A \psi \tag{2.20}$$

and that the $[(n-1/n]$ P.A. are obtained by replacing the exact solution of the linear equation (2.20) with the approximate solution $\hat{\psi}$

$$[(n-1)/n]_U = \langle \phi, \hat{\psi} \rangle , \tag{2.21}$$

where

$$\hat{\psi} = \phi - \lambda PAP\hat{\psi} \tag{2.22}$$

and P is a projector into the n-dimensional subspace $\mathscr{E}_n$ spanned by the vectors (2.18). The proof is immediate if one observes that $\langle \phi, (1 + \lambda PAP)^{-1}\phi \rangle$ is a rational function in λ of the right order and that $\langle \phi, (PAP)^k \phi \rangle = \langle \phi, A^k \phi \rangle$ for $k = 0,1,..., 2n-1$ imply the agreement of the Taylor series of $\langle \phi, \psi \rangle$ and $\langle \phi, \hat{\psi} \rangle$ up to order $2n-1$ in λ. Detailed proofs of the above properties and further information about P. A. can be found in Ref. 9.

3. EXISTENCE OF THE $[(n-1)/n]$ P. A.

Following the definitions of the previous sections we can state the basic existence result.

Theorem 1: The $[(n-1)/n]$ P. A. to the N-soliton solution exists and is unique for $n \leqslant N$ provided that the $\{k_1,...,k_N\}$ are all distinct.

Proof: The result follows if we can prove that the vectors

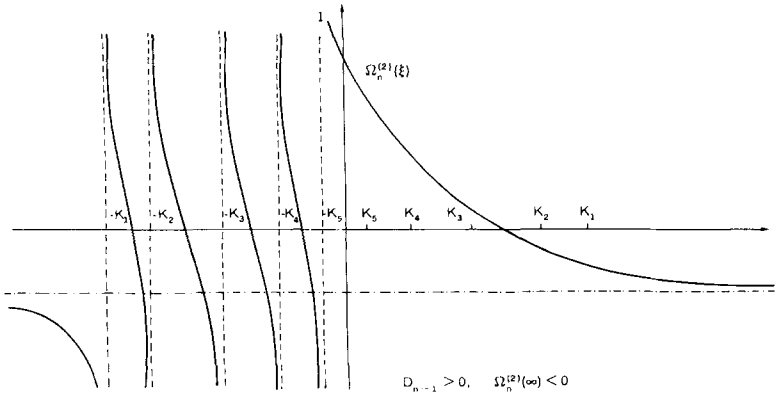


FIG. 1.

(2.19) are linearly independent for $n \leqslant N$. Letting

$$\phi^{(k)} = A^{k-1}\phi, \quad k = 1,...,n, \tag{3.1}$$

a sufficient condition for these vectors to be linearly independent is that the determinant

$$D_n = \begin{vmatrix} \phi_1^{(1)} & \cdots & \phi_n^{(1)} \\ \vdots & & \vdots \\ \phi_1^{(n)} & \cdots & \phi_n^{(n)} \end{vmatrix} \tag{3.2}$$

does not vanish.

We use an inductive procedure. It is evident that $D_1 = \phi_1^{(1)} = \phi_1$ never vanishes (for any finite values of x and t). Assuming that $D_{n-1} \neq 0$ for $1 \leqslant n \leqslant N$ we shall prove that $D_n \neq 0$.

Let us consider the functions $\Omega_n^{(m)}(y)$ defined by

$$\Omega_n^{(m)}(y) = \begin{vmatrix} \phi_1^{(1)} & \cdots & \phi_{n-1}^{(1)} & 0 \\ \vdots & & \vdots & \vdots \\ \phi_1^{(n-m)} & \cdots & \phi_{n-1}^{(n-m)} & 0 \\ \phi_1^{(n-m+1)} & \cdots & \phi_{n-1}^{(n-m+1)} & 1 \\ \phi_1^{(n-m+2)} & \cdots & \phi_{n-1}^{(n-m+2)} & \sum_{j=1}^{N} \dfrac{\phi_j \phi_j^{(1)}}{y + k_j} \\ \vdots & & \vdots & \vdots \\ \phi_1^{(n)} & \cdots & \phi_{n-1}^{(n)} & \sum_{j=1}^{N} \dfrac{\phi_j \phi_j^{(m-1)}}{y + k_j} \end{vmatrix}. \tag{3.3}$$

From (3.1) we have

$$\phi^{(k)} = A\phi^{(k-1)} \tag{3.4}$$

and taking (2.12) into account we can write

$$\phi_l^{(k)}/\phi_l = \sum_{j=1}^{N} \frac{\phi_j \phi_j^{(k-1)}}{k_l + k_j}. \tag{3.5}$$

It is immediate to check that

$$\Omega_n^{(1)}(y) = D_{n-1} \tag{3.6}$$

and

$$\Omega_n^{(n)}(k_n) = D_n/\phi_n. \tag{3.7}$$

In fact, accounting for (3.5) the last column of $\Omega_n^{(n)}(k_n)$ is given by $1, \phi_n^{(2)}/\phi_n, ..., \phi_n^{(n)}/\phi_n$. Since we have assumed $D_{n-1} \neq 0$ and since $\phi_n \neq 0$ the task of proving that $D_n \neq 0$ amounts to proving that $\Omega_n^{(n)}(k_n) \neq 0$ knowing that $\Omega_n^{(1)} \neq 0$.

The functions $\Omega_n^{(m)}(y)$ are for $m \geqslant 2$ rational functions of y; the degrees of the numerator and denominator polynomials are not greater than N. The residues at the poles at $-k_1, ..., -k_N$ are given by

$$\lim_{y \to -k_l} \Omega_n^{(m)}(y)(y + k_l) = \Omega_n^{(m-1)}(k_l)\phi_l^2, \quad l = 1,...,N. \tag{3.8}$$

Indeed, after taking the limit in the l.h.s. of (3.8) we have a determinant whose last column is precisely $0,...,0, \phi_l^2, \phi_l\phi_l^{(2)},...,\phi_l\phi_l^{(m-1)}$, where the zero appears $n - m + 1$ times. After factoring ϕ_l^2 we obtain a determinant, whose last column is $0,...,0,1,\phi_l^{(2)}/\phi_l,...,\phi_l^{(m-1)}/\phi_l$, which is identical to $\Omega_n^{(m-1)}(k_l)$, as one can see from (3.3) and (3.5). It is obvious that if $\Omega_n^{(m-1)}(k_l) = 0$ then the residue of the pole at $y = -k_l$ vanishes, so that the degrees of the numerator and denominator polynomials of $\Omega_n^{(m)}(y)$ are at most $N - 1$.

We claim that $\Omega_n^{(m)}(y)$ has at most $m - 1$ zeros on $\mathbb{R}_+$.

This result also can be proved by induction. In fact for $m = 2$ all the residues are of equal sign since from (3.8) and (3.6) the residue at $y = -k_l$ for $l = 1,...,N$ is given by $\phi_l^2 D_{n-1} \neq 0$. As a consequence, on $\mathbb{R}_+$ $\Omega_n^{(2)}(y)$ has at most one zero [none if $\Omega^{(2)}(\infty)D_{n-1} > 0$] (see Fig. 1). Assuming that $\Omega_n^{(m-1)}(y)$ has at most $m - 2$ zeros on $\mathbb{R}_+$ we can then prove that $\Omega_n^{(m)}(y)$ has at most $m - 1$ zeros on $\mathbb{R}_+$. A preliminary remark is that $\Omega_n^{(m)}(y)$ cannot vanish identically since all of its N residues cannot be zero. The residues for (3.8) are proportional to $\Omega_n^{(m-1)}(k_l)$ and at least $N - (m - 2) \geqslant 2$ of them are different from zero since $m \leqslant n \leqslant N$. Let us consider the case in which no one of the $\{k\}$ is a zero of $\Omega_n^{(m-1)}(y)$ and let j_1 of them $k_N, ..., k_{N-j_1+1}$ fall between 0 and the first zero of $\Omega_n^{(m-1)}(y)$, j_2 of them fall between the first two zeros of $\Omega_n^{(m-1)}(y)$, and finally j_{m-1} of them fall after the last zero of $\Omega_n^{(m-1)}(y)$. The poles corresponding to the same group of $\{k\}$ (for example $-k_N, -k_{N-1}, ..., -k_{N-j_1+1}$) have, due to (3.8), residues of equal sign and at least one zero has to lie between each pair. Therefore to j_l poles there correspond at least $j_l - 1$ zeros. Two contiguous poles not belonging to the same group have opposite residues and an even number of zeros or none can fall between them. As a consequence a lower bound to the number N_- of zeros in $\mathbb{R}_-$ of $\Omega_n^{(m)}(y)$ is given by

$$N_- \geqslant \sum_{l=1}^{m-1} (j_l - 1) = N - m + 1. \tag{3.9}$$

We conclude that the number N_+ of zeros in $\mathbb{R}_+$ is given by

$$N_+ = N - N_- \leqslant m - 1. \tag{3.10}$$

If p of the $\{k\}$ are zeros of $\Omega_n^{(m-1)}(y)$ the estimate (3.10) still holds since $\Omega_n^{(m)}(y)$ is then a rational fraction of order $N - p$ and one has simply to replace N by $N - p$ in (3.9) and (3.10). In Fig. 2 we illustrate the behavior of $\Omega_n^{(3)}(y)$ when $p = 0$.

We can now end our proof since we know that $\Omega_n^{(n)}(y)$ will have at most $n-1$ zeros on $\mathbb{R}_+$. Moreover, $\Omega_n^{(n)}(k_l)=0$ for $l=1,\ldots,n-1$ since the elements of the last column are 1, $\phi_l^{(2)}/\phi_l,\ldots,\phi_l^{(n-1)}/\phi_l$, identical to the l th column up to the factor $1/\phi_l$. We can argue that $\Omega_n^{(n)}(k_n)\neq 0$. However due to (3.7) this implies that $D_n\neq 0$ and our induction is complete.

4. ASYMPTOTICS OF THE EXACT SOLUTION

In order to investigate the asymptotic behavior of the multisoliton solution one cannot simply take the limit of $U(x,t)$ for $t\to\infty$, but must rather follow the signal by moving with a given speed v and look at this picture for very large times. The collection of these pictures for different values of v will give a complete description of the multisoliton asymptotic state which can be visualized as a superposition of individual solitons.

Let us first recall that the single soliton solution, given by (2.13) for $N=1$, reads

$$U(x,t)=\frac{\phi_1^2(x,t)}{1+(\lambda/2k_1)\phi_1^2(x,t)}=\frac{2k_1/\lambda}{1+\exp[k_1(x-v_1t)+2\delta]}, \tag{4.1}$$

where δ is given by

$$\delta=\tfrac{1}{2}\ln\left[2k_1/\lambda a_1^2\right]. \tag{4.2}$$

By differentiating one recovers the actual soliton solution of K.d.V. and identifies δ with the phase factor. Indeed one obtains

$$v(x,t)=-\frac{\partial}{\partial x}U(x,t)=\frac{k_1^2}{2\lambda}\cdot\frac{1}{\cosh^2[\tfrac{1}{2}k_1(x-v_1t)+\delta]}. \tag{4.3}$$

Let us move along with velocity v; namely, choose $x=vt+\xi$ and consider the limit for $t\to\infty$

$$U_\infty(\xi)=\lim_{t\to+\infty}U(vt+\xi,t). \tag{4.4}$$

In the case of a single soliton $U_\infty(\xi)$ follows from (4.1) and reads

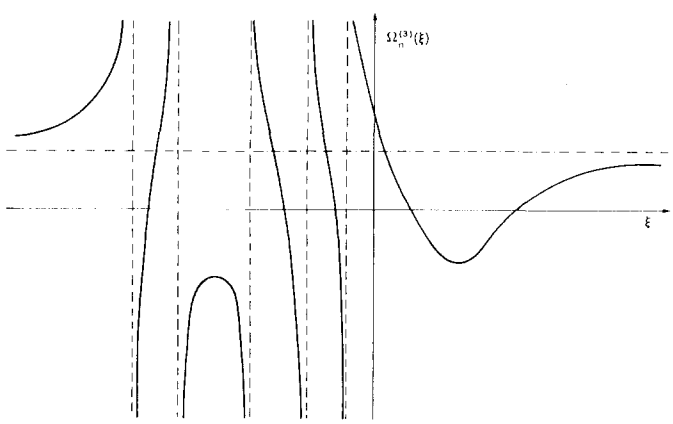


FIG. 2.

$$U_\infty(\xi)=\begin{cases}0 & v>v_1\\ U(\xi,0) & v=v_1.\\ 2k_1/\lambda & v<v_1\end{cases} \tag{4.5}$$

In order to investigate the N-soliton behavior we distinguish four different regions $v>v_1$, $v_{l+1}<v<v_l$, $v=v_l$, and $v<v_N$. In fact, we first observe that

$$\phi_l\equiv\phi_l(vt+\xi,t)=\phi_l(\xi,0)\exp[\tfrac{1}{2}k_l(v_l-v)t] \tag{4.6}$$

and that

$$\lim_{t\to+\infty}\phi_l=\begin{cases}0 & v>v_l\\ \phi_l(\xi,0) & v=v_l.\\ \infty & v<v_l\end{cases} \tag{4.7}$$

The asymptotic behavior of $U(vt+\xi,t)$ for $t\to+\infty$ is given by

$$U_\infty(\xi)=\lim_{t\to+\infty}U(vt+\xi,t)=\begin{cases}0 & v>v_1\\ \frac{2}{\lambda}\sum_{j=1}^{l}k_j & v_{l+1}<v<v_l\\ \frac{2}{\lambda}\sum_{j=1}^{l-1}k_j+\frac{2k_l/\lambda}{1+\exp[k_l\xi+2\delta_l]} & v=v_l\\ \frac{2}{\lambda}\sum_{j=1}^{N}k_j & v<v_N\end{cases}, \tag{4.8}$$

where δ_l is defined by

$$\delta_l=\frac{1}{2}\ln\left[\frac{\Delta_{l-1}}{\lambda a_l^2\Delta_l}\right];\ \Delta_j=\begin{vmatrix}\frac{1}{2k_1} & \cdots & \frac{1}{k_1+k_j}\\ \vdots & & \vdots\\ \frac{1}{k_1+k_j} & \cdots & \frac{1}{2k_j}\end{vmatrix}$$

$$\Delta_0=1. \tag{4.9}$$

From (4.8) we understand that the N-soliton solution is asymptotically a superposition of N free solitons, whose phases, however, are modified, as one can see by comparing (4.9) with (4.2). In view of the study of the asymptotics and in order to set some basic notations, we carry out a proof of (4.8), which was first obtained by Zackarov.[10]

A. $v>v_1$

In this case $\phi_j\to 0$ for $j=1,\ldots,N$ and from (2.12) and (2.13) we see that $U_\infty=0$.

B. $v_{l+1}<v<v_l$

We need to write the equations (2.19), (2.20), equivalent to (2.13), in a different form, since neither ϕ nor A has a limit for $t\to+\infty$. Let us define the $N\times N$ matrices Γ,B and the vectors l, η according to

$$\Gamma_{ij}=\phi_i\,\delta_{ij},\quad B_{ij}=\frac{1}{k_i+k_j},\quad i,j=1,\ldots,N \tag{4.10}$$

and

$$e_i = 1, \quad i = 1,...,N, \ \eta = \Gamma\psi. \tag{4.11}$$

After noticing that $A = \Gamma B \Gamma$ and that $\phi = \Gamma e$ one can replace (2.19) by

$$U = \langle e,\eta \rangle \tag{4.12}$$

and (2.20) by

$$e = (\Gamma^{-2} + \lambda B)\eta \,. \tag{4.13}$$

Since the ϕ_j have different behaviors according to $j \leqslant l$ or $j > l$ we introduce the diagonal matrices Γ_+ and Γ_- according to

$$\Gamma_+ = \begin{vmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \phi_{l+1} & & \\ & & & & \ddots & \\ & & & & & \phi_N \end{vmatrix}$$

$$\Gamma_- = \begin{vmatrix} \phi_1^{-1} & & & & & \\ & \ddots & & & & \\ & & \phi_l^{-1} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{vmatrix} \tag{4.14}$$

So that $\Gamma = \Gamma_+ \Gamma_-^{-1}$ and Eq. (4.13) can be written

$$\Gamma_+^2 \, e = (\Gamma_-^2 + \lambda \Gamma_+^2 \, B)\eta \,. \tag{4.15}$$

Denoting that Π_l the projector defined by

$$(\Pi_l)_{ij} = \begin{cases} \delta_{ij} & i,j \leqslant l \\ 0 & \text{otherwise} \end{cases} . \tag{4.16}$$

We observe that $\Gamma_+ \to \Pi_l, \Gamma_- \to (1 - \Pi_l)$ for $t \to +\infty$. It is easy to check that the limit η_∞ of η for $t \to +\infty$ exists since the limits of $\Pi_l \, \eta$ and $(1 - \Pi_l)\eta$ both exist. Taking the limit of (4.15) for $t \to +\infty$, we obtain

$$\Pi_l e = (1 - \Pi_l)\eta_\infty + \lambda \Pi_l B \eta_\infty \,, \tag{4.17}$$

which imples

$$(1 - \Pi_l)\eta_\infty = 0; \quad \Pi_l e = \lambda \Pi_l B \Pi_l \, \eta_\infty \,. \tag{4.18}$$

Letting $B^{-1}(l)$ be a matrix whose first principal minor of order l is the inverse of the first principal minor of order l of B, namely,

$$\Pi_l B^{-1}(l) \Pi_l \, B \Pi_l = \Pi_l \,, \tag{4.19}$$

one can express the solution of the second equation in (4.18) and finally obtain

$$\begin{aligned} U_\infty &= \langle e,(1 - \Pi_l)\eta_\infty \rangle + \langle e,\Pi_l \, \eta_\infty \rangle \\ &= (1/\lambda)\langle \Pi_l e, B^{-1}(l) \, \Pi_e \, e \rangle \end{aligned} \tag{4.20}$$

By using the results of Appendix A the last scalar product can be explicity evaluated in agreement with (4.8).

C. $v = v_l$

In this case ϕ_l is independent of t and we introduce the diagonal matrices Γ_+ according to (4.14), Γ_- and Γ_0 according to

$$\Gamma_- = \begin{vmatrix} \phi_l^{-1} & & & & & \\ & \ddots & & & & \\ & & \phi_{l-1}^{-1} & & & \\ & & & 0 & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{vmatrix} \qquad \Gamma_0 = \begin{vmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \phi_l^{-1}(\xi,0) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{vmatrix} , \tag{4.21}$$

so that $\Gamma^{-2} = \Gamma_+^{-2} \Gamma_-^2 + \Gamma_0^2$. Since the asymptotic limit of Γ_+ and Γ_- is the same as is the previous case, η, which satisfies the equation

$$\Gamma_+^2 \, e = [\Gamma_-^2 + \Gamma_+^2 \, \Gamma_0^2 + \lambda \Gamma_+^2 \, B \,] \, \eta \tag{4.22}$$

has a limit of η_∞ for $t \to +\infty$. It is easy to verify that

$$(1 - \Pi_l)\,\eta_\infty = 0; \Pi_l \, e = \Pi_l [\Gamma_0^2 + \lambda B \,] \, \Pi_l \, \eta_\infty . \tag{4.23}$$

As a consequence, letting $C^{-1}(l)$ be a matrix such that

$$\Pi_l \, C^{-1}(l) \, \Pi_l (\Gamma_0^2 + \lambda B)\Pi_l = \Pi_l \,, \tag{4.24}$$

one obtains

$$U_\infty = \langle e,\eta_\infty \rangle = \langle \, \Pi_l e, C^{-1}(l) \, \Pi_l e \rangle . \tag{4.25}$$

Using the results of the Appendix one finds that (4.25) is in agreement with (4.8).

D. $v < v_N$.

In this case all the ϕ_j diverge for $t \to +\infty$ so that $\Gamma^{-2} \to 0$. As a consequence, Eq. (4.13) defines the limit η_∞

$$e = \lambda B \eta_\infty \tag{4.26}$$

and U_∞ is given by

$$U_\infty = \frac{1}{\lambda} \langle e, B^{-1} e \rangle \,, \tag{4.27}$$

still in agreement with (4.8).

5. PRELIMINARY RESULTS FOR THE P. A. ASYMPTOTICS

In order to investigate the asymptotic behavior of $[(n-1)/n]_U$ for $t \to +\infty$ we replace (2.21) and (2.22) by

$$[(n-1)/n]_U = \langle e,\hat{\eta} \rangle \,, \tag{5.1}$$

where $\hat{\eta} = \Gamma\hat{\psi}$ satisfies the equation

$$e = (\Gamma^{-2} + \lambda TBT^{+})\hat{\eta}\,. \tag{5.2}$$

The matrix T is defined by

$$T = \Gamma^{-1}P\Gamma \tag{5.3}$$

and enjoys the following properties:

$$T^2 = T, \quad Te = e. \tag{5.4}$$

As we did previously for the exact solution we evaluate the P.A. for $x = vt + \xi$ and we consider the behavior for fixed ξ and large t.

Theorem 2: The vector $\hat{\eta}$ satisfies $T^{+}\hat{\eta} = \hat{\eta}$ and is uniformly bounded for any finite t.

Proof: From (5.2) and (5.4) we obtain

$$e = (T\Gamma^{-2} + \lambda TBT^{+})\,\hat{\eta} = (\Gamma^{-2} + \lambda TBT^{+})T^{+}\hat{\eta}\,, \tag{5.5}$$

and subtracting from (5.2)

$$(\Gamma^{-2} + \lambda TBT^{+})(\hat{\eta} - T^{+}\hat{\eta}) = 0\,. \tag{5.6}$$

Since $\Gamma^{-2} + \lambda TBT^{+}$ is, for any finite t, positive definite and therefore invertible, (5.6) implies that $T^{+}\hat{\eta} = \hat{\eta}$.

In order to prove the other property we remark that for any finite t, using the Schwartz inequality, (5.2) and $T^{+}\hat{\eta} = \hat{\eta}$ we have

$$\begin{aligned}\|\hat{\eta}\|\|e\| \geqslant |\langle e,\hat{\eta}\rangle| &= \langle\hat{\eta},(\Gamma^{-2} + \lambda B)\hat{\eta}\rangle \\ &\geqslant \lambda\mu_0[B]\|\hat{\eta}\|^2\,,\end{aligned} \tag{5.7}$$

where $\mu_0[\,\cdot\,]$ denotes the smallest eigenvalue of a symmetric matrix and the inequality $\mu_0[\Gamma^{-2} + \lambda B] \geqslant \lambda\mu_0[B]$ follows from the Rayleigh–Ritz principle. As a consequence

$$\|\hat{\eta}\| \leqslant \frac{\|e\|}{\lambda\mu_0[B]} = \frac{\|e\|\|B^{-1}\|}{\lambda} \quad \text{Q.E.D.} \tag{5.8}$$

The next step towards the P. A. asymptotics is the asymptotic behavior of T that will be stated by Theorem 4. All the results ranging from Eq. (5.12) to (5.48) prepare the proof of this Theorem and could be skipped in a quick reading.

A. Ordering of ϕ_j

The asymptotic limit of $[(n-1)/n]_U$ can be performed if we know the behavior of P and T for large t. For this purpose the basis (2.19) on which P projects must be orthonormalized. The first step, however, is the ordering of the ϕ_j for v fixed and t large. This behavior is determined, according to (4.6) by the arguments $k_j(v_j - v)$.

Letting

$$f(y) = y(y^2 - v)\,, \tag{5.9}$$

we see that for t large enough

$$\begin{cases} f(\mathrm{k}_i) > f(k_j) \Rightarrow \lim\limits_{t \to +\infty} \dfrac{\phi_j}{\phi_i} = 0 \\ f(k_i) = f(k_j) \Rightarrow \dfrac{\phi_j}{\phi_i} = \text{const} \end{cases} \tag{5.10}$$

When the first condition in (5.10) occurs we shall also use the notation $\phi_i > {}^{*}\phi_j$ while the second will be denoted by $\phi_i \sim \phi_j$. Since the function (5.9) is monotonic increasing for $y > (v)^{1/2}$ then if $k_1 > k_2 > \cdots > k_N > (v)^{1/2}$ the sequence of ϕ_j is ordered $\phi_1 > {}^{*}\phi_2 > {}^{*}\ldots > {}^{*}\phi_N$.

When $k_1 > k_2 > \cdots > k_l > (v)^{1/2} > k_{l+1} > \cdots > k_N$, the first l of the ϕ are ordered; the remaining are not.

However we can relabel the k_j for $j > l$ so that ϕ_j are ordered according to $\phi_1 > {}^{*}\phi_2 > {}^{*}\cdots > {}^{*}\phi_l > {}^{*}\phi_{l+1} \tilde{>} {}^{*}\phi_{l+2} \tilde{>} {}^{*}\cdots \tilde{>} {}^{*}\phi_N$. After ϕ_{l+1} the occurrence of two ϕ with the same behavior is not excluded. However, there can be at most two ϕ with the same behavior since there are only two points in the interval $]0,(v)^{1/2}[$ at which $f(y)$ can assume the same value. Therefore, after reordering the ϕ one must have $\phi_j > {}^{*}\phi_{j+2}$ for $j > l$. Finally if $(v)^{1/2} > k_N$ all the k must relabeled if we wish to order the ϕ. To conclude, we write

$$v_{l+1} < v < v_l \Rightarrow \begin{cases} \phi_1 > {}^{*}\phi_2 > {}^{*}\cdots > {}^{*}\phi_l > {}^{*}\phi_{l+1} \tilde{>} {}^{*}\phi_{l+2} \tilde{>} {}^{*}\cdots \tilde{>} {}^{*}\phi_N, \\ \phi_j > {}^{*}\phi_{j+2}, \quad j > l. \end{cases} \tag{5.11}$$

The reordering of the ϕ obviously can change by changing v.

B. Changes of basis

Let us define an array $n \times N$ formed with the components of the vectors $\phi^{(i)}$ for $i = 1,\ldots,n$, and denote it by X

$$X = \left|\left| \begin{matrix} \phi_1^{(1)} & \cdots & \phi_N^{(1)} \\ \vdots & & \vdots \\ \phi_1^{(n)} & \cdots & \phi_N^{(n)} \end{matrix} \right|\right|. \tag{5.12}$$

Let us denote by $X^{(p)}$, for $1 \leqslant p \leqslant n$, the matrix $n \times N$ formed by vectors $\phi^{(i,p)}$ $i = 1,\ldots,n$, obtained by linear combinations of the original vectors $\phi^{(i)}$ $i = 1,\ldots,n$, such that $X_{ij}^{(p)} = \delta_{ij}$ for $i = 1,\ldots,p, j = 1,\ldots,p$. In extended form we write

$$X^{(p)} = \begin{vmatrix} & & & \phi_{p+1}^{(1,p)} & \cdots & \phi_N^{(1,p)} \\ & \mathbb{1}^{(p)} & & \vdots & & \vdots \\ & & & \phi_{p+1}^{(p,p)} & \cdots & \phi_N^{(p,p)} \\ \phi_1^{(p+1,p)} & \cdots & \phi_p^{(p+1,p)} & \phi_{p+1}^{(p+1,p)} & \cdots & \phi_N^{(p+1,p)} \\ \vdots & & \vdots & \vdots & & \vdots \\ \phi_1^{(n,p)} & \cdots & \phi_p^{(n,p)} & \phi_{p+1}^{(n,p)} & \cdots & \phi_N^{(n,p)} \end{vmatrix} . \tag{5.13}$$

The recursion algorithm from $X^{(p-1)}$ to $X^{(p)}$ explicitly reads

$$\begin{aligned} &\omega^{(p)} = \phi^{(p,p-1)} - \sum_{i=1}^{p-1} \phi^{(i,p-1)} \phi_i^{(p,p-1)}, \\ &\phi^{(p,p)} = \omega^{(p)} / \omega_p^{(p)}, \\ &\phi^{(i,p)} = \phi^{(i,p-1)} - \phi^{(p,p)} \phi_p^{(i,p-1)}, \quad i = 1,\ldots,p-1, \\ &\phi^{(i,p)} = \phi^{(i,p-1)}, \quad i \geqslant p+1 . \end{aligned} \tag{5.14}$$

For convenience we shall also denote X by $X^{(0)}$ and $\phi^{(i)}$ by $\phi^{(i,0)}$. Let $D^{(p)}(1,\ldots,p;j_1,\ldots,j_p)$ be the determinant of the minor of order p of $X^{(p)}$ formed by the first p rows and the columns $j_1,\ldots,j_p$ and let $D(1,\ldots,p;j_1,\ldots,j_p)$ be the corresponding determinant for X, that is,

$$D(1,\ldots,p;j_1,\ldots,j_p) = \begin{vmatrix} \phi_{j_1}^{(1)} & \cdots & \phi_{j_p}^{(1)} \\ \vdots & & \vdots \\ \phi_{j_1}^{(p)} & \cdots & \phi_{j_p}^{(p)} \end{vmatrix} . \tag{5.15}$$

Since $X^{(p)}$ is obtained from $X^{(p-1)}$ by linear combinations of its rows and by dividing the pth row by $\omega_p^{(p)}$ it is evident by recursion that $D^{(p)}(1,\ldots,p;j_1,\ldots,j_p)$ will be equal to $D(1,\ldots,p;j_1,\ldots,j_p)$ divided by $\Pi_{i=1}^{p} \omega_i^{(i)}$. The last product is different from zero. Indeed, by comparing (3.2) and (5.15), Theorem 1 states that $D(1,\ldots,p;1,\ldots,p) \neq 0$, while from the definitions of $X^{(p)}$ and $D^{(p)}$ it follows that $D^{(p)}(1,\ldots,p;1,\ldots,p) = 1$ so that

$$\prod_{i=1}^{p} \omega_i^{(i)} = D(1,\ldots,p;1,\ldots,p) \neq 0 . \tag{5.16}$$

Therefore we can write

$$D^{(p)}(1,\ldots,p;j_1,\ldots,j_p) = \frac{D(1,\ldots,p;j_1,\ldots,j_p)}{D(1,\ldots,p;1,\ldots,p)} \tag{5.17}$$

and notice that the following relation holds:

$$\begin{aligned} \phi_j^{(p,p)} &= D^{(p)}(1,\ldots,p;1,\ldots,p-1,j) \\ &= \frac{D(1,\ldots,p;1,\ldots,p-1,j)}{D(1,\ldots,p;1,\ldots,p-1,p)}, \quad j = 1,\ldots,N. \end{aligned} \tag{5.18}$$

The basic strategy to obtain an asymptotic estimate of $\phi_i^{(i,p)}$ for $i = 1,\ldots,N$ is the following. First we establish a recursion relation relating the D of order p to the D of order $p-1$. Then we can estimate the ratios of D, appearing in the r.h.s. of (5.18) in terms of the ϕ_j only, for t large, and obtain the required behavior of $\phi_j^{(p,p)}$. Through the third of the relations (5.14) we then obtain the estimate on the $\phi^{(i,p)}$ for $i = 1,\ldots,p-1$. A Gram–Schmidt orthogonalization will change the final basis given by $\phi^{(i,n)}$, $i = 1,\ldots,n$, to a new basis $\tau^{(i)}$, $i = 1,\ldots,n$, for which the asymptotic estimates can be given. Finally the asymptotic structure of P and T are determined. For the next developments we need to define the determinant of order p $K(r_1,\ldots,r_{p-1};j_1,\ldots,j_p)$ according to

$$K(r_1,\ldots,r_{p-1};j_1,\ldots,j_p) = \begin{vmatrix} 1 & \cdots & 1 \\ \dfrac{1}{k_{r_1}+k_{j_1}} & \cdots & \dfrac{1}{k_{r_1}+k_{j_p}} \\ \vdots & & \vdots \\ \dfrac{1}{k_{r_{p-1}}+k_{j_1}} & \cdots & \dfrac{1}{k_{r_{p-1}}+k_{j_p}} \end{vmatrix} \tag{5.19}$$

Lemma 1: The determinants D of orders p and $p-1$ are related by

$$\begin{aligned} &D(1,\ldots,p;j_1,\ldots,j_p) \\ &\quad = \phi_{j_1} \cdots \phi_{j_p} \sum_{\{\tilde{r}_1,\ldots,\tilde{r}_{p-1}\}_1^N} \phi_{\tilde{r}_1} \cdots \phi_{\tilde{r}_{p-1}} \\ &\qquad \times D(1,\ldots,p-1;\tilde{r}_1,\ldots,\tilde{r}_{p-1}) K(\tilde{r}_1,\ldots,\tilde{r}_{p-1};j_i,\ldots,j_p) , \end{aligned} \tag{5.20}$$

where $\{\tilde{r}_1,\ldots,\tilde{r}_{p-1}\}_1^N$ is an ordered set of $p-1$ natural numbers, all distinct, chosen among 1,2,...,N and the sum runs over all possible such sets.

Proof: We first observe that from (2.12) and (3.1)

$$\phi_j^{(i)} = \sum_{r=1}^{N} \frac{\phi_j \phi_r}{k_j + k_r} \phi_r^{(i-1)} . \tag{5.21}$$

From the definition of determinant one can write

$$\begin{aligned} &D(1,\ldots,p;j_1,\ldots,j_p) \\ &\quad = \sum_{r_1=1}^{N} \cdots \sum_{r_p=1}^{N} \phi_{r_1} \phi_{r_1}^{(1)} \phi_{r_2} \phi_{r_2}^{(2)} \cdots \phi_{r_{p-1}} \phi_{r_{p-1}}^{(p-1)} \\ &\qquad \times \sum_{\{s_1,\ldots,s_p\}_1^p} (-1)^P \phi_{j_{s_1}} \frac{\phi_{j_{s_2}}}{k_{j_{s_1}} + k_{r_1}} \cdots \frac{\phi_{j_{s_p}}}{k_{j_{s_p}} + k_{r_{p-1}}} , \end{aligned} \tag{5.22}$$

where $\{s_1,\ldots,s_p\}_1^p$ is a permutation of 1,...,p and P is its parity. Accounting for $\phi_{j_{s_1}} \phi_{j_{s_2}} \cdots \phi_{j_{s_p}} = \phi_{j_1} \phi_{j_2} \cdots \phi_{j_p}$ and for (5.19) we have

$$\begin{aligned} &D(1,\ldots,p;j_1,\ldots,j_p) \\ &\quad = \sum_{r_1=1}^{N} \cdots \sum_{r_{p-1}=1}^{N} \phi_{r_1} \phi_{r_1}^{(1)} \phi_{r_2} \phi_{r_2}^{(2)} \cdots \phi_{r_{p-1}} \phi_{r_{p-1}}^{(p-1)} \\ &\qquad \times \phi_{j_1} \phi_{j_2} \cdots \phi_{j_p} K(r_1,\ldots,r_{p-1};j_1,\ldots j_p) . \end{aligned} \tag{5.23}$$

We notice that the terms of the sum in which two or more of the integers $r_1,...,r_{p-1}$ are equal, vanish since in this case the determinant K has two equal rows. Therefore we can write a sum over the distinct combinations of ordered $p-1$ integers $\{\tilde r_1,...,\tilde r_{p-1}\}_1^N$ out of 1,...,N and a sum over all possible permutations $\{r_1,...,r_{p-1}\,|\tilde r_1,...,\tilde r_{p-1}\}$ of $\tilde r_1,...,\tilde r_{p-1}$. Letting P be the parity of this permutation we consider the identity $K(r_1,...,r_{p-1};j_1,...,j_p)=(-1)^P K(\tilde r_1,...,\tilde r_{p-1};j_1,...j_p)$ so that we can write

$$\begin{aligned}D(1,...,p;j_1,...,j_p)&=\phi_{j_1}\cdots\phi_{j_p}\\&\times\sum_{\{\tilde r_1,...,\tilde r_{p-1}\}_1^N}K(\tilde r_1,...,\tilde r_{p-1};j_1,...,j_p)\cdot\\&\times\sum_{\{r_1,...,r_{p-1}\,|\tilde r_1,...,\tilde r_{p-1}\}}(-1)^P\phi_{r_1}\phi_{r_2}\cdots\phi_{r_{p-1}}\\&\times\phi_{r_1}^{(1)}\phi_{r_2}^{(2)}\cdots\phi_{r_{p-1}}^{(p-1)}\\&=\phi_{j_1}\cdots\phi_{j_p}\sum_{\{\tilde r_1,...,\tilde r_{p-1}\}_1^N}\phi_{\tilde r_1}\cdots\phi_{\tilde r_{p-1}}\\&\times D(1,...,p-1;\tilde r_1,...,\tilde r_{p-1})\,K(\tilde r_1,...,\tilde r_{p-1};j_1,...,j_p),\end{aligned}\tag{5.24}$$

where we used the identity

$$\phi_{r_1}\cdots\phi_{r_{p-1}}=\phi_{\tilde r_1}\cdots\phi_{\tilde r_{p-1}}\,.\quad\text{Q.E.D.}\tag{5.25}$$

Lemma 2: For the determinants D the following estimates hold for t large enough:

$$\begin{aligned}m_p\frac{\phi_{j_1}\cdots\phi_{j_p}}{\phi_1\cdots\phi_p}&\leqslant\frac{|D(1,...,p;j_1,...,j_p)|}{|D(1,...,p;1,...,p)|}\\&\leqslant M_p\frac{\phi_{j_1}\cdots\phi_{j_p}}{\phi_1\cdots\phi_p},\end{aligned}\tag{5.26}$$

where M_p and m_p converge to the same limit for $t\to+\infty$. More precisely one can write

$$M_p=C_p\frac{1+\epsilon_p}{1-\epsilon_p}\qquad m_p=C_p\frac{1-\epsilon_p}{1+\epsilon_p},\tag{5.27}$$

where C_p is a constant and ϵ_p goes to zero with $t\to+\infty$ as ϕ_p^2/ϕ_{p-1}^2 if $\phi_{p-1}>^*\phi_p$; as $\alpha\phi_{p-1}^2/\phi_{p-2}^2+\beta\phi_{p+1}^2/\phi_{p-1}^2$ if $\phi_{p-1}\sim\phi_p$. In addition one has $m_1=M_1=1$, $\epsilon_1=0$.

Proof: We use an inductive argument. The result is obvious for $p=1$ since $D(1,j)=\phi_j$. Assuming that (5.26) holds for $p-1$ we use the result of Lemma 1, separating in the sum the leading term from the remainder. If $\phi_{p-1}>^*\phi_p$ there is just one leading term which we denote by $A(j_1,...,j_p)$, while the remainder is $B(j_1,...,j_p)$, namely

$$\begin{aligned}D(1,...,p;j_1,...,j_p)&=\phi_{j_1}\cdots\phi_{j_p}\\&\times[A(j_1,...,j_p)+B(j_1,...,j_p)],\end{aligned}\tag{5.28}$$

where

$$\begin{aligned}A(j_1,...,j_p)&=\phi_1\cdots\phi_{p-1}D(1,...,p-1;1,...,p-1)\\&\times K(1,...,p-1;j_1,...,j_p)\end{aligned}\tag{5.29}$$

and

$$\begin{aligned}B(j_1,...,j_p)&=\sum_{\{\tilde r_1,...,\tilde r_{p-1}\}_1^N\neq\{1,...,p-1\}}\phi_{\tilde r_1}\cdots\phi_{\tilde r_{p-1}}\\&\times D(1,...,p-1;\tilde r_1,...,\tilde r_{p-1})\\&\times K(\tilde r_1,...,\tilde r_{p-1};j_1,...,j_p).\end{aligned}\tag{5.30}$$

As a consequence, using (5.26) for $p-1$, we have

$$\begin{aligned}\frac{|B(j_1,...,j_p)|}{|A(j_1,...,j_p)|}&\leqslant M_{p-1}\sum_{\{\tilde r_1,...,\tilde r_{p-1}\}_1^N\neq\{1,...,p-1\}}\\&\times\frac{\phi_{\tilde r_1}^2\cdots\phi_{\tilde r_{p-1}}^2}{\phi_1^2\cdots\phi_{p-1}^2}\frac{|K(\tilde r_1,...,\tilde r_{p-1};j_1,...,j_p)|}{|K(1,...,p-1;j_1,...,j_p)|}\\&\leqslant M_{p-1}\binom{p-1}{N}\chi_p\frac{\phi_p^2}{\phi_{p-1}^2}=\epsilon_p,\end{aligned}\tag{5.31}$$

where by χ_p we denote the largest of the ratios of K after noticing that the denominator is nonzero (see Appendix B). We conclude that (5.26) holds where

$$C_p=\frac{|K(1,...,p-1;j_1,...,j_p)|}{|K(1,...,p-1;1,...,p)|}.\tag{5.32}$$

If $\phi_{p-1}\sim\phi_p$ we have two leading terms in the r.h.s. of (5.20) namely, $\phi_1\cdots\phi_{p-2}\,\phi_{p-1}$ and $\phi_1\cdots\phi_{p-2}\,\phi_p$. As a consequence, we have to go one step further and use (5.20) twice in order to write $D(1,...,p;j_1,...,j_p)$ as a sum of $D(1,...,p-2;\tilde r_1,...,\tilde r_{p-2})$. Of course this case can occur only for $p\geqslant3$. If $p=2$ then the factors of the equivalent leading terms are constants and one does not need to go one step further. Let us denote by $A(j_1,...,j_p)$ the leading term and by $B_1(j_1,...,j_p)$, $B_2(j_1,...,j_p)$ two remainders according to

$$\begin{aligned}D(1,...,p;j_1,...,j_p)&=\phi_{j_1}\cdots\phi_{j_p}\\&\times[A(j_1,...,j_p)+B_1(j_1,...,j_p)+B_2(j_1,...,j_p)],\end{aligned}\tag{5.33}$$

where

$$\begin{aligned}A(j_1,...,j_p)&=\phi_1^2\cdots\phi_{p-2}^2\,\phi_{p-1}^2\,\phi_1\cdots\phi_{p-2}\\&\times D(1,...,p-2;1,...,p-2)\\&\times\Delta(1,...,p-2;j_1,...,j_p)\end{aligned}\tag{5.34}$$

and

$$\begin{aligned}B_1(j_1,...,j_p)&=\phi_1^2\cdots\phi_{p-1}^2\,\phi_{p-1}^2\sum_{\{\tilde r_1,...,\tilde r_{p-2}\}_1^N\neq\{1,...,p-2\}}\\&\times\phi_{\tilde r_1}\cdots\phi_{\tilde r_{p-2}}D(1,...,p-2;\tilde r_1,..\tilde r_{p-2})\\&\times\Delta(\tilde r_1,...,\tilde r_{p-2};j_1,..j_p)\end{aligned}\tag{5.35}$$

and

$$\begin{aligned}B_2(j_1,...,j_p)&=\sum_{\substack{\{\tilde r_1,...,\tilde r_{p-1}\}_1^N\\\neq\{1,...,p-2,p-1\}\\\neq\{1,...,p-2,p\}}}\phi_{\tilde r_1}^2\cdots\phi_{\tilde r_{p-1}}^2\\&\times\sum_{\{\tilde s_1,...,\tilde s_{p-2}\}_1^N}\phi_{\tilde s_1}\cdots\phi_{\tilde s_{p-2}}D(1,...,p-2;\tilde s_1,...,\tilde s_{p-2})\\&\times K(\tilde s_1,...,\tilde s_{p-2};\tilde r_1,...,\tilde r_{p-1})K(\tilde r_1,...,\tilde r_{p-1};j_1,...,j_p).\end{aligned}\tag{5.36}$$

The constants Δ are defined by

$$\Delta(\tilde{r}_1,...,\tilde{r}_{p-2};j_1,...,j_p) = K(\tilde{r}_1,...,\tilde{r}_{p-2};\,1,...,p-2,p-1) \times K(1,...,p-2,p-1;j_1,...,j_p) + \frac{\phi_p^2}{\phi_{p-1}^2} K(\tilde{r}_1,...,\tilde{r}_{p-2};1,...,p-2,p) \times K(1,...,p-2,p;j_1,...,j_p), \tag{5.37}$$

where the ratio ϕ_p/ϕ_{p-1} is independent of t since $\phi_{p-1} \sim \phi_p$ and the coefficient of the leading term $\Delta(1,...,p-2;j_1,...j_p)$ is different from zero as shown in Appendix B. Using (5.26) for $p-2$ we have

$$\frac{|B_1(j_1,...,j_p)|}{|A(j_1,...,j_p)|} \leqslant M_{p-2} \sum_{\{\tilde{r}_1,...,\tilde{r}_{p-2}\}_1^N \neq \{1,...,p-2\}} \frac{\phi_{\tilde{r}_1}^2 \cdots \phi_{\tilde{r}_{p-2}}^2}{\phi_1^2 \cdots \phi_{p-2}^2} \frac{|\Delta(\tilde{r}_1,...,\tilde{r}_{p-2};j_1,...,j_p)|}{|\Delta(1,...,p-2;j_1,...,j_p)|} \leqslant M_{p-2} \binom{p-2}{N} \chi_p^{(1)} \frac{\phi_{p-1}^2}{\phi_{p-2}^2} = \epsilon_p^{(1)}, \tag{5.38}$$

where $\chi_p^{(1)}$ is the largest of the ratios of Δ occuring in the sum. For the ratio $|B_2/A|$ we obtain

$$\frac{|B_2(j_1,...,j_p)|}{|A(j_1,...,j_p)|} \leqslant M_{p-2} \sum_{\substack{\{\tilde{r}_1,...,\tilde{r}_{p-1}\}_1^N \\ \neq\{1,...,p-2,p-1\} \\ \neq\{1,...,p-2,p\}}} \frac{\phi_{\tilde{r}_1}^2 \cdots \phi_{\tilde{r}_{p-1}}^2}{\phi_1^2 \cdots \phi_{p-1}^2} \times \sum_{\{\tilde{s}_1,...,\tilde{s}_{p-2}\}_1^N} \frac{\phi_{\tilde{s}_1}^2 \cdots \phi_{\tilde{s}_{p-2}}^2}{\phi_1^2 \cdots \phi_{p-2}^2} \times \frac{|K(s_1,...,s_{p-2};\tilde{r}_1,...,\tilde{r}_{p-1})K(\tilde{r}_1,...,\tilde{r}_{p-1};j_1,...,j_p)|}{|\Delta(1,...,p-2;j_1,...,j_p)|} \leqslant M_{p-2} \binom{p-1}{N} \binom{p-2}{N} \chi_p^{(2)} \frac{\phi_{p+1}^2}{\phi_{p-1}^2} = \epsilon_p^{(2)}, \tag{5.39}$$

where $\chi_p^{(2)}$ is the largest of the ratios of K and Δ occurring in the sum. Finally, (5.27) is recovered with $\epsilon_p = \epsilon_p^{(1)} + \epsilon_p^{(2)}$ and

$$C_p = \frac{|\Delta(1,...,p-2;j_1,...,j_p)|}{|\Delta(1,...,p-2;\,1,...,p)|}. \tag{5.40}$$

We can now quote the basic result stating the following theorem.

Theorem 3: The asymptotic behavior of the vectors $\phi^{(i,p)}$ whose components define the matrix $X^{(p)}$ is given by

$$\phi_j^{(i,p)} = O(\phi_j/\phi_i), \quad p < j \leqslant N, \; i \leqslant p \leqslant n. \tag{5.41}$$

Proof: For $i = p$, using (5.18) and the results of Lemma 2, the estimate (5.41) is satisfied. For $i < p$ we use the third equation of (5.14) and an inductive argument. Indeed, for $p = 1$ the relation (5.41) is trivially verified since $\phi_j^{(1,1)} = \phi_j/\phi_1$ by construction. Assuming the relation valid for $p-1$ we have

$$\phi_j^{(i,p)} = \phi_j^{(i,p-1)} - \phi_j^{(p,p)} \phi_p^{(i,p-1)} = O(\phi_j/\phi_i) + O(\phi_j/\phi_p) O(\phi_p/\phi_i) = O(\phi_j/\phi_i), \quad i \leqslant p-1, \; j > p. \tag{5.42}$$

C. The orthogonal basis

The orthogonalization procedure is applied to the final set of vectors $\phi^{(i,n)}$ for which we can write when t is large, see (5.13) and (5.41),

$$\phi_j^{(i,n)} = \begin{cases} O(\phi_j/\phi_i) & j > n \\ \delta_{ij} & j \leqslant n \end{cases} \quad i \leqslant n. \tag{5.43}$$

Letting $\tau^{(i)}$ be the orthogonal vectors, we can obtain them by using the Gram–Schmidt recursive procedure

$$\tau^{(i)} = \phi^{(i,n)} - \sum_{l=1}^{i-1} \hat{\tau}^{(l)} \langle \hat{\tau}^{(l)}, \phi^{(i,n)} \rangle \; i = 1,...,n, \tag{5.44}$$

where $\hat{\tau}^{(l)}$ are the normalized vectors

$$\hat{\tau}^{(l)} = \tau^{(l)}/\|\tau^{(l)}\|. \tag{5.45}$$

The asymptotic behavior of $\tau^{(i)}$ is easily computed from (5.43) and one has to distinguish four different regions

$$\tau_j^{(i)} = \begin{cases} O(\phi_{n+1}^2/\phi_i\,\phi_j) & j < i, & \text{I} \\ 1 & j = 1, & \text{II} \\ 0 & i < j \leqslant n, & \text{III} \\ O(\phi_j/\phi_i) & n < j \leqslant N, & \text{IV} \end{cases}. \tag{5.46}$$

In addition the norm of $\tau^{(l)}$ has the following estimate:

$$\|\tau^{(l)}\|^2 = 1 + O(\phi_{n+1}^2/\phi_l^2). \tag{5.47}$$

We are in a position now to specify the asymptotic structure of the matrix T defined by (5.3). We can also write

$$T = \sum_{p=1}^{n} T^{(p)},$$

where

$$T_{ij}^{(p)} = \phi_i^{-1} \tau_i^{(p)} \tau_j^{(p)} \phi_j \, \|\tau^{(p)}\|^{-2}. \tag{5.48}$$

Theorem 4: The matrix T has the following asymptotic structure:

$$T = \Pi_n + (1 - \Pi_n) Z \Pi_n + \mathscr{E} \tag{5.49}$$

if $\phi_n >^* \phi_{n+1}$, that is, $\phi_{n+1}/\phi_n \to 0$ for $t \to +\infty$;

$$T = \Pi_{n-1} + (1 - \Pi_{n-1}) Z \Pi_{n+1} + \mathscr{E} \tag{5.50}$$

if $\phi_n \sim \phi_{n+1}$, that is, ϕ_n/ϕ_{n+1} is independent of t. The matrix Z is bounded while any $\mathscr{E}_{ij}$ is exponentially small for $t \to +\infty$.

Proof: In order to estimate the matrix elements $T_{ij}^{(p)}$ we have to distinguish 16 possibilities according to the region to which the couples of indices (p,i), (p,j) belong [see (5.46)]. Then it is not hard to check that

[I–I]	$i<p,\ j<p$	$T_{ij}^{(p)}=O\left(\frac{\phi_{n+1}^4}{\phi_p^2\phi_i^2}\right)$
[I–II]	$i<p,\ j=p$	$T_{ij}^{(p)}=O\left(\frac{\phi_{n+1}^2}{\phi_i^2}\right)$
[I–III]	$i<p,\ p<j\leqslant n$	$T_{ij}^{(p)}=0$
[I–IV]	$i<p,\ n<j\leqslant N$	$T_{ij}^{(p)}=O\left(\frac{\phi_{n+1}^2\phi_j^2}{\phi_p^2\phi_i^2}\right)$
[II–I]	$i=p,\ j<p$	$T_{ij}^{(p)}=O\left(\frac{\phi_{n+1}^2}{\phi_p^2}\right)$
[II–II]	$i=p,\ j=p$	$T_{ij}^{(p)}=1+O\left(\frac{\phi_{n+1}^2}{\phi_p^2}\right)$
[II–III]	$i=p,\ p<j\leqslant n$	$T_{ij}^{(p)}=0$
[II–IV]	$i=p,\ n<j\leqslant N$	$T_{ij}^{(p)}=O\left(\frac{\phi_j^2}{\phi_p^2}\right)$
[III–I],[III,II] [III,III],[III,IV]	$p<i\leqslant n,\ 1\leqslant j\leqslant N$	$T_{ij}^{(p)}=0$
[IV–I]	$n<i\leqslant N,\ j<p$	$T_{ij}^{(p)}=O\left(\frac{\phi_{n+1}^2}{\phi_p^2}\right)$
[IV–II]	$n<i\leqslant N,\ j=p$	$T_{ij}^{(p)}=O(1)$
[IV–III]	$n<i\leqslant N,\ p<j\leqslant n$	$T_{ij}^{(p)}=0$
[IV–IV]	$n<i\leqslant N,\ n<j\leqslant N$	$T_{ij}^{(p)}=O\left(\frac{\phi_j^2}{\phi_p^2}\right)$

(5.51)

We remark that in the above expressions if $\phi_n >^* \phi_{n+1}$ then all the terms are exponentially small for t large except for [II,II] and [IV,II]. When $\phi_n \sim \phi_{n+1}$ we have $T_{ij}^{(n)}=O(1)$ in the regions [II,I], [II,II], and [II,IV], [IV,IV] only for $j=n+1$.

6. ASYMPTOTIC LIMIT OF THE $[(n-1)/n]$ P.A.

The asymptotic limit for $t\to\infty$ of the $[(n-1)/n]$ P.A. to the N-soliton solution for $n\leqslant N$ is given by

$$\lim_{t\to\infty}[(n-1)/n]_U=\begin{cases}0 & v>v_1\\ \frac{2}{\lambda}\sum_{j=1}^{l}k_j & v_{l+1}<v<v_l \quad l<n\\ \frac{2}{\lambda}\sum_{j=1}^{l-1}k_j+\frac{2k_l/\lambda}{1+\exp[k_l\xi+2\delta_l]} & v=v_l \quad l\leqslant n\\ \frac{2}{\lambda}\sum_{j=1}^{n}k_j & v<v_n\end{cases} \tag{6.1}$$

where δ_l is defined by (4.9). One can easily recognize that asymptotically the $[(n-1)/n]_U$ is given by an ensemble of n free solitons whose parameters are $k_1>k_2>\ldots>k_n$, the same as the n leading solitons of the exact solution. We shall give a distinct proof of the limit in each of the above regions.

A. $v>v_1$

In this case it is sufficient to recall the inequality

$$0\leqslant[(n-1)/n]_U\leqslant U \tag{6.2}$$

valid for any finite t and to use (4.8) to show that the P.A. converges to zero when $t\to+\infty$

B. $v_{l+1}<v<v_l, \quad l<n$

In this case both $\phi_n >^* \phi_{n+1}$ and $\phi_n\sim\phi_{n+1}$ are allowed. However, from (5.49) and (5.50) the following relations are obtained:

$$\begin{aligned}&\Pi_{n-1}T=\Pi_{n-1}+\Pi_{n-1}\mathscr{E};\\ &T(1-\Pi_{n+1})=\mathscr{E}(1-\Pi_{n+1}).\end{aligned} \tag{6.3}$$

We define Γ_+ and Γ_- according to (4.14) and observe that

$$\Gamma_+^2=\Pi_l+(1-\Pi_l)\mathscr{E}_+;\ \Gamma_-^2=(1-\Pi_l)+\Pi_l\mathscr{E}_-, \tag{6.4}$$

where $\|\mathscr{E}_+\|$ and $\|\mathscr{E}_-\|$ are exponentially small for large t. The basic equation (5.2) then reads

$$\Gamma_+^2\, e=[\Gamma_-^2+\Gamma_+^2\,\lambda TB\,]\hat{\eta}. \tag{6.5}$$

Acting with $(1-\Pi_l)$ on (6.5) we obtain the equation

$$(1-\Pi_l)\hat{\eta}=(1-\Pi_l)\mathscr{E}_+[e-\lambda TB\hat{\eta}]\,, \tag{6.6}$$

which shows that $\|(1-\Pi_l)\hat{\eta}\|$ is exponentially small for large t since both T and $\hat{\eta}$ are uniformly bounded according to Theorems 2 and 4.

Acting with Π_l on Eq. (6.5) and accounting for (6.3), (6.6), and $\Pi_l\,\Pi_{n-1} = \Pi_l$, we have

$$\Pi_l\, e = \lambda \Pi_l\, B \Pi_l\, \hat{\eta} + \rho\,, \tag{6.7}$$

where ρ is a remainder defined by

$$\rho = \lambda \Pi_l\, B(1-\Pi_l)\,\hat{\eta} + \Pi_l(\mathscr{E}_- + \lambda \mathscr{E} B)\hat{\eta}\,, \tag{6.8}$$

and its norm $\|\rho\|$ is exponentially small with t. As a consequence, using the same symbols as in Sec. 4, we finally obtain

$$[(n-1)/n]_U = \frac{1}{\lambda}\,\langle \Pi_l e, B^{-1}(l)\Pi_l e\rangle + r\,, \tag{6.9}$$

where r is exponentially small with t.

C. $v = v_l$, $l \leqslant n$

We introduce the matrices Γ_- and Γ_0 according to (4.21), that is,

$$\Gamma^2_- = (1-\Pi_l) + \Pi_{l-1}\mathscr{E}_-;$$
$$\Gamma^2_0 = (\Pi_l - \Pi_{l-1})\,\phi_l^{-2}(\xi,0)\,. \tag{6.10}$$

If $l < n$ then both $\phi_n >^* \phi_{n+1}$ and $\phi_n \sim \phi_{n+1}$ can occur while for $l = n$ the second is excluded. As a consequence in both cases we have

$$\Pi_l\, T = \Pi_l + \Pi_l \mathscr{E}\,. \tag{6.11}$$

Letting Γ_+ still be defined by (6.4) the basic equation (5.2) now reads

$$\Gamma^2_+\, e = [\Gamma^2_- + \Gamma^2_+\,\Gamma^2_0 + \Gamma^2_+\,\lambda TB\,]\hat{\eta}. \tag{6.12}$$

Applying Π_l and $1-\Pi_l$ respectively to the last equation we finally have

$$(1-\Pi_l)\,\hat{\eta} = (1-\Pi_l)\mathscr{E}_+[-\lambda TB\hat{\eta} + e] \tag{6.13}$$

and

$$\Pi_l e = \Pi_l(\lambda + \Gamma^2_0)\Pi_l\,\hat{\eta} + \rho\,, \tag{6.14}$$

where ρ is given by (6.8). The P.A. then reads

$$[(n-1)/n]_U = \langle \Pi_l e, C^{-1}(l)\Pi_l e\rangle + r, \tag{6.15}$$

where $C(l)$ is defined by (4.24) and is in agreement with (6.1)

D. $v < v_n$

In this case $\phi_n >^* \phi_{n+1}$ and from (5.49) we obtain

$$\Pi_n\, T = \Pi_n + \Pi_n \mathscr{E};\; T(1-\Pi_n) = \mathscr{E}(1-\Pi_n)\,. \tag{6.16}$$

Using the second equation of (6.16) and Theorem 2 we have

$$(1-\Pi_n)\,\hat{\eta} = (1-\Pi_n)\,T^+\hat{\eta} = (1-\Pi_n)\mathscr{E}^+\hat{\eta}\,. \tag{6.17}$$

We observe that $\hat{\eta}$ still satisfies Eq. (6.5) where for $v_{l+1} < v < v_l$ with $l \geqslant n$ the matrices Γ_+ and Γ_- are given by (6.4) while for $v = v_l$ with $l > n$, Γ_+ does not change and Γ_- is given by

$$\Gamma^2_- = (1-\Pi_{l-1}) + (\Pi_l - \Pi_{l-1})\,\phi_l^2(\xi,0) + \Pi_{l-1}\,\mathscr{E}_-\,. \tag{6.18}$$

By applying Π_n to Eq. (6.5) one obtains in each case

$$\Pi_n e = \lambda \Pi_n\, B \Pi_n\,\hat{\eta} + \rho\,, \tag{6.19}$$

where ρ is a remainder given by (6.8) where l is replaced by n. From (6.19) we finally obtain

$$[(n-1)/n]_U = \frac{1}{\lambda}\langle \Pi_n e, B^{-1}(n)\,\Pi_n e\rangle + r,$$

where r is exponentially small with t.

7. CONCLUSIONS

The asymptotic behavior of the $[n - 1/n]$ P.A. to the multisoliton solution for the potential K.d.V. does not seem to be an isolated accident. In fact the structure of exact multisoliton solutions and of the related perturbation series for other equations (modified K.d.V., cubic Schrödinger) would suggest that most of the arguments used here can be extended. For the actual solution of K.d.V. corresponding asymptotic statements could be obtained that are physically quite reasonable. In fact the summation method we propose extracts from a finite number of terms a nonperturbative feature of the solution, as the solitons are.

An important and yet unsolved question concerns the persistence of the described asymptotic behavior also in presence of a background. We hope an answer will be given in spite of the nontrivial mathematical difficulties.

APPENDIX A

Let B, K, E, be $l \times l$ matrices defined by

$$B_{ij} = \frac{1}{k_i + k_j};\; K_{ij} = k_i\,\delta_{ij};\; E_{ij} = 1 \quad i,j = 1,\ldots,l \tag{A1}$$

and e be the vector of $\mathbb{R}^l$ defined by $e_i = 1$ for $i = 1,\ldots,l$. We can prove that the following relation is satisfied:

$$\langle e, B^{-1}e\rangle = T_r(B^{-1}E) = 2\sum_{j=1}^{l} k_j\,. \tag{A2}$$

Indeed if we observe that

$$BK + KB = E, \tag{A3}$$

then $\mathrm{Tr}(B^{-1}E) = 2\,\mathrm{Tr}(K)$ follows. Let C be the $l \times l$ matrix defined by

$$C_{ij} = \lambda B_{ij} + \phi_l^{-2}(\xi,0)\,\delta_{il}\,\delta_{jl}\,, \tag{A4}$$

where $\phi_l^{-2}(\xi,0) = a_l^{-2}\exp(k_l\,\xi)$ in agreement with (2.11). The following relation holds

$$\langle e, C^{-1}e\rangle = \mathrm{Tr}(C^{-1}E) = \frac{2}{\lambda}\sum_{j=1}^{l-1} k_j + \frac{2k_l/\lambda}{1+\exp[k_l\,\xi + 2\delta_l]}\,, \tag{A5}$$

where δ_l is defined by (4.9). In fact we observe that

$$KC + CK = \lambda E + \Phi\,, \tag{A6}$$

where

$$\Phi_{ij} = 2k_l\,\phi_l^{-2}(\xi,0)\delta_{il}\,\delta_{jl}\,. \tag{A7}$$

As a consequence we have

$$\lambda\,\mathrm{Tr}(C^{-1}E) = 2\mathrm{Tr}(K) - \mathrm{Tr}(C^{-1}\Phi) = 2\sum_{j=1}^{l} k_j - 2k_l\,\phi_l^{-2}(\xi,0)(C^{-1})_{ll} \tag{A8}$$

so that accounting for

$$(C^{-1})_{ll} = \frac{\Delta_{l-1}}{\lambda\Delta_l + \phi_l^{-2}(\xi,0)\Delta_{l-1}}\,, \tag{A9}$$

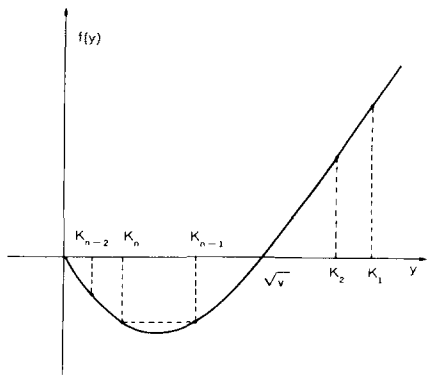


FIG. 3.

where Δ_j is defined by (4.9), after simple algebra (A5) is obtained.

APPENDIX B

We show that the following relations are satisfied:

$$K(1,...,p-1;j_1,...,j_p)\neq 0;\ K(1,...,p-2,p;j_1,...j_p)\neq 0\,. \tag{B1}$$

It suffices to consider the function

$$\Lambda(y)=\begin{vmatrix} 1 & \cdots & 1 \\ \dfrac{1}{k_1+k_{j_1}} & \cdots & \dfrac{1}{k_1+k_{j_p}} \\ \vdots & & \vdots \\ \dfrac{1}{k_{p-2}+k_{j_1}} & \cdots & \dfrac{1}{k_{p-2}+k_{j_p}} \\ \dfrac{1}{y+k_{j_1}} & \cdots & \dfrac{1}{y+k_{j_p}} \end{vmatrix} \tag{B2}$$

and to remark that it is a rational fraction $[(p-2)/p)]$. In fact by expanding the determinant on the last row one would find a rational fraction $[(p-1)/p]$; however the numerator is indeed of order $p-2$ since one verifies that

$$\lim_{y\to\infty} y\Lambda(y)=0\,. \tag{B3}$$

The $p-2$ zeros are at $k_1,k_2,...,k_{p-2}$ and one has

$$\Lambda(k_{p-1})=K(1,...,p-1;j_1,...,j_p);$$
$$\Lambda(k_p)=K(1,...,p-2,p;j_1,...,j_p). \tag{B4}$$

Then (B1) follows provided that $\Lambda(y)$ is not identically zero; this occurrence can be excluded since the residue of $\Lambda(y)$ at $y=-k_{j_p}$ is given by $K(1,...,p-2;j_1,...,j_{p-1})$ and the induction argument applies since $K(1,j_1)=1$.

Next we observe that if $\phi_{p-1}\sim\phi_p$ then, see also Fig. 3, it is evident that $k_1,...,k_{p-2}$ are external to the interval $[k_{p-1},k_p]$; as a consequence $\Lambda(y)$ must have the same sign throughout the interval itself and one has

$$\operatorname{sign}[K(1,...,p-1;j_1,...,j_p)]$$
$$=\operatorname{sign}[K(1,...,p-2,p;\ j_1,...,j_p)]\,.$$

We then consider the function $\Omega(y)$ defined by

$$\Omega(y)=\begin{vmatrix} 1 & \cdots & 1 & 1 \\ \dfrac{1}{2k_1} & \cdots & \dfrac{1}{k_1+k_{p-2}} & \dfrac{1}{k_1+y} \\ \vdots & & \vdots & \vdots \\ \dfrac{1}{k_{p-2}+k_1} & \cdots & \dfrac{1}{2k_{p-2}} & \dfrac{1}{k_{p-2}+y} \end{vmatrix} \tag{B5}$$

which is rational of type $[(p-2)/(p-2)]$ and vanishes for $k_1,k_2,...,k_{p-2}$. Using the previous arguments one concludes that

$$K(1,...,p-2;\ 1,...,p-2,p-1)\neq 0;$$
$$K(1,...,p-2;\ 1,...,p-2,p)\neq 0 \tag{B6}$$

and if $\phi_{p-1}\sim\phi_p$, namely, $k_1,k_2,...,k_{p-2}$ do not belong to the interval $[k_{p-1},k_p]$,

$$\operatorname{sign}[K(1,...,p-2;\ 1,...,p-2,p-1)]$$
$$=\operatorname{sign}[K(1,...,p-2;\ 1,...,p-2,p)]\,. \tag{B7}$$

[1] J. M. Gardner, J. M. Green, M. D. Kruskal, and R. M. Miura, Phys. Rev. Lett. **19**, 1095 (1967).
[2] V. E. Zakharov and A. B. Shabat, Sov. Phys. JETP **34**, 62 (1972).
[3] R. Hirota, Phys. Rev. Lett. **27**, 1192 (1971).
[4] R. Rosales, Stud. Appl. Math. **58**, 117 (1978).
[5] F. Lambert, Z. Phys. C **5**, 147 (1980).
[6] G. Turchetti, Lett. Nuovo Cimento **27**, 107 (1980).
[7] C. Liverani and G. Turchetti, Lett. Nuovo Cimento **30**, 9 (1981).
[8] F. Lambert and M. Musette, Z. Phys. C **10**, 357 (1981).
[9] S. Wall, *Continued fractions* (Van Nostrand, New York, 1946); C. Brezinski, *Padè-Type Approximation and General Orthogonal Polynomials*, International Series of Numerical Mathematics, No. 50 (Birkhauser, Cambridge, MA, 1979); G. Baker, *Essentials of Padè approximants* (Academic, New York, 1977).
[10] V. E. Zakharov, Sov. Phys. JETP. **33**, 538 (1970); M. J. Ablowitz and A. G. Newell, J. Math. Phys. **14**, 1277 (1973).

Manifestly parity invariant electromagnetic theory and twisted tensors

William L. Burke
Lick Observatory, Boards of Studies in Astronomy and Astrophysics and in Physics, University of California, Santa Cruz, California 95064

(Received 15 June 1981; accepted for publication 11 September 1981)

We develop here the calculus of twisted tensors and in particular twisted differential forms, treating them as tensors with complementary orientations. These geometrical objects give us the proper language for electromagnetic theory in a 3-space plus time representation. The parity properties of the fields are simplified and many graphical illustrations are given.

PACS numbers: 03.50.De, 02.40. + m, 04.90. + e

The student's image of the electromagnetism teacher has him wildly waving his *right* hand at every B field and cross product in sight. Since classical electromagnetism is a parity-invariant theory, this handedness must all cancel out. One might think that modern geometric language, especially differential forms, would clear this up, but an inspection of the "egg-crate" pictures in Misner, Thorne, and Wheeler[1] shows this not to be the case. There charge density, for example, is represented by a 3-form. A 3-form has a screw sense, and a right-hand rule is needed to choose one such screw sense to represent positive charge. Now the 4-vector formalism is manifestly parity-invariant, but lacks the numerous advantages of a space/time splitting. Thus arises the question: can one find a 3 + 1 representation that is naturally parity-invariant from start to finish?

This question has a resolution in an old, nearly forgotten class of geometric objects, ones whose transformation law includes the sign of the Jacobian of the transformation in addition to the usual tensorial terms. They were introduced by Weyl[2] and developed by Schouten,[3] who called them W tensors (for Weyl). Synge and Schild[4] refer to them as oriented tensors. DeRahm[5] used differential forms of this type and called them odd differential forms. Sorkin[6] uses these differential forms to discuss magnetic monopoles, and calls them axial forms. Steenrod[7] constructs the bundles for these tensors, as does Eells,[8] who along with Frankel[9] calls them *twisted* tensors. I will use twisted as the most apt description of them.

Twisted tensors are usually introduced abstractly. To a physicist they are sets of components with transformation laws which depend on the sign of the Jacobian of the transformation. To a mathematician, they are cross sections of fiber bundles. In this paper we give a concrete development of twisted tensors, showing the twisted tensors as independent geometric objects. They do not depend upon a choice of orientation, only their conventional representation does. Explicit rules for the operations of the exterior calculus for twisted forms will be given. Pullback in particular will be carefully discussed. Instead of using oriented maps, we will find it convenient for the applications to define it in terms of a transverse orientation for the subspace. The explicit treatment of twisted tensors given here lends itself to simple but accurate graphical representations. In the last part of the paper these twisted forms are applied to classical electromagnetism. The pullback rule, for example, gives us a nice representation of the junction conditions.

TWISTED TENSORS

An orientation for a vector space is a choice of an ordered set of basis vectors to represent positive orientation. A subspace of a vector space can be oriented in two ways. One can orient it as a vector space itself. Alternatively, one can orient its complement in the entire vector space. One needs an orientation for the entire vector space to go back and forth between these two types of orientations, called *inner* and *outer* orientations by Schouten. An outer orientation is often called a *transverse* orientation. Tensors of all types have representations in the tangent space,[1,10] and the above procedure can be used to generate from any oriented tensor a geometric object with complementary orientation. These are the twisted tensors.

A tangent vector is represented by an arrow. A twisted vector in three dimensions is represented by a line with a definite length and a sense of circulation around it. Figure 1 shows vectors and twisted vectors and their law of addition. To appreciate that a twisted vector is an independent notion, consider the problem of finding a continuous nonzero vector field on the Moebius strip which is everywhere transverse to the edge. No such vector field exists, but a twisted vector field with these properties does. See Fig. 2.

A 1-form in three dimensions is represented by a pair of planes with a definite spacing and an outer orientation. A

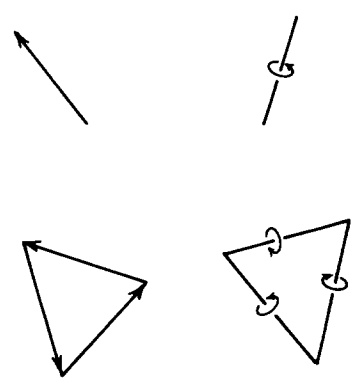


FIG. 1. Vectors and twisted vectors. Their addition is shown by triples which sum to zero.

 0022-2488/83/010065-05$02.50

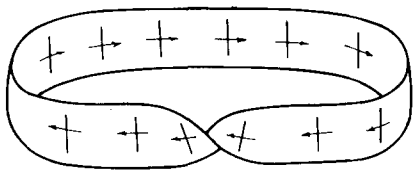


FIG. 2. A continuous twisted vector field on the Moebius strip.

twisted 1-form has instead an inner orientation. See Fig. 3.

Given an orientation of the tangent space, one can map tensors into twisted tensors and vice versa. The basis of twisted vectors generated by the usual right-hand orientation is shown in Fig. 4. I will indicate taking the complementary orientation by putting a tilde over the symbol: $\tilde{\omega}$ is thus the twisted differential form associated with ω by a specific orientation. Let me denote any ordered set of vectors which represents the orientation of the object τ by $\{\tau\}$. If Ω is the unit volume form, then $\{\Omega\}$ is an orientation of the entire tangent space, and the map tilde is given by

$$\{\{\tilde{\alpha}\}, \{\alpha\}\} = \{\Omega\}, \tag{1}$$

as we shall see. I will use a tilde over the symbol to indicate twisted forms in general. The most convenient representation for twisted forms is to pick an orientation $\{\Omega\}$ and to use $\widetilde{dx}$, $\widetilde{dy}$, and $\widetilde{dz}$ as a basis.

INTEGRATION OF TWISTED FORMS

Ordinary differential forms can be integrated over regions having an inner orientation. The most natural application of this is to line integrals. For each little piece of the integrand one compares the orientation of the differential form with the orientation of the region to find the sign of its contribution to the integral. We have Stokes' Theorem

$$\int_\Gamma d\omega = \int_{\partial\Gamma} \omega, \tag{2}$$

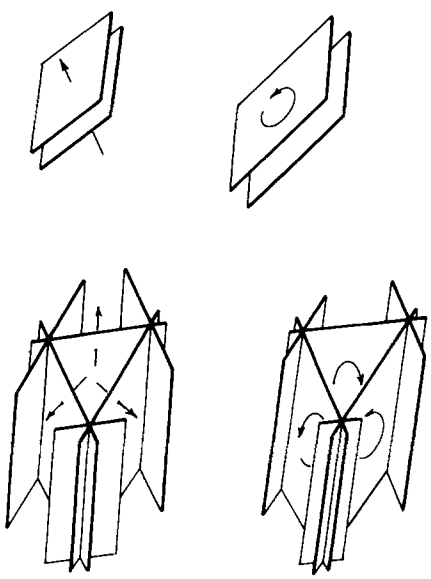


FIG. 3. A 1-form, a twisted 1-form, and their rule of addition, shown by triples which sum to zero.

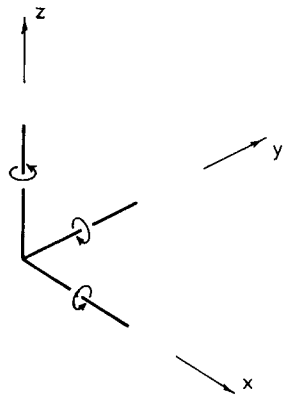


FIG. 4. A right-handed basis for twisted vectors.

where the orientation of the boundary is related to the orientation of Γ by

$$\{n, \{\partial\Gamma\}\} = \{\Gamma\}. \tag{3}$$

Here n is the outward pointing normal. Ordinary differential forms cannot be integrated over nonorientable regions.

Twisted forms, on the other hand, can be integrated over regions which have an outer orientation. The most natural application here is to volume integrals. Let us consider first integration over a region with the same dimension as the space itself. Twisted n-forms already have a sign and so they may be integrated directly, even over nonorientable regions. To integrate a twisted $(n-1)$-form, we compare the orientation of the form with a vector giving the outer orientation of the region, and from this find the sign of the integrand. In the next section we will define the exterior derivative of twisted forms so that we have what I think should be called the divergence theorem,

$$\int_{\tilde{\Gamma}} d\tilde{\omega} = \int_{\widetilde{\partial\Gamma}} \tilde{\omega}. \tag{4}$$

The tilde here indicates that the regions have an outer orientation. The boundary $\partial\Gamma$ is given an outer orientation using the outward-pointing normal. This ensures that the contribution of an internal boundary cancels, and that an integration region can be freely cut up into cells. We will soon define pullback so that the divergence theorem can be applied also to subspaces.

OPERATIONS ON TWISTED FORMS

All of the operations of the exterior calculus readily extend to twisted differential forms. The general rule is that the tilde factors through products, and that its square is unity. We want it to commute with the operations of exterior differentiation,

$$d\tilde{\omega} = \widetilde{d\omega}, \tag{5}$$

wedge product

$$\widetilde{\alpha\wedge\beta} = \tilde{\alpha}\wedge\beta = \alpha\wedge\tilde{\beta}, \tag{6}$$

and pullback.

If we look at the monomial

$$\tilde{\omega} = f(x^1)\,\widetilde{dx}^2\wedge dx^3\wedge\cdots\wedge dx^n \tag{7}$$

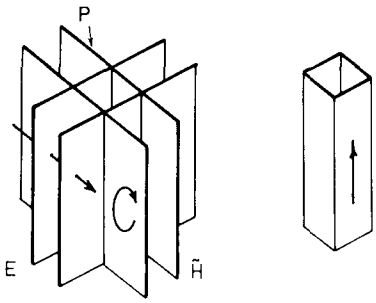


FIG. 5. The wedge product of a 1-form E and a twisted 1-form $\tilde{H}$ is a twisted 2-form. A regular 2-form would not behave properly under reflection in the plane P.

and use the commutivity of d and tilde we find

$$d\tilde{\omega} = \frac{\partial f}{\partial x^1} d\tilde{x}^1 \wedge dx^2 \wedge \cdots \wedge dx^n. \tag{8}$$

The divergence theorem will only be satisfied if we assign to $\tilde{\omega}$ the orientation $\{\tilde{x}^1\}$ (just compare with the corresponding Stokes' theorem). This forces the map tilde to have the form

$$\{\{\tilde{\omega}\},\{\omega\}\} = \{\Omega\}. \tag{9}$$

The orientation of the wedge product $\alpha \wedge \beta$ is given by

$$\{\alpha \wedge \beta\} = \{\{\alpha\},\{\beta\}\}, \tag{10}$$

where the vectors in $\{\alpha\}$ should be in the kernel of $\{\beta\}$ and vice versa. From Eq. (9) we find the relation

$$\{\{\beta\}, \widetilde{\{\alpha \wedge \beta\}}\} = \{\tilde{\alpha}\}. \tag{11}$$

The wedge product of a 1-form and a twisted 1-form is a twisted 2-form. See Fig. 5. The wedge product of two twisted 1-forms is an ordinary 2-form, as shown in Fig. 6. The contraction of a vector with a twisted 1-form gives, not a signed number, but an *oriented* number (screw-sense).

One operation not specified by the tilde-rule is pullback. While ordinary differential forms can be pulled back onto subspaces directly, twisted forms require an outer orientation of the subspace. If this outer orientation is given by $\{n\}$, then the orientation of the pullback $\psi^*(\tilde{\alpha})$ is given by

$$\{\{n\}, \{\psi^*(\tilde{\alpha})\}\} = \{\tilde{\alpha}\}. \tag{12}$$

If we give the subspace the orientation $\{\Omega'\}$ satisfying

$$\{\{n\}, \{\Omega'\}\} = \{\Omega\}, \tag{13}$$

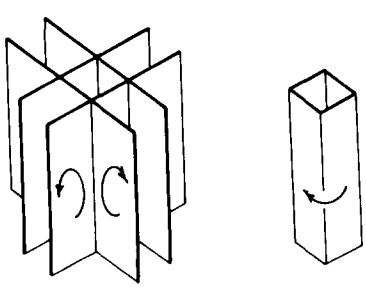


FIG. 6. The wedge product of two twisted 1-forms is an ordinary 2-form. A twisted 2-form would not behave properly under reflection in a horizontal plane.

then pullback commutes with tilde. The pullback behavior of twisted forms will fit naturally into the junction conditions of electromagnetism.

To use the divergence theorem on subspaces, we need an outer orientation for the subspace $\tilde{\Gamma}$, and an outward normal vector n for the boundary $\partial\tilde{\Gamma}$. The outer orientation of the boundary is then

$$\{\partial\tilde{\Gamma}\} = \{\{\tilde{\Gamma}\}, n\}. \tag{14}$$

A metric is usually introduced into the exterior calculus by the Hodge star operator. This operator involves both the metric *and* a choice of orientation. We can define an unoriented version of the Hodge star by mapping forms to twisted forms, and vice versa. This can be handled by writing the new operator as, $\tilde{*}$, and using the tilde-rule to simplify products. Despite its appearance, $\tilde{*}$ is independent of any choice of orientation.

ELECTROMAGNETISM

The manifestly parity invariant representation of electromagnetism comes from the work of van Dantzen[11] and Schouten,[3] although their work is not in modern notation and sometimes hard to follow. The representation in parity invariant form uses the geometric objects shown in Table I and in Fig. 7. The co-orientations are all taken with respect to 3-space. Time enters here just as a parameter. These "egg-crate" representations of 2-forms are the same as those given in Schouten[3] or Misner, Thorne, and Wheeler.[1] The development of electromagnetism here follows Frankel[9] except for the units.

Maxwell's equations for the evolution of the electric and magnetic fields read

$$\frac{\partial B}{\partial t} = -dE, \tag{15}$$

$$\frac{\partial \tilde{D}}{\partial t} = d\tilde{H} - 4\pi\tilde{J}, \tag{16}$$

with initial-value equations

$$dB = D, \tag{17}$$

$$d\tilde{D} = 4\pi\tilde{\rho}. \tag{18}$$

We are using unrationalized units with $c = 1$ (thus avoiding the e.s.u./e.m.u. distinction). The operator d is exterior differentiation in 3-space. The Lorentz force law is

$$F = q(E - v \cdot B). \tag{19}$$

The geometric objects were chosen as follows. The current is represented by a twisted 2-for $\tilde{J}$, an "egg-crate" en-

TABLE I. Geometric objects for electromagnetism.

E	1-form
$\tilde{D}$	twisted 2-form
B	2-form
$\tilde{H}$	twisted 1-form
$\tilde{J}$	twisted 2-form
$\tilde{\rho}$	twisted 3-form
ϕ	scalar
A	1-form

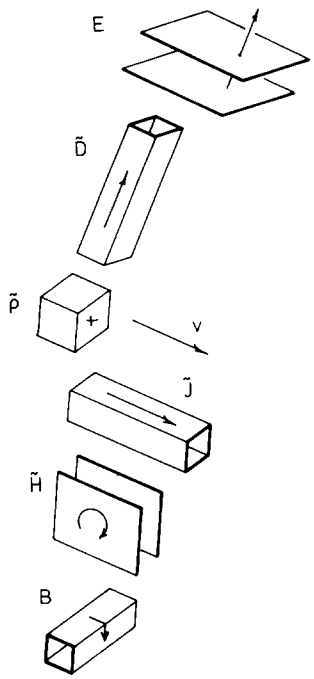


FIG. 7. The geometric objects representing electric and magnetic fields. The signs of the fields have been carefully chosen. The charge shown creates the $\tilde{D}$ field shown, and that $\tilde{D}$ leads to the E shown, etc.

closing unit current with an orientation given by the direction of current flow. Using an ordinary 2-form to represent current density would improperly describe the current flow by a screw sense. The charge density is represented by a twisted 3-form enclosing unit charge, with a sign for positive or negative charge. Again a twisted 3-form is used because an ordinary 3-form would describe charge with a screw sense. An ordinary 3-form would represent magnetic charge. Charge conservation follows by taking the exterior derivative of Eq. (16):

$$\frac{\partial \tilde{\rho}}{\partial t} = -d\tilde{J}, \tag{20}$$

and using Eq. (18). It guarantees that the initial-value equations are preserved by the dynamical equations.

With the above geometric structures for charges and currents, we see that $\tilde{D}$ and $\tilde{H}$ must also be twisted. The

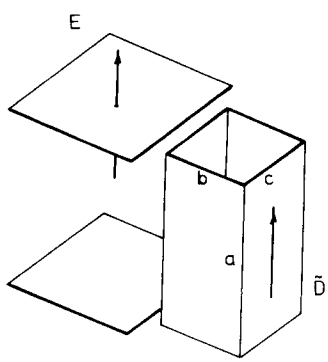


FIG. 8. The action of $\tilde{*}$ in three dimensions. The 1-form E and the twisted 2-form $\tilde{D}$ form a rectangular parallelopiped with sides satisfying $a = bc$.

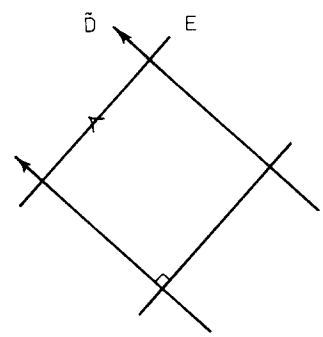


FIG. 9. The action of $\tilde{*}$ in two dimensions. The 1-form E and the twisted 1-form $\tilde{D}$ form a square.

parity properties are now nicely straightened out. The B field is a 2-form and does not change sign under inversion. The $\tilde{D}$ field is a twisted 2-form and does. Likewise, E is a 1-form and does change sign under inversion, while $\tilde{H}$ does not. Note that we give the B field the correct parity by making it a 2-form rather than a twisted 1-form.

No covariant differentiations are needed and thus far no metric has appeared. This is the usual advantage of using differential forms in electromagnetism. The metric must enter, and it does in relating E to $\tilde{D}$ and B to $\tilde{H}$. The usual formalism uses the Hodge star operator for this, but for our parity invariant formalism we will use $\tilde{*}$. The construction in vacuum is shown in Fig. 8. We have

$$\tilde{D} = \tilde{*}E, \tag{21}$$

$$\tilde{H} = \tilde{*}B. \tag{22}$$

In Fig. 9 we show the construction in two space dimensions. This is the familiar square construction. It requires only a conformal structure, not a full metric. Indeed, conformal transformations are only a symmetry of electromagnetism in two and four dimensions.

The junction conditions of electromagnetism are of two types. E and B are continuous across any surface. Their pullbacks onto any surface from either side must be equal. $\tilde{D}$ and $\tilde{H}$, however, can have discontinuities if there are surface

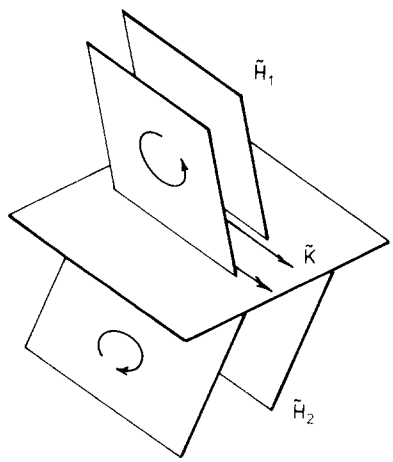


FIG. 10. A surface current $\tilde{K}$ and a suitably discontinuous $\tilde{H}$ field. The addition ignores the 4π factor for clarity.

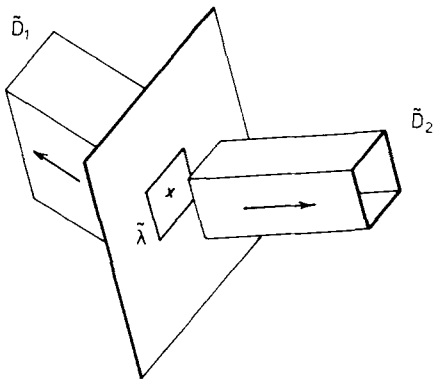


FIG. 11. A surface charge $\tilde{\lambda}$ and a suitably discontinuous $\tilde{D}$ field.

charges or currents. Surface charge can be represented by a twisted 2-form on the surface, with the orientation taken in the two dimensional surface. A surface current is represented by a twisted 1-form on the surface. We pullback the fields on the two sides of a surface, orienting the surface with a vector pointing from the surface *to* the side where the field is defined. The field on the other side of the surface uses the opposite orientation. The junction condition is that the *sum* of the pullbacks of $\tilde{H}$ from the two sides is 4π times the surface current, and the sum of the pullbacks of $\tilde{D}$ is 4π times the surface charge. This natural relation between surface current and the $\tilde{H}$ field is shown in Fig. 10; the relation between surface charge and the $\tilde{D}$ field is shown in Fig. 11. We orient the surface with a vector n pointing from the surface *to* the side where the field is defined, and use the opposite orientation to pullback the field on the other side of the surface.

Note how naturally these two types of junction conditions fit into our formalism. We cannot naturally give a sign to the discontinuity of either E or B, nor do we need to. For $\tilde{D}$ and $\tilde{H}$ the sign is forced on us by the behavior of twisted forms under pullback. It is amusing to note that magnetic charge does *not* fit naturally into this formalism. A surface current of magnetic charge has an outer orientation and cannot be described by any geometric object intrinsic to the surface. How is this difference between electric and magnetic charge reconciled with duality rotations? These are the transformations:

$$E \to (\cos\theta)E + (\sin\theta)*B, \tag{23}$$

$$B \to -(\sin\theta)*E + (\cos\theta)B. \tag{24}$$

Note that the duality rotation must involve $*$ and not $\tilde{*}$. It is *not* a parity invariant transformation.

[1] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
[2] H. Weyl, *Space, Time, Matter* (Dover, New York, 1922).
[3] J. A. Schouten, *Tensor Analysis for Physicists* (Oxford University, Oxford, England, 1951).
[4] J. L. Synge and A. Schild, *Tensor Calculus* (University of Toronto, Toronto, 1949).
[5] G. DeRham, *Varieties Differentiables* (Hermann, Paris, 1955).
[6] R. Sorkin, J. Phys. A **10**, 717 (1977).
[7] N. Steenrod, *The Topology of Fibre Bundles* (Princeton University, Princeton, New Jersey, 1951), p. 23.
[8] J. Eells, mimeographed lecture notes from the University of Amsterdam (1966) (unpublished).
[9] T. Frankel, *Gravitational Curvature* (Freeman, San Francisco, 1979).
[10] W. L. Burke, *Spacetime Geometry Cosmology* (University Science Books, Mill Valley, California, 1980).
[11] D. van Dantzen, Proc. K. Ned. Akad. Wet. **37**, 521, 526, 644, 825 (1934).

On the stability problem of a pair of adjoint operators

Per-Olov Löwdin
Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32611 and Department of Quantum Chemistry, Uppsala University, Box 518, S-751 20, Uppsala, Sweden

(Received 21 December 1981; accepted for publication 21 April 1982)

As an introduction, the eigenvalue problem for a linear operator T having a discrete point spectrum and a complete set of eigenfunctions is studied. The bivariational principle for T and its adjoint operator $T^\dagger$ is derived, and the biorthogonal properties of their eigenfunctions are discussed. The main part of the paper is then concerned with the problem whether these features can be extended also to a general pair of adjoint operators, T and $T^\dagger$, in which case the eigenvalue problem is replaced by the more general stability problem. The stability problem for a pair of adjoint operators—T and $T^\dagger$—is first formulated in terms of nonorthogonal projectors—O and $O^\dagger$—which decompose these operators and satisfy the commutation relations $TO = OT$ and $T^\dagger O^\dagger = O^\dagger T^\dagger$. In the case of a finite space, these skew-projectors may be explicitly expressed in product forms derived from the reduced Cayley–Hamilton equation for the operator T. It is shown that, if the stable subspaces defined by these projectors are properly classified by their Segre characteristics, one may explicitly derive the form of the projectors for the irreducible stable subspaces associated with the individual Jordan blocks of the so-called classical canonical forms of the matrix representations of T and $T^\dagger$. It is further shown that, in such a case, the biorthonormality property of the generalized eigenfunctions is still valid, and that a bivariational principle may be derived. The extension of these results to infinite spaces is finally briefly discussed.

PACS numbers: 03.65. – w

1. INTRODUCTION

A fundamental mathematical tool in quantum theory is represented by the linear operators T defined on a Hilbert space $\mathfrak{H} = \{f\}$ having the positive definite binary product $\langle f|g\rangle$. Such an operator T has a domain $D(T)$, and it has further an adjoint operator $T^\dagger$ with the domain $D(T^\dagger)$ defined through the relation

$$\langle f|Tg\rangle = \langle T^\dagger f|g\rangle. \tag{1.1}$$

Considering the physical applications, however, one is particularly interested in such operators F which have real expectation values $\langle F\rangle_{av} = \langle f|F|f\rangle : \langle f|f\rangle$ for any state vector f within $D(F)$. Using the well-known polarization identity,[1] it is easily shown that such operators are *self-adjoint*, $F^\dagger = F$, which means that $\langle f|Fg\rangle = \langle Ff|g\rangle$ and that $D(F^\dagger) = D(F)$. In the case of a purely discrete point spectrum, these operators have real eigenvalues and orthogonal eigenfunctions and, in many cases, the latter form a complete set which may be used as a basis for the space $\mathfrak{H} = \{f\}$ in formal studies as well as in practical applications. These properties may also be generalized to the case when the spectrum of the operator F is partly or fully continuous.

In addition, one also studies sometimes in physics *normal operators* Λ characterized by the relation $\Lambda\Lambda^\dagger = \Lambda^\dagger\Lambda$. They have complex eigenvalues and orthogonal eigenfunctions, and they may be considered as combinations $\Lambda = A + iB$ of two self-adjoint operators A and B having the property $AB = BA$. Of particular importance are, of course, the unitary operators U characterized by the relation $UU^\dagger = U^\dagger U = 1$ having their eigenvalues in the unit circle in the complex plane. The spectra of the normal operators may be either discrete point spectra or partly or fully continuous, and the main properties are still essentially the same.

From the point of view of physics, it seems hence sufficient to study the properties of the self-adjoint and the normal operators. However, for the mathematical treatment of many problems in the quantum theory of matter, one has during the last decades become interested also in the *general linear operators* T which are neither self-adjoint nor normal, in the partitioning technique[2] for solving the eigenvalue problem with the aid of a complex parameter or in the complex scaling method[3] for studying resonance phenomena in scattering problems.

The question is under what conditions one can generalize the highly useful properties of the self-adjoint and normal operators—particularly the orthogonality and expansion properties of the eigenfunctions—also to general linear operators T. The general treatment of this question is a comparatively difficult mathematical problem which may still have to wait for some time for its final solution. In this paper, we will only try to familiarize ourselves with certain aspects of the problem in some particularly simple cases which still may be of interest to physicists and quantum chemists.

2. OPERATORS WITH DISTINCT POINT SPECTRA AND COMPLETE SETS OF EIGENFUNCTIONS

Let us start by considering an operator T having a discrete point spectrum $\{\lambda_k\}$ consisting of distinct (nondegenerate) eigenvalues λ_k in the complex plane associated with the eigenfunctions C_k, so that $TC_k = C_k\lambda_k$. The eigenfunctions $\{C_k\}$ are assumed to be complete in the sense that they form a basis for the space $\mathfrak{H} = \{f\}$, so that one has an expan-

 0022-2488/83/010070-18$02.50

sion of the type

$$f=\sum_k C_k a_k \tag{2.1}$$

for every element f in $\mathfrak{H}$. In such a case, the operator T has an adjoint $T^\dagger$ defined through the relation (1.1), and it will turn out to be convenient to study the two operators simultaneously as a pair $(T,T^\dagger)$. The operator $T^\dagger$ is assumed to have the eigenvalues μ and the eigenfunctions D. Starting from the eigenvalue problems

$$TC_k=\lambda_k C_k, \quad T^\dagger D=\mu D, \tag{2.2}$$

one gets immediately

$$\begin{aligned}\mu^*\langle D|C_k\rangle &= \langle T^\dagger D|C_k\rangle = \langle D|TC_k\rangle \\ &= \lambda_k\langle D|C_k\rangle,\end{aligned} \tag{2.3}$$

i.e.,

$$(\mu^*-\lambda_k)\langle D|C_k\rangle=0, \tag{2.4}$$

which means that

$$\langle D|C_k\rangle=0 \quad \text{if } \mu\neq\lambda_k^*. \tag{2.5}$$

This is the so-called general *biorthogonality theorem*, and we will now show that—in the special case considered in this section—it replaces the orthogonality theorem for the self-adjoint and normal operators in the most useful way.

So far, we have not made any assumptions about the set $\{\mu\}$, i.e., about the spectrum of the operator $T^\dagger$. It may now be shown that, to each eigenfunction D to $T^\dagger$, there exists one and only one eigenfunction C_k to T, such that

$$\langle D|C_k\rangle\neq 0. \tag{2.6}$$

If all the eigenfunctions C_k would be orthogonal to D, one would have $\langle D|f\rangle=\Sigma_k\langle D|C_k\rangle a_k=0$ for all f, which is impossible since $D\neq 0$. Combining (2.4) and (2.6), one gets further $\lambda_k=\mu^*$, and since all the eigenvalues λ_k are distinct, the function C_k is unique except for a trivial constant. This means that one has a unique pairing between the eigenfunctions to T and $T^\dagger$, and we will introduce the notation D_k for the eigenfunction to $T^\dagger$ having the eigenvalue $\mu_k=\lambda_k^*$ and the property (2.6). Hence the spectrum $\{\mu_k\}$ is also distinct and discrete.

Since the spectra $\{\lambda_k\}$ and $\{\mu_k\}$ are both ennumerable, it may be convenient to arrange them in a specific order which is invariant under complex conjugation ($\mu_k=\lambda_k^*$), for instance after the properties of their absolute values or their real components, or both.

It is now clear that the eigenfunctions C_k must necessarily be linearly independent. Multiplying the relation

$$\sum_k C_k\alpha_k=0 \tag{2.7}$$

to the left by D_l, one gets $\langle D_l|C_l\rangle\alpha_l=0$, i.e., $\alpha_l=0$. This implies that the expansion (2.1) must be unique. Multiplying this relation to the left by D_l, one gets further

$$\langle D_l|f\rangle=\sum_k\langle D_l|C_k\rangle a_k=\langle D_l|C_l\rangle a_l, \tag{2.8}$$

i.e.,

$$a_l=\langle D_l|C_l\rangle^{-1}\langle D_l|f\rangle. \tag{2.9}$$

Substitution of (2.9) into (2.1) then gives the relation

$$f=\sum_k C_k a_k=\sum_k C_k\langle D_k|C_k\rangle^{-1}\langle D_k|f\rangle, \tag{2.10}$$

which corresponds to the following "resolution of the identity":

$$1=\sum_k|C_k\rangle\langle D_k|C_k\rangle^{-1}\langle D_k|=\sum_k O_k. \tag{2.11}$$

Here the operators

$$O_k=\frac{|C_k\rangle\langle D_k|}{\langle D_k|C_k\rangle} \tag{2.12}$$

satisfy the following fundamental relations:

$$O_k^2=O_k, \quad O_kO_l=0, \quad \mathrm{Tr}\,O_k=1, \tag{2.13}$$

$$TO_k=O_kT=\lambda_kO_k, \quad T=\sum_k\lambda_kO_k. \tag{2.14}$$

The operator O_k is hence an *eigenprojector* to T associated with the eigenvalue λ_k; the last relation (2.14) is the "spectral resolution" of the operator T.

For the adjoint operator $O_k^\dagger\neq O_k$, one has

$$O_k^\dagger=\frac{|D_k\rangle\langle C_k|}{\langle C_k|D_k\rangle} \tag{2.15}$$

and it satisfies the relations

$$(O_k^\dagger)^2=O_k^\dagger, \quad O_k^\dagger O_l^\dagger=0, \quad \mathrm{Tr}\,O_k^\dagger=1, \tag{2.16}$$

$$T^\dagger O_k^\dagger=O_k^\dagger T^\dagger=\lambda_k^*O_k^\dagger, \tag{2.17}$$

$$1=\sum_k O_k^\dagger, \quad T^\dagger=\sum_k\lambda_k^*O_k^\dagger, \tag{2.18}$$

which are the adjoint of the relations (2.13) and (2.14) and (2.11). Using the first relation (2.18), one obtains

$$f=1\cdot f=\sum_k O_k^\dagger f=\sum_k D_k\langle C_k|D_k\rangle^{-1}\langle C_k|f\rangle, \tag{2.19}$$

which also implies that the eigenelements $\{D_k\}$ form a complete set and may serve as a basis.

The eigenfunction D_k is determined except for a constant factor. Putting $D_k'=\langle C_k|D_k\rangle^{-1}D_k$, one obtains the normalization

$$\langle D_k'|C_k\rangle=1, \tag{2.20}$$

which will be assumed to be automatically fulfilled in the following.

Introducing the bold symbols[1] $\mathbf{C}=\{C_1,C_2,C_3,\cdots\}$ and $\mathbf{D}=\{D_1,D_2,D_3,\cdots\}$ one may now write the biorthonormality theorem (2.5) and the resolution of the identity (2.11):

$$\langle D_l|C_k\rangle=\delta_{lk}, \quad 1=\sum_k|C_k\rangle\langle D_k| \tag{2.21}$$

in the condensed form:

$$\langle\mathbf{D}|\mathbf{C}\rangle=\mathbf{1}, \quad 1=|\mathbf{C}\rangle\langle\mathbf{D}|. \tag{2.22}$$

Since the set $\mathbf{C}$ is complete, it is evident that the set $\mathbf{D}$ can be expressed in the form $\mathbf{D}=\mathbf{C}\boldsymbol{\alpha}$, which gives $\langle\mathbf{D}|\mathbf{C}\rangle=\langle\mathbf{C}\boldsymbol{\alpha}|\mathbf{C}\rangle=\boldsymbol{\alpha}^\dagger\langle\mathbf{C}|\mathbf{C}\rangle=\mathbf{1}$ and $\boldsymbol{\alpha}^\dagger=\langle\mathbf{C}|\mathbf{C}\rangle^{-1}=\boldsymbol{\alpha}$. Hence

$$\mathbf{D}=\mathbf{C}\langle\mathbf{C}|\mathbf{C}\rangle^{-1}=\mathbf{C}_r, \tag{2.23}$$

i.e., the set $\mathbf{D}$ is the *reciprocal basis* of the basis $\mathbf{C}$, and it is

completely determined by the set **C**. Substituting (2.23) into the second relation (2.22), one obtains

$$1 = |\mathbf{C}\rangle\,\langle\mathbf{C}|\mathbf{C}\rangle^{-1}\langle\mathbf{C}|, \tag{2.24}$$

which is the completeness relation for a nonorthonormal basis **C**. Letting $\boldsymbol{\lambda}$ be the diagonal matrix with the elements $\{\lambda_k\}$, one can finally write the eigenvalue relations (2.2) in the condensed form

$$T\mathbf{C} = \mathbf{C}\boldsymbol{\lambda}, \quad T^\dagger\mathbf{D} = \mathbf{D}\boldsymbol{\lambda}^*. \tag{2.25}$$

It should be observed that some of the theorems treated in this section may be generalized also to operators having partly or fully continuous spectra on the real axis or in the complex plane. Some of these questions will be treated in a forthcoming paper.

3. BIVARIATIONAL PRINCIPLE FOR A PAIR OF ADJOINT OPERATORS

The variational principle is a fundamental tool in evaluating approximate eigenvalues and eigenfunctions to self-adjoint and normal operators. It is remarkable that—to some extent—it may be generalized also to general linear operators and their adjoints. Starting from the eigenvalue relations

$$TC = C\lambda, \quad T^\dagger D = D\mu, \tag{3.1}$$

one gets directly

$$\lambda = \frac{\langle D|T|C\rangle}{\langle D|C\rangle} = \mathrm{Tr}\,T\Gamma, \tag{3.2}$$

$$\mu = \frac{\langle C|T^\dagger|D\rangle}{\langle C|D\rangle} = \mathrm{Tr}\,T^\dagger\Gamma^\dagger = \lambda^*, \tag{3.3}$$

where

$$\Gamma = \frac{|C\rangle\langle D|}{\langle D|C\rangle}, \quad \Gamma^\dagger = \frac{|D\rangle\langle C|}{\langle C|D\rangle} \tag{3.4}$$

are "transition operators" satisfying relations of the type

$$\Gamma^2 = \Gamma, \quad \mathrm{Tr}\,\Gamma = 1, \quad \Gamma \neq \Gamma^\dagger. \tag{3.5}$$

In connection with the exact expressions (3.2) and (3.3), it is now convenient to study the variational forms:

$$I_1 = \frac{\langle x_2|T|x_1\rangle}{\langle x_2|x_1\rangle}, \quad I_2 = \frac{\langle x_1|T^\dagger|x_2\rangle}{\langle x_1|x_2\rangle} = I_1^*, \tag{3.6}$$

where x_1 and x_2 are a pair of elements of $\mathfrak{H} = \{f\}$ having the property $\langle x_2|x_1\rangle \neq 0$. Assuming x_1 and x_2 are variations of the true eigenfunctions C and D, respectively, so that

$$x_1 = C + \delta C, \quad x_2 = D + \delta D, \tag{3.7}$$

one obtains

$$\begin{aligned}(T-\lambda\cdot 1)x_1 &= (T-\lambda\cdot 1)\delta C,\\ (T-\lambda\cdot 1)^\dagger x_2 &= (T-\lambda\cdot 1)^\dagger\delta D.\end{aligned} \tag{3.8}$$

Starting from (3.6), one gets

$$\begin{aligned}I_1 - \lambda &= \frac{\langle x_2|T-\lambda\cdot 1|x_1\rangle}{\langle x_2|x_1\rangle} = \frac{\langle x_2|T-\lambda\cdot 1|\delta C\rangle}{\langle x_2|x_1\rangle}\\ &= \frac{\langle (T-\lambda\cdot 1)^\dagger x_2|\delta C\rangle}{\langle x_2|x_1\rangle} = \frac{\langle (T-\lambda\cdot 1)^\dagger\delta D\,|\delta C\rangle}{\langle x_2|x_1\rangle},\end{aligned} \tag{3.9}$$

i.e.,

$$I_1 = \lambda + \frac{\langle \delta D\,|T-\lambda\cdot 1|\delta C\rangle}{\langle x_2|x_1\rangle}. \tag{3.10}$$

Since the first-order variation does not appear in this expression, one has

$$\delta I_1 = 0. \tag{3.11}$$

Similarly, one gets $\delta I_2 = 0$.

The reverse theorem is also true. If $\delta I_1 = 0$ for all variations δx_1 and δx_2 of a given pair x_1 and x_2, then

$$(T - I_1\cdot 1)x_1 = 0, \quad (T^\dagger - I_1^*\cdot 1)x_2 = 0, \tag{3.12}$$

i.e., x_1 and x_2 are eigenelements to T and $T^\dagger$, respectively, associated with the eigenvalues I_1 and $I_2 = I_1^*$. For the proof, one observes that $I_1 = A/B$, which gives

$$\delta I_1 = \frac{B\delta A - A\delta B}{B^2} = \frac{1}{B}(\delta A - I_1\delta B) = 0, \tag{3.13}$$

where

$$\begin{aligned}\delta A - I_1\delta B &= \langle \delta x_2|T - I_1\cdot 1|x_1\rangle\\ &\quad + \langle x_2|T - I_1\cdot 1|\delta x_1\rangle\\ &= \langle \delta x_2|T - I_1\cdot 1|x_1\rangle\\ &\quad + \langle \delta x_1|(T - I_1\cdot 1)^\dagger|x_2\rangle^* = 0\end{aligned} \tag{3.14}$$

for all variations δx_1 and δx_2, including also the cases when either $\delta x_1 = 0$ or $\delta x_2 = 0$. This gives

$$\langle \delta x_2|(T - I_1\cdot 1)x_1\rangle = 0, \quad \langle \delta x_1|(T - I_1\cdot 1)^\dagger x_2\rangle = 0 \tag{3.15}$$

for all δx_2 and δx_1, which implies that the relations (3.12) are true.

The variation principle (3.11) may now also be used to obtain *approximate* eigenvalues and eigenelements. If $\boldsymbol{\phi} = \{\phi_1,\phi_2,\ldots,\phi_m\}$ and $\boldsymbol{\psi} = \{\psi_1,\psi_2,\ldots,\psi_m\}$ are linearly independent sets, one may try expansions of the type

$$x_1 = \boldsymbol{\phi}\mathbf{c}, \quad x_2 = \boldsymbol{\psi}\mathbf{d} \tag{3.16}$$

and look for the "best approximations" to the true eigenelements C and D, respectively. Using the bivariational principle, one has

$$I_1 = \frac{\langle x_2|T|x_1\rangle}{\langle x_2|x_1\rangle} = \frac{\mathbf{d}^\dagger\langle\boldsymbol{\psi}|T|\boldsymbol{\phi}\rangle\mathbf{c}}{\mathbf{d}^\dagger\langle\boldsymbol{\psi}|\boldsymbol{\phi}\rangle\mathbf{c}} = \frac{A}{B} \tag{3.17}$$

and further

$$\begin{aligned}&\delta A - I_1\delta B\\ &\quad = \delta\mathbf{d}^\dagger\langle\boldsymbol{\psi}|T - I_1\cdot 1|\boldsymbol{\phi}\rangle\mathbf{c} + \mathbf{d}^\dagger\langle\boldsymbol{\psi}|T - I_1\cdot 1|\boldsymbol{\phi}\rangle\delta\mathbf{c} = 0\end{aligned} \tag{3.18}$$

for all variations $\delta\mathbf{c}$ and $\delta\mathbf{d}$. This gives

$$\langle\boldsymbol{\psi}|T - I_1\cdot 1|\boldsymbol{\phi}\rangle\mathbf{c} = \mathbf{0}, \tag{3.19}$$

$$\langle\boldsymbol{\phi}|T^\dagger - I_1^*\cdot 1|\boldsymbol{\psi}\rangle\mathbf{d} = \mathbf{0}. \tag{3.20}$$

These equations are generalizations of the standard secular equations in quantum mechanics, and the approximate eigenvalues I_1 are the roots to the polynomial equation:

$$P(z) \equiv |\langle\boldsymbol{\psi}|T - z\cdot 1|\boldsymbol{\phi}\rangle| = 0. \tag{3.21}$$

Properties of the approximate eigenfunctions. In order to study the solutions to Eqs. (3.19)–(3.21) in greater detail, we will introduce the notations

$$\mathscr{T} = \langle\boldsymbol{\psi}|T\boldsymbol{\phi}\rangle, \quad \mathscr{T}^\dagger = \langle T\boldsymbol{\phi}|\boldsymbol{\psi}\rangle = \langle\boldsymbol{\phi}|T^\dagger\boldsymbol{\psi}\rangle, \tag{3.22}$$

$$\Delta = \langle\psi|\phi\rangle, \quad \Delta^\dagger = \langle\phi|\psi\rangle. \tag{3.23}$$

Indicating approximate quantities by a bar, one can write $I_1 = \bar{\lambda}$. One may now write (3.19) and (3.20) in the form

$$\mathscr{T}\mathbf{c}_i = \Delta\mathbf{c}_i\bar{\lambda}_i, \quad \mathscr{T}^\dagger\mathbf{d}_i = \Delta^\dagger\mathbf{d}_i\bar{\lambda}_i^*, \tag{3.24}$$

which equation systems should be solved for $i = 1,2,...,m$, where $z = \bar{\lambda}_1,\bar{\lambda}_2,...,\bar{\lambda}_m$ are the roots to (3.21). Assuming that the matrix $\Delta = \langle\psi|\phi\rangle$ is *nonsingular*, and introducing the matrix

$$\mathbf{T} = \Delta^{-1}\mathscr{T}, \tag{3.25}$$

one can write the first relation (3.24) in the form $\mathbf{T}\mathbf{c}_i = \mathbf{c}_i\bar{\lambda}_i$. Forming the quadratic matrix $\gamma = \{\mathbf{c}_1,\mathbf{c}_2,...,\mathbf{c}_m\}$ of order $m\times m$, one has $T\gamma = \gamma\bar{\lambda}$, where $\bar{\lambda}$ is the diagonal matrix formed by the approximate eigenvalues $\bar{\lambda}_i$. This gives the relation

$$\gamma^{-1}\mathbf{T}\gamma = \bar{\lambda}, \tag{3.26}$$

which shows that the matrix $\mathbf{T}$ is brought to diagonal form by means of a similarity transformation γ. The approximate eigenfunctions are given by the relations $\bar{\mathbf{C}}_i = \phi\mathbf{c}_i$ for $i = 1,2,3,...,m$ and for the row vector $\bar{\mathbf{C}} = \{\bar{\mathbf{C}}_1,\bar{\mathbf{C}}_2,...,\bar{\mathbf{C}}_m\}$, one obtains

$$\bar{\mathbf{C}} = \phi\gamma. \tag{3.27}$$

For the approximate eigenfunctions to $T^\dagger$, one has similarly $\bar{\mathbf{D}} = \psi\mathbf{d}_i$ and, for the row vector $\mathbf{D} = \{\bar{\mathbf{D}}_1,\bar{\mathbf{D}}_2,...,\bar{\mathbf{D}}_m\}$, this gives

$$\mathbf{D} = \psi\delta, \tag{3.28}$$

where $\delta = \{\mathbf{d}_1,\mathbf{d}_2,...,\mathbf{d}_m\}$ is a quadratic matrix of order $m\times m$ formed by the eigenvectors $\mathbf{d}_i$.

In order to study the connection between γ and δ, we will take the adjoint of (3.26), which gives

$$\gamma^\dagger\mathbf{T}^\dagger(\gamma^\dagger)^{-1} = \bar{\lambda}^\dagger \tag{3.29}$$

or

$$\gamma^\dagger\mathscr{T}^\dagger(\Delta^\dagger)^{-1}(\gamma^\dagger)^{-1} = \bar{\lambda}^*, \tag{3.30}$$

i.e.,

$$\mathscr{T}^\dagger(\Delta^\dagger)^{-1}(\gamma^\dagger)^{-1} = (\gamma^\dagger)^{-1}\bar{\lambda}^*. \tag{3.31}$$

A comparison with the second relation (3.24) in the form $\mathscr{T}^\dagger\delta = \Delta^\dagger\delta\bar{\lambda}^*$ shows that one has the connection

$$\delta = (\Delta^\dagger)^{-1}(\gamma^\dagger)^{-1}, \quad \delta^\dagger = \gamma^{-1}\Delta^{-1}, \tag{3.32}$$

where one has also chosen a convenient normalization of the eigenvectors $\mathbf{d}_i$.

The exact solutions satisfy the relations (2.22) and (2.23), and we will now study the behavior of the approximate eigenfunctions. Using (3.27), (3.28), and (3.32), one gets directly

$$\langle\bar{\mathbf{D}}|\bar{\mathbf{C}}\rangle = \langle\psi\delta|\phi\gamma\rangle = \delta^\dagger\langle\psi|\phi\rangle\gamma = \gamma^{-1}\Delta^{-1}\Delta\gamma = \mathbf{1}_m, \tag{3.33}$$

which means that the basic orthonormality relation is fulfilled. Since we have used only truncated sets of order m, it is evident that the second relation (2.22) cannot be valid, since it expresses a resolution of the identity with respect to the entire Hilbert space $\mathfrak{H}$. Instead one has

$$\begin{aligned} Q = |\bar{\mathbf{C}}\rangle\langle\bar{\mathbf{D}}| &= |\phi\gamma\rangle\langle\psi\delta| = |\phi\rangle\gamma\delta^\dagger\langle\psi| \\ &= |\phi\rangle\gamma\gamma^{-1}\Delta^{-1}\langle\psi| = |\phi\rangle\langle\psi|\phi\rangle^{-1}\langle\psi|, \end{aligned} \tag{3.34}$$

where Q is an operator having the properties

$$Q^2 = Q, \quad \mathrm{Tr}\,Q = m, \quad Q^\dagger \neq Q. \tag{3.35}$$

Since $Q\phi = \phi$, it is clear that Q is the projector on the subspace M_ϕ spanned by the elements ϕ. Similarly

$$Q^\dagger = |\psi\rangle\langle\phi|\psi\rangle^{-1}\langle\phi| \tag{3.36}$$

is the projector on the subspace M_ψ spanned by the element ψ. This means that, instead of the second formula (2.22), one has the relations

$$Q = |\bar{\mathbf{C}}\rangle\langle\bar{\mathbf{D}}| = \sum_{i=1}^{m}|\bar{C}_i\rangle\langle\bar{D}_i|, \tag{3.37}$$

$$Q^\dagger = |\bar{\mathbf{D}}\rangle\langle\bar{\mathbf{C}}| = \sum_{i=1}^{m}|\bar{D}_i\rangle\langle\bar{C}_i|, \tag{3.38}$$

which may be described as "resolutions" of the projectors associated with the subspaces M_ϕ and M_ψ, respectively, in terms of one-dimensional projectors. However, unless the sets ϕ and ψ span the same space one has $Q \neq Q^\dagger$, which means that these projectors are *not* orthogonal projectors but of a more general character which will be further discussed in Sec. 5.

In order to study the analog of relation (2.23), we will form the reciprocal basis $\bar{\mathbf{C}}_r$ to the basis $\bar{\mathbf{C}}$ through the relation

$$\begin{aligned} \bar{\mathbf{C}}_r &= \bar{\mathbf{C}}\langle\mathbf{C}|\mathbf{C}\rangle^{-1} = \phi\gamma[\gamma^\dagger\langle\phi|\phi\rangle\gamma]^{-1} \\ &= \phi\gamma\cdot\gamma^{-1}\langle\phi|\phi\rangle^{-1}(\gamma^\dagger)^{-1} = \phi\langle\phi|\phi\rangle^{-1}(\gamma^\dagger)^{-1}. \end{aligned} \tag{3.39}$$

Taking the projection of $\bar{\mathbf{C}}_r$ on the subspace M_ψ, one obtains

$$\begin{aligned} Q^\dagger\bar{\mathbf{C}}_r &= |\psi\rangle\langle\phi|\psi\rangle^{-1}\langle\phi|\phi\rangle\langle\phi|\phi\rangle^{-1}(\gamma^\dagger)^{-1} \\ &= |\psi\rangle(\Delta^\dagger)^{-1}(\gamma^\dagger)^{-1} = |\psi\rangle\delta = \bar{\mathbf{D}}, \end{aligned} \tag{3.40}$$

i.e.,

$$\bar{\mathbf{D}} = Q^\dagger\bar{\mathbf{C}}_r, \tag{3.41}$$

which is the relation desired. In the case when $m\to\infty$ and the two sets ϕ and ψ become complete, (3.39) goes over into (2.23).

It should be observed that these results are independent of any linear transformations $\phi' = \phi\alpha$ and $\psi' = \psi\beta$ of the basic sets introduced. Starting from a fixed set ϕ, it may be convenient to introduce the transformed set $\psi_r = \psi\langle\phi|\psi\rangle^{-1}$, since this gives

$$\langle\psi_r|\phi\rangle = \mathbf{1}. \tag{3.42}$$

One can then describe the set ψ_r as the set in M_ψ which is biorthonormal to the set ϕ in M_ϕ. Using (3.34) and (3.36), one gets directly

$$Q = |\phi\rangle\langle\psi_r|, \quad Q^\dagger = |\psi_r\rangle\langle\phi|. \tag{3.43}$$

According to (3.25), one gets for the fundamental matrix $\mathbf{T}$

$$\mathbf{T} = \langle\psi|\phi\rangle^{-1}\langle\psi|T|\phi\rangle = \langle\psi_r|T|\phi\rangle, \tag{3.44}$$

whereas the approximate solutions are given by the formulas

$$\bar{\mathbf{C}} = \phi\gamma, \quad \bar{\mathbf{D}} = \psi_r(\gamma^\dagger)^{-1}. \tag{3.45}$$

In concluding this section, it should be observed that the variational quantity I_1, which gives the approximate eigenvalues $\bar{\lambda}_i$, is a complex number, and this means that the optimum value defined by $\delta I_1 = 0$ is not a simple "maxi-

mum" or "minimum" but of a much more complicated character, the nature of which is hidden in the "hessian" given by the second-order term in formula (3.10). In fact, little research has been done so far to investigate how the quantity I_1 approaches an eigenvalue λ_i as $m \to \infty$ and the basic sets become complete. In the case of self-adjoint operators bounded from below, one has the Hylleraas–Undheim separation theorem,[4] and it would be interesting to study what happens to the roots of the secular equation (31) when one more function is added to the sets ϕ and ψ, and the order is changed from m to $(m+1)$, etc. Even some computer studies may be helpful to get a hint how to approach this problem.

4. STABILITY PROBLEM FOR A PAIR OF ADJOINT OPERATORS

In the treatment of self-adjoint and normal operators, the eigenvalue problem (2.2) was a convenient starting point for the construction of sets of eigenfunctions which were complete and which hence could be used as a basis for a further study of the properties of the Hilbert space $\mathfrak{H}$. In Sec. 2, we have studied a particular family of linear operators T which by assumption could be treated in a similar way. It should be observed, however, that—in the study of a general linear operator T—the eigenvalue problem (2.2) is too narrowly formulated to serve as a basis for the theory, and that it has to be replaced by more general concepts. Some of these will be discussed in this section.

A subspace V of $\mathfrak{H}$ is said to be *stable* under the operator T, if for any element f' out of V also Tf' belongs to V. A stable subspace V is said to be *irreducible* with respect to T, if there is no proper subspace of V which is also stable under T; otherwise it is said to be *reducible*. The stability problem and the search for irreducible subspaces of T is apparently a generalization of the eigenvalue problem $TC = \lambda C$, which corresponds to the existence of one-dimensional stable subspaces. The self-adjoint and normal operators are characterized by the fact that all their irreducible subspaces are one-dimensional, whereas this is usually not true for linear operators T in general.

Let us assume that the stable subspace V is of finite order p and that it may be spanned by the linearly independent set $\mathbf{f} = \{f_1, f_2, \ldots, f_p\}$. The stability property implies that

$$Tf_l = \sum_{k=1}^{p} f_k T_{kl}, \tag{4.1}$$

where the coefficients form a matrix $\mathbf{T}_f = \{T_{kl}\}$ of order $p \times p$, which may be considered as the matrix representation of the operator T in the subspace V with respect to the set $\mathbf{f}$. In "fat symbols," one may write (4.1) in the condensed form $T\mathbf{f} = \mathbf{f}\mathbf{T}_f$ and, if there is no risk for misunderstanding, we will omit the index f on the matrix.

If the basis for V undergoes a linear transformation $\mathbf{f}' = \mathbf{f}\boldsymbol{\alpha}$, one has $T\mathbf{f}' = T\mathbf{f}\boldsymbol{\alpha} = \mathbf{f}\mathbf{T}\boldsymbol{\alpha} = \mathbf{f}'(\boldsymbol{\alpha}^{-1}\mathbf{T}\boldsymbol{\alpha}) = \mathbf{f}'\mathbf{T}'$, i.e.,

$$\mathbf{T}' = \boldsymbol{\alpha}^{-1}\mathbf{T}\boldsymbol{\alpha}, \tag{4.2}$$

which is referred to as a *similarity transformation*. If $\mathbf{A}$ and $\mathbf{B}$ are quadratic matrices, their determinants fulfill the multiplication rule $|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|$. Considering the *characteristic polynomial*

$$P(z) \equiv |\mathbf{T} - z \cdot \mathbf{1}|, \tag{4.3}$$

where z is a complex variable, one gets immediately that

$$\begin{aligned} P'(z) &\equiv |\mathbf{T}' - z \cdot \mathbf{1}| \equiv |\boldsymbol{\alpha}^{-1}(\mathbf{T} - z \cdot \mathbf{1})\boldsymbol{\alpha}| \\ &\equiv |\boldsymbol{\alpha}^{-1}| \cdot |\mathbf{T} - z \cdot \mathbf{1}| \cdot |\boldsymbol{\alpha}| \\ &\equiv |\mathbf{T} - z \cdot \mathbf{1}| \equiv P(z), \end{aligned} \tag{4.4}$$

which means that the coefficients of the characteristic polynomial are invariant under linear transformations; they may hence be considered as characteristics of the operator T with respect to the subspace V which is, of course, the reason for the name of $P(z)$. Other characteristic quantities are the roots $z = \lambda_1, \lambda_2, \ldots, \lambda_p$ of the equation $P(z) = 0$, which are called *eigenvalues* also in the general case, even if there are no eigenvalue relations in the ordinary sense. The multiplicity of a specific root $z = \lambda_k$ will be denoted by g_k and referred to as the "order of degeneracy" of this root. According to the factorial theorem, one has then

$$P(z) = \prod_k (\lambda_k - z)^{g_k}, \tag{4.5}$$

which relation gives the connection between the coefficients and the eigenvalues.

The question is now whether the subspace V is reducible or not; in the former case it may be decomposed into two subspaces V_1 and V_2 which are both stable under the operator T. In such a case it should be possible to find a basis $\mathbf{f}_1$ for the subspace V_1 of order p_1, so that the elements of $\mathbf{f}_1$ transform among themselves under the operator T, and similarly for the subspace V_2 of order p_2. In such a case, there exists also a similarity transformation (4.2), which changes the matrix $\mathbf{T}$ into two diagonal blocks:

$$\boldsymbol{\alpha}^{-1}\mathbf{T}\boldsymbol{\alpha} = \begin{pmatrix} \mathbf{T}_1 & 0 \\ 0 & \mathbf{T}_2 \end{pmatrix}, \tag{4.6}$$

where $\mathbf{T}_1$ and $\mathbf{T}_2$ are of order p_1 and p_2, respectively. It is now evident that, if one wants to decompose the space V into irreducible subspaces, one should try to find a similarity transformation $\boldsymbol{\gamma}$ which *block-diagonalizes* the matrix $\mathbf{T}$ as far as ever possible.

At this point we observe that one has the elementary theorem that, if all the eigenvalues λ_k are *distinct* or nondegenerate with $g_k = 1$, the matrix $\mathbf{T}$ may be completely diagonalized, and all the irreducible subspaces are hence one-dimensional. In such a case, one may apply the theory of the two previous sessions. This means also that all complications in the general case are related to the existence of *degenerate eigenvalues*. By considering the simple examples,

$$\mathbf{J}_2 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

$$\mathbf{J}_3 = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \tag{4.7}$$

one can easily convince onself there exist elementary matrices which cannot be diagonalized. The eigenvalues are $z = \lambda$ with $g = 2, 3, \cdots$, and—if they could be diagonalized—one would have $\boldsymbol{\gamma}^{-1}\mathbf{J}_n\boldsymbol{\gamma} = \lambda \cdot \mathbf{1}_n$ and $\mathbf{J}_n = \boldsymbol{\gamma} \cdot \lambda \cdot \mathbf{1}_n \boldsymbol{\gamma}^{-1} = \lambda \cdot \mathbf{1}_n$,

which is certainly a contradiction. Matrices of the type (4.7) are known as *Jordan blocks*, and we will later see that they play a fundamental role in the theory of degenerate eigenvalues.

It should be observed that the existence of the stable subspace V to the operator T does not given an immediate clue as to existence of a subspace $V^\dagger$ which is stable under the operator $T^\dagger$ and, for the moment, we will consider this as an independent problem. Assuming that the subspace $V^\dagger$ is spanned by the set $\mathbf{g} = \{ g_1, g_2, ..., g_q \}$, the stability condition takes the form

$$T^\dagger g_l = \sum_{k=1}^{q} g_k R_{kl}, \tag{4.8}$$

where $\mathbf{R} = \{R_{kl}\}$ may be interpreted as the matrix representation of the operator $T^\dagger$ in the space $V^\dagger$ with respect to the basis $\mathbf{g}$. One may then write the relation (4.8) in the condensed form

$$T^\dagger \mathbf{g} = \mathbf{gR}. \tag{4.9}$$

In order to proceed, we will now develop some more mathematical tools and auxiliary concepts.

Projectors associated with a pair of subspaces. Before proceeding with the stability problem, we will now study whether one can construct a pair of adjoint projectors O and $O^\dagger$ which are associated with two arbitrary subpaces M_f and M_g, respectively, of the same order p. As we will see, a necessary and sufficient condition for such a construction is that there is *no* element in M_f—except the zero element—which is orthogonal to all the elements of M_g.

Let us assume that M_f and M_g are spanned by the linearly independent sets $\mathbf{f} = \{ f_1, f_2, ..., f_p \}$ and $\mathbf{g} = \{ g_1, g_2, ..., g_p \}$. Since there is no element $f' = \mathbf{fa}$ in M_f—except the zero element—which satisfies the orthogonality condition $\langle \mathbf{g} | f' \rangle = 0$ the equation system

$$\langle \mathbf{g} | f' \rangle = \langle \mathbf{g} | \mathbf{f} \rangle \mathbf{a} = \mathbf{0} \tag{4.10}$$

should have only the trivial solution $\mathbf{a} = \mathbf{0}$, which means that $|\langle \mathbf{g} | \mathbf{f} \rangle| \neq 0$ and that $\langle \mathbf{g} | \mathbf{f} \rangle$ is a nonsingular matrix having an inverse.

The problem is to construct a pair of operators O and $O^\dagger$ satisfying the relations

$$O^2 = O, \quad \mathrm{Tr}\, O = p, \tag{4.11}$$

$$O\mathbf{f} = \mathbf{f}, \quad O^\dagger \mathbf{g} = \mathbf{g}. \tag{4.12}$$

Taking an arbitrary element x out of the Hilbert space $\mathfrak{H} = \{x\}$, one has the decompositions

$$x = \mathbf{fc} + r_1 = \mathbf{gd} + r_2, \tag{4.13}$$

where

$$Ox = \mathbf{fc}, \quad O^\dagger x = \mathbf{gd} \tag{4.14}$$

are the projections of x on M_f and M_g, respectively. For the remainders r_1 and r_2, one has $Or_1 = 0$ and $O^\dagger r_2 = 0$, i.e., they are obviously orthogonal to M_g and M_f, respectively:

$$\langle \mathbf{g} | r_1 \rangle = \langle O^\dagger \mathbf{g} | r_1 \rangle = \langle \mathbf{g} | O r_1 \rangle = 0,$$
$$\langle \mathbf{f} | r_2 \rangle = \langle O\mathbf{f} | r_2 \rangle = \langle \mathbf{f} | O^\dagger r_2 \rangle = 0. \tag{4.15}$$

This gives

$$\langle \mathbf{g} | x \rangle = \langle \mathbf{g} | \mathbf{fc} + r_1 \rangle = \langle \mathbf{g} | \mathbf{f} \rangle \mathbf{c},$$
$$\langle \mathbf{f} | x \rangle = \langle \mathbf{f} | \mathbf{gd} + r_2 \rangle = \langle \mathbf{f} | \mathbf{g} \rangle \mathbf{d}, \tag{4.16}$$

i.e.,

$$\mathbf{c} = \langle \mathbf{g} | \mathbf{f} \rangle^{-1} \langle \mathbf{g} | x \rangle, \tag{4.17}$$

$$\mathbf{d} = \langle \mathbf{f} | \mathbf{g} \rangle^{-1} \langle \mathbf{f} | x \rangle. \tag{4.18}$$

Hence one has for the projections:

$$Ox = \mathbf{fc} = \mathbf{f} \langle \mathbf{g} | \mathbf{f} \rangle^{-1} \langle \mathbf{g} | x \rangle, \tag{4.19}$$

$$O^\dagger x = \mathbf{gd} = \mathbf{g} \langle \mathbf{f} | \mathbf{g} \rangle^{-1} \langle \mathbf{f} | x \rangle, \tag{4.20}$$

i.e.,

$$O = |\mathbf{f}\rangle \langle \mathbf{g} | \mathbf{f} \rangle^{-1} \langle \mathbf{g}|, \tag{4.21}$$

$$O^\dagger = |\mathbf{g}\rangle \langle \mathbf{f} | \mathbf{g} \rangle^{-1} \langle \mathbf{f}|. \tag{4.22}$$

It is immediately checked that these projectors satisfy the relations (4.11) and (4.12), and that each one is the adjoint of the other. If the sets spanning M_f and M_g undergo linear transformations, $\mathbf{f}' = \mathbf{f}\alpha$ and $\mathbf{g}' = \mathbf{g}\beta$, the operators O and $O^\dagger$ stay invariant. Introducing the particular set $\mathbf{g}_r = \mathbf{g} \langle \mathbf{f} | \mathbf{g} \rangle^{-1}$, one gets the relations

$$\mathbf{g}_r = \mathbf{g} \langle \mathbf{f} | \mathbf{g} \rangle^{-1}, \quad \langle \mathbf{g}_r | \mathbf{f} \rangle = \langle \mathbf{f} | \mathbf{g}_r \rangle = \mathbf{1}, \tag{4.23}$$

i.e., the set $\mathbf{g}_r$ is the basis in M_g which is *biorthonormal* to the basis $\mathbf{f}$ in M_f. In such a case, one has the simplified relations

$$O = |\mathbf{f}\rangle \langle \mathbf{g}_r|, \quad O^\dagger = |\mathbf{g}_r\rangle \langle \mathbf{f}|. \tag{4.24}$$

We note that the previously studied operators O and $O^\dagger$, defined by (3.34) and (3.36) or (3.43), are projectors of this general type.

In the special case when the set $\mathbf{g}$ may be expanded in the set $\mathbf{f}$, and vice-versa, one has apparently $\mathbf{g}_r = \mathbf{f} \langle \mathbf{f} | \mathbf{f} \rangle^{-1}$; in such a case, the subspaces M_f and M_g are identical and $O = O^\dagger = |\mathbf{f}\rangle \langle \mathbf{f} | \mathbf{f} \rangle^{-1} \langle \mathbf{f}|$, which means that O has become an *orthogonal projector* of a more conventional type.

Stability problem formulated in terms of projectors. Let us consider a subspace V which is stable under the operator T, and let us assume that it is described by a projector O having V as its range. The results of the previous subsection have shown that there is an infinite family of projectors having the property, and it is hence important to have this nonuniqueness in mind. The only exception is the orthogonal projector which is self-adjoint and hence not general enough for our purposes.

For any element x of $\mathfrak{H} = \{x\}$, the projection $f' = Ox$ belongs to V, and the stability condition implies then also that $Tf' = TOx$ belongs to V, i.e., $OTf' = Tf'$ or $OTOx = TOx$ for all x. This gives the operator relation

$$TO = OTO \tag{4.25}$$

as an expression for the stability condition. Its implications will be studied in greater detail below. A projector O which satisfies (4.25) is said to *reduce* the operator T.

The operator $P = 1 - O$ fulfills the relations $P^2 = P$ and $PO = 0$, and it may be interpreted as the projector for the *complement* V_c to the subspace V defined by the projector O. Since $O \neq O^\dagger$, one has usually four different projectors O, $O^\dagger$, P, and $P^\dagger$ defining the subspaces V, $V^\dagger$, V_c, and $V_c^\dagger$,

respectively. One may now write the stability condition (4.25) for the subspace V in the special form

$$PTO = 0. \tag{4.26}$$

Taking the adjoint relation $O^\dagger T^\dagger P^\dagger = 0$, one realizes that the projector $P^\dagger$ reduces the adjoint operator $T^\dagger$, and that hence the subspace $V_c^\dagger$ is stable under $T^\dagger$.

Since $(P^\dagger)^\dagger O = PO = 0$, the subspace $V = \{f'\}$ and $V_c^\dagger = \{g''\}$ are automatically *orthogonal*:

$$\langle g''|f'\rangle = \langle P^\dagger g''|Of'\rangle = \langle g''|PO|f'\rangle = 0, \tag{4.27}$$

which is obviously an extension of the previously derived biorthogonality theorem (2.5).

In studying the subspace V and $V^\dagger$, we note that they are obviously of the *same order*, since $\mathrm{Tr}\, O = \mathrm{Tr}\, O^\dagger$. We observe further that there is no element of V (except the zero element) which can be orthogonal to all elements of $V^\dagger$, since it would then be orthogonal to $V^\dagger$ as well as $V_c^\dagger$, i.e., to the entire Hilbert space $\mathfrak{H}$, which is impossible. At this stage there is, of course, no reason for the subspace $V^\dagger$ to be stable under the operator $T^\dagger$.

If the subspace V is of finite order p, the same applies to the subspace $V^\dagger$. Spanning the subspace $V^\dagger$ by the linearly independent set $\mathbf{g} = \{g_1, g_2, \ldots, g_p\}$, we observe that, according to (4.8) and the reasoning above, the condition

$$|\langle \mathbf{g}|\mathbf{f}\rangle| \neq 0 \tag{4.28}$$

is automatically fulfilled. In such a case, one can construct the projector according to (4.21):

$$O = |\mathbf{f}\rangle\, \langle \mathbf{g}|\mathbf{f}\rangle^{-1} \langle \mathbf{g}|. \tag{4.29}$$

Multiplying (4.25) to the right by $\mathbf{f}$ and observing that $O\mathbf{f} = \mathbf{f}$, one obtains $T\mathbf{f} = OT\mathbf{f}$, i.e.,

$$T\mathbf{f} = \mathbf{f}\mathbf{T}, \quad \mathbf{T} = \langle \mathbf{g}|\mathbf{f}\rangle^{-1} \langle \mathbf{g}|T|\mathbf{f}\rangle, \tag{4.30}$$

which is analogous to (4.1) with an explicit expression for the matrix $\mathbf{T}$. The second relation may look somewhat unfamiliar, but it may be obtained from the first by multiplying to the left by $\langle \mathbf{g}|$ and solving for $\mathbf{T}$.

In the case when not only the space V but also the complimentary space V_c is stable under T, one has the relation

$$(1 - P)TP = OTP = 0, \tag{4.31}$$

in which case the projectors O and P are said to *decompose* the operator T. Taking the adjoint relation of (4.31), one obtains $P^\dagger T^\dagger O^\dagger = 0$ or

$$T^\dagger O^\dagger = O^\dagger T^\dagger O^\dagger, \tag{4.32}$$

which means that the subspace $V^\dagger$ defined by $O^\dagger$ is *stable* under the operator $T^\dagger$. Multiplying (4.32) to the right by $\mathbf{g}$ and observing that $O^\dagger \mathbf{g} = \mathbf{g}$, one obtains $T^\dagger \mathbf{g} = O^\dagger T^\dagger \mathbf{g}$ or

$$T^\dagger \mathbf{g} = \mathbf{g}\mathbf{R}, \quad \mathbf{R} = \langle \mathbf{f}|\mathbf{g}\rangle^{-1} \langle \mathbf{f}|T^\dagger \mathbf{g}\rangle. \tag{4.33}$$

Since $\langle \mathbf{f}|T^\dagger \mathbf{g}\rangle = \langle T\mathbf{f}|\mathbf{g}\rangle = \langle \mathbf{f}\mathbf{T}|\mathbf{g}\rangle = \mathbf{T}^\dagger \langle \mathbf{f}|\mathbf{g}\rangle$, one has

$$\mathbf{R} = \langle \mathbf{f}|\mathbf{g}\rangle^{-1} \mathbf{T}^\dagger \langle \mathbf{f}|\mathbf{g}\rangle, \tag{4.34}$$

which implies that $\mathbf{R}$ is a similarity transformation of the adjoint matrix $\mathbf{T}^\dagger$. Introducing the special set $\mathbf{g}_r = \mathbf{g}\langle \mathbf{f}|\mathbf{g}\rangle^{-1}$ characterized by the relations (4.23) and (4.24), one gets finally

$$T\mathbf{f} = \mathbf{f}\mathbf{T}, \quad T^\dagger \mathbf{g}_r = \mathbf{g}_r \mathbf{T}^\dagger, \tag{4.35}$$

where

$$\mathbf{T} = \langle \mathbf{g}_r|T|\mathbf{f}\rangle. \tag{4.36}$$

It is evident that it is by no means trivial to determine the projectors O and $P = 1 - O$, so that both V and V_c become stable under the operator T and that, in practice, it may be easier to determine O and $O^\dagger$ so that V and $V^\dagger$ become stable under T and $T^\dagger$, respectively. Fortunately, there is one more aspect to the problem. It is evident that the necessary and sufficient condition for the validity of the two relations $PTO = OTP = 0$ or $TO = OTO = OT$ is that

$$TO = OT. \tag{4.37}$$

In such a case, one has also $O^\dagger T^\dagger = T^\dagger O^\dagger$, i.e., if O decomposes T, then $O^\dagger$ decomposes $T^\dagger$. From these two commutation relations and the explicit expressions for O and $O^\dagger$ according to (4.29), one can again derive the relations (4.30) and (4.33).

Through the relation (4.37), the problem of finding the projectors which decompose the operators T and $T^\dagger$ may be reduced to the problem of finding the projectors which *commute* with T. If a projector O is the sum of two projectors, i.e., $O = O_1 + O_2$, which both commute with T, the projector O is said to be reducible—otherwise it is irreducible.

The problem of finding the irreducible subspaces of T and $T^\dagger$ is then essentially reduced to finding a resolution of the identity operator in terms of irreducible projectors, which commute with the operator T, analogous to (2.11). In the special case when there exist one-dimensional stable subspaces, one has according to (2.14) and (2.17) the relation

$$TO_k = O_k T = \lambda_k O_k, \tag{4.38}$$

$$T^\dagger O_k^\dagger = O_k^\dagger T^\dagger = \lambda_k^* O_k^\dagger, \tag{4.39}$$

i.e., the operators O_k and $O_k^\dagger$ are eigenprojectors to T and $T^\dagger$, respectively. In the following, we will see that, in the study of general linear operators T, the relation (4.37) for irreducible projectors O is going to replace the eigenvalue problem (2.2) as a basis for the theory. In the next section, we will try to approach the problem of the proper "resolution of the identity."

5. TREATMENT OF THE STABILITY PROBLEM BY USING PRODUCT PROJECTION OPERATORS IN THE FINITE CASE

For the sake of simplicity, we will start by considering a space $A = \{x\}$ of finite order n. If a linear operator T defined on this space has the eigenvalues $\lambda_1, \lambda_2, \lambda_3, \cdots$ with the degeneracies $g_1, g_2, g_3, \cdots$, one may write the characteristic polynomial in the form

$$F(z) \equiv (-1)^n \cdot P(z) \equiv |z\cdot\mathbf{1} - \mathbf{T}| \equiv \prod_k (z - \lambda_k)^{g_k}. \tag{5.1}$$

Of essential importance now is the expansion of the inverse of $F(z)$ in terms of partial fractions:

$$\frac{1}{F(z)} \equiv \sum_k \frac{q_k(z)}{(z - \lambda_k)^{g_k}}, \tag{5.2}$$

where $q_k(z)$ is a polynomial in z of a degree equal to or less

than $(g_k - 1)$, which is easily determined by standard algebraic methods. Multiplying (5.2) by $F(z)$, one gets the identity

$$1 \equiv \sum_k q_k(z) \frac{F(z)}{(z-\lambda_k)^{g_k}} \equiv \sum_k q_k(z) \prod_{l \neq k} (z-\lambda_l)^{g_l}$$
$$\equiv \sum_k O_k(z), \tag{5.3}$$

where

$$O_k(z) \equiv q_k(z) \prod_{l \neq k} (z-\lambda_l)^{g_l} \tag{5.4}$$

is a polynomial of a degree equal to or less than $(n-1)$, which may be expressed in the form

$$O_k(z) \equiv a_0^{(k)} + a_1^{(k)} z + a_2^{(k)} z^2 + \cdots + a_{n-1}^{(k)} z^{n-1}. \tag{5.5}$$

Because of the relation (5.3), the coefficients $a_r^{(k)}$ have the following simple properties:

$$\sum_k a_0^{(k)} = 1, \quad \sum_k a_r^{(k)} = 0, \quad \text{for } r = 1,2,\ldots,n-1. \tag{5.6}$$

Instead of the factor $(z - \lambda_k \cdot 1)$, we will now introduce the operator

$$N_k = T - \lambda_k \cdot I, \tag{5.7}$$

where we have replaced the complex variable z by the operator T and the number 1 by the identity operator I. Instead of (5.4), we will now consider the polynomial operator

$$O_k(T) \equiv g_k(T) \prod_{l \neq k} (T - \lambda_l \cdot I)^{g_l}$$
$$\equiv a_0^{(k)} I + a_1^{(k)} \cdot T + a_2^{(k)} \cdot T^2 + \cdots + a_{n-1}^{(k)} \cdot T^{n-1}. \tag{5.8}$$

Using the relations (5.6), one gets immediately

$$\sum_k O_k(T) \equiv I, \tag{5.9}$$

and we will now show that it represents a "resolution of the identity" of the type desired.

The operator T satisfies the standard Cayley–Hamilton theorem, which means that

$$F(T) \equiv \prod_k (T - \lambda_k \cdot I)^{g_k} = 0, \tag{5.10}$$

i.e., $F(T)$ is a zero operator. Combining (5.8) and (5.10), one gets directly

$$(T - \lambda_k \cdot I)^{g_k} O_k(T) = O_k(T)(T - \lambda_k \cdot I)^{g_k} = 0. \tag{5.11}$$

Since further the factor $(T - \lambda_k \cdot I)^{g_k}$ is contained in all the operators $O_l(T)$ for $l \neq k$, one gets also

$$O_l(T) O_k(T) = 0. \tag{5.12}$$

Multiplying the relation (5.9) to the left by $O_l(T)$, one has

$$O_l(T) = \sum_k O_l(T) O_k(T) = O_l(T) O_l(T), \tag{5.13}$$

i.e., the operator $O_l(T)$ is an idempotent. According to (5.8), the operator $O_k(T)$ is a *polynomial* in T, which implies that it commutes with T:

$$T O_k(T) = O_k(T) T. \tag{5.14}$$

It is now evident that the operators $O_1, O_2, O_3, \cdots$ form a set of mutually exclusive *projectors* which commute with T and which form a resolution of the identity. The "product projection operators" $O_1, O_2, O_3, \cdots$ defined by (5.8) represent hence a solution to the problem stated in the previous section in the finite case, and they define a sequence of subspaces $V_1, V_2, V_3, \cdots$ which are not only *stable* under the operator T but also decompose the space.

Each one of the subspaces $V_1, V_2, V_3, \cdots$ is characterized by an "eigenvalue" $\lambda_1, \lambda_2, \lambda_3, \cdots$, which is defined as a root to the characteristic equation $P(z) = 0$. In the case of a nondegenerate root, Eq. (5.11) corresponds to an ordinary eigenvalue relation of the type (4.38), but in the degenerate case things are considerably more complicated. It should be observed, however, that—in the case of a general linear operator T—the relation (5.11) replaces the simple eigenvalue problem which was characteristic for the self-adjoint and normal operators. This means also that, even if one has a resolution of the identity (5.9), there is no simple "spectral resolution" of the operator T in the general case.

Connection between the two types of projectors. The product projection operator $O_k(T)$ defined by (5.8) and its adjoint operator look very different from the projectors (4.21) and (4.22) previously considered, and it may hence be interesting to study the connection between them.

For this purpose, we need some elementary theorems about projectors $P \neq 1$ in general satisfying the relation $P^2 = P$. If $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ is a basis for the entire space $A = \{x\}$, then the subspace $A_1 = PA$ is spanned by the set $\mathbf{X}' = P\mathbf{X} = \{PX_1, PX_2, \ldots, PX_n\}$, which is certainly linearly dependent. However, it may be replaced by a linearly independent set $\mathbf{X}_1$, if one goes through all elements $X_k' = PX_k$ in order, and leaves out all elements X_k' for which either $X_k' = PX_k = 0$ or X_k' is a linear combination of the preceding elements $X_1', X_2', \ldots, X_{k-1}'$. Similarly the subspace $A_2 = (1-P)A$ is spanned by the set $\mathbf{X}'' = (1-P)\mathbf{X}$, which may be replaced by the linearly independent set $\mathbf{X}_2$.

For any element $x = \mathbf{Xa}$, one has then the resolution

$$x = Px + (1-P)x = P\mathbf{Xa} + (1-P)\mathbf{Xa}$$
$$= \mathbf{X}_1\mathbf{a}_1 + \mathbf{X}_2\mathbf{a}_2, \tag{5.15}$$

where the linear dependencies have been removed. This that the combination $(\mathbf{X}_1, \mathbf{X}_2)$ forms another basis for the entire space $A = \{x\}$. Since further

$$P(\mathbf{X}_1, \mathbf{X}_2) = (\mathbf{X}_1, \mathbf{0}), \tag{5.16}$$

it is evident that the operator P with respect to this basis has a matrix of the form

$$\mathbf{P} = \begin{pmatrix} \mathbf{1}_g & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{5.17}$$

where $\mathbf{1}_g$ is a unit matrix of the same order g as the subspace $A_1 = PA$. Hence every projector P may be *diagonalized* with the eigenvalues 0 or 1, and one has

$$\mathrm{Tr}\, P = g, \tag{5.18}$$

where the integer g gives the order of the range of P. In the special case when $g = 0$, one has also $P = 0$.

Let us now start by considering the product projection operator $O_k(T)$ and its adjoint operator which have the

ranges V_k and $V_k^\dagger$, respectively. In the following, we will omit the index k if there is no risk of misunderstanding. The subspaces V and $V^\dagger$ are spanned by the linearly dependent sets $O\mathbf{X}$ and $O^\dagger\mathbf{X}$, which are then replaced by the linearly independent sets $\mathbf{f}$ and $\mathbf{g}$, respectively, by using the procedure described above. One has

$$O\mathbf{f}=\mathbf{f}, \quad O^\dagger\mathbf{g}=\mathbf{g}. \tag{5.19}$$

Since further

$$\mathrm{Tr}\, O=\mathrm{Tr}\, O^\dagger=g, \tag{5.20}$$

the subspaces V and $V^\dagger$ are of the same order. According to (4.21) and (4.22), one may now introduce the projectors:

$$Q=|\mathbf{f}\rangle\langle\mathbf{g}|\mathbf{f}\rangle^{-1}\langle\mathbf{g}|, \tag{5.21}$$

$$Q^\dagger=|\mathbf{g}\rangle\langle\mathbf{f}|\mathbf{g}\rangle^{-1}\langle\mathbf{f}|, \tag{5.22}$$

where $\mathrm{Tr}\, Q=\mathrm{Tr}\, Q^\dagger=g$. According to (5.19), one has further $O|\mathbf{f}\rangle=|\mathbf{f}\rangle$ and $\langle\mathbf{g}|O=\langle O^\dagger\mathbf{g}|=\langle\mathbf{g}|$, which gives

$$OQ=QO=Q. \tag{5.23}$$

For the difference $P=O-Q$, one gets hence that $P^2=O^2+Q^2-OQ-QO=O-Q=P$, and that $\mathrm{Tr}\, P=\mathrm{Tr}\, O-\mathrm{Tr}\, Q=0$. Hence $P=0$, which implies

$$O=Q. \tag{5.24}$$

Hence the product projector operator $O(T)$ and the projector Q given by (5.21) are different expressions for one and the same projector.

In the construction above, the sets $\mathbf{f}$ and $\mathbf{g}$ are obtained from the basis $\mathbf{X}$ by means of the projectors O and $O^\dagger$, respectively, independently of each other. Instead of the original set $\mathbf{g}$, one may find it convenient to introduce the reciprocal set $\mathbf{g}_r=\mathbf{g}\langle\mathbf{f}|\mathbf{g}\rangle^{-1}$ having the property $\langle\mathbf{g}_r|\mathbf{f}\rangle=\mathbf{1}$. According to (4.35) and (4.36), one then has the relations

$$T\mathbf{f}=\mathbf{f}\mathbf{T}, \quad T^\dagger\mathbf{g}_r=\mathbf{g}_r\mathbf{T}^\dagger, \tag{5.25}$$

where

$$\mathbf{T}=\langle\mathbf{g}_r|T|\mathbf{f}\rangle, \tag{5.26}$$

as expressions for the stability properties of the sets $\mathbf{f}$ and $\mathbf{g}_r$ under the operators T and $T^\dagger$, respectively.

In certain connections, it may be convenient to use a slightly different approach, in which the linearly independent set $\mathbf{f}$ is first established and a new set $\bar{\mathbf{g}}$ is then introduced by the formula

$$\bar{\mathbf{g}}=O^\dagger\mathbf{f}. \tag{5.27}$$

Since $\langle\bar{\mathbf{g}}|\mathbf{f}\rangle=\langle O^\dagger\mathbf{f}|\mathbf{f}\rangle=\langle\mathbf{f}|O|\mathbf{f}\rangle=\langle\mathbf{f}|\mathbf{f}\rangle$, the matrix $\langle\bar{\mathbf{g}}|\mathbf{f}\rangle$ is also nonsingular, and the elements in $\bar{\mathbf{g}}$ are linearly independent. For the projector Q, one obtains

$$Q=|\mathbf{f}\rangle\langle\bar{\mathbf{g}}|\mathbf{f}\rangle^{-1}\langle\bar{\mathbf{g}}|=|\mathbf{f}\rangle\langle\mathbf{f}|\mathbf{f}\rangle^{-1}\langle\mathbf{f}|O=Q_fO, \tag{5.28}$$

where O_f is the self-adjoint (orthogonal) projector on the space spanned by the set $\mathbf{f}$. For the reciprocal set $\bar{\mathbf{g}}_r=\bar{\mathbf{g}}\langle\mathbf{f}|\bar{\mathbf{g}}\rangle^{-1}$, one obtains finally

$$\bar{\mathbf{g}}_r=O^\dagger\mathbf{f}\langle\mathbf{f}|\bar{\mathbf{g}}\rangle^{-1}=O^\dagger\mathbf{f}\langle\mathbf{f}|\mathbf{f}\rangle^{-1}. \tag{5.29}$$

Nilpotent operators. It is interesting to observe that, even if the relation (5.11) is to be considered as a generalization of the ordinary eigenvalue problem for $g_k>1$, it leads to considerations of a rather different type. In treating a specific stable subspace V_k, we will again in the following temporarily omit the index k, if there is no risk for misunderstanding. Using the notation (5.7), one may now write (5.11) in the form

$$N^g f=0, \tag{5.30}$$

where $N=T-\lambda I$, for all elements f of the subspace V. If there exists at least one element f_s in V, for which

$$N^{g-1}f_s\neq 0, \tag{5.31}$$

one says that the operator N is *nilpotent* of order g within the subspace V. In such a case, it is convenient to introduce a sequence of elements $\mathbf{f}=\{f_1,f_2,f_3,\ldots,f_g\}$ through the recursion formula $f_{r-1}=Nf_r$, i.e.,

$$f_1=Nf_2, \quad f_2=Nf_3,\ldots,f_{g-1}=Nf_g, \quad f_g=f_s, \tag{5.32}$$

which implies that $f_r=N^{g-r}f_s$. We note particularly that, according to (5.31), one has $f_1=N^{g-1}f_s\neq 0$, whereas $Nf_1=N^2f_2=\cdots=N^gf_g=0$. The elements in the sequence $\mathbf{f}$ are certainly linearly independent for, if one assumes the existence of a linear relation

$$\mathbf{f}\alpha=f_1\alpha_1+f_2\alpha_2+\cdots f_g\alpha_g=0, \tag{5.33}$$

and multiplies it successively to the left by N^{g-1}, $N^{g-2},\ldots,N$, one gets a sequence of equations from which one may conclude that

$$\alpha_g=\alpha_{g-1}=\cdots=\alpha_1=0, \tag{5.34}$$

which proves the statement.

Choosing the linearly independent set $\mathbf{f}$ as a *basis* for the subspace V, one has $N\mathbf{f}=N\{f_1,f_2,\ldots,f_g\}=\{0,f_1,f_2,\ldots,f_{g-1}\}$. The corresponding matrix $\mathbf{N}$ hence consists of a sequence of 1's in the diagonal one step above the main diagonal, whereas all other elements are vanishing; we note that such matrices are common in physics as representations of "ladder operators", and that N is a typical step-up operator. One has further $N^2\mathbf{f}=\{0,0,f_1,f_2,\ldots,f_{g-2}\}$, etc., which implies that the sequence of subspaces V, $V'=NV$, $V''=NV'$,$\cdots$ have the orders g, $g-1$, $g-2$,...,1; the order of the subspace hence decreases by one unit every time the operator N is applied. This is an important result which we will use in the following subsection. Going back to the operator $T=N+\lambda I$, one has $T\mathbf{f}=\lambda\mathbf{f}+\{0,f_1,f_2,\ldots,f_{g-1}\}$, i.e.,

$$\begin{aligned}
Tf_1&=\lambda f_1,\\
Tf_2&=\lambda f_2+f_1,\\
Tf_3&=\lambda f_3+f_2,\\
&\cdots\\
Tf_g&=\lambda f_g+f_{g-1},
\end{aligned} \tag{5.35}$$

which gives the matrix representation

$$\mathbf{T}=\begin{pmatrix}
\lambda & 1 & 0 & 0 & \cdots & 0 & 0\\
0 & \lambda & 1 & 0 & \cdots & 0 & 0\\
\cdots & & & & & \cdots & \\
0 & 0 & \cdots & & & \lambda & 1\\
0 & 0 & \cdots & & & 0 & \lambda
\end{pmatrix}, \tag{5.36}$$

for the operator T in the subspace V with respect to the basis **f**. It should be observed that, even if the first relation in (5.35) is an ordinary eigenvalue problem for the eigenelement f_1, the same function f_1 occurs also in the line below—hence the subspace $\{f_1\}$ is stable under T, but it reduces without decomposing the operator.

In elementary matrix theory, one actually starts from the fact that—even in the case of a degenerate eigenvalue λ—the eigenvalue problem $TC = \lambda C$ has at least one solution C, which then is chosen as the first element of a basis $(C, Y_2, Y_3, ..., Y_m)$. Considering the eigenvalue problem in the space spanned by the elements $(Y_2, Y_3, ..., Y_m)$ and repeating the reasoning, one finds another eigenelement C' associated with another eigenvalue λ', etc. By repeating the process one shows that, by means of a proper choice of basis—i.e., by means of a similarity transformation—every matrix **T** may be *triangularized* to a form in which all the elements *below* the main diagonal are identically vanishing, whereas the elements on the main diagonal may be identified with the eigenvalues as defined by the secular equation $P(z) = 0$.

Using this theorem, one can now easily show that the matrix (5.36) representing a Jordan block of order g in accordance with (4.5) cannot be further block-diagonalized. Let us assume temporarily that it may be transformed into two diagonal blocks of order p_1 and p_2, respectively, with $p_1 \geqslant p_2$, which are subsequently triangularized. Studying the powers of the matrix $\mathbf{N} = \mathbf{T} - \lambda \cdot \mathbf{1}$, one obtains directly that $\mathbf{N}^{p_1} = 0$, which means that the operator N is nilpotent of order p_1 or less within the entire subspace V, which contradicts the relation (5.31). Hence the Jordan block (5.36) cannot be further block-diagonalized, and the stable subspace V is *irreducible*.

Segré characteristics. If the relation (5.31) is not fulfilled for any element f_s in V, things are going to be more complicated, and the space V is going to turn out to be reducible. To every element f in the stable subspace $V = \{f\}$ defined by the projector $O(T)$, one may now assign a specific exponent $m(f)$ or *index* such that:

$$N^{m(f)} f = 0, \quad N^{m(f)-1} f \neq 0, \tag{5.37}$$

where $m(f) \leqslant g$. Starting out from an arbitrary element f and using the construction of the previous subsection, one can now construct a stable and irreducible subspace $V(f)$ or order $m(f)$ which is associated with the element f. Such a subspace $V(f)$ reduces the operator T, and the question now is how the element f should be chosen so that the subspace $V(f)$ also *decomposes* the operator T.

The *minimal index m* is the smallest number having the property that $N^m f = 0$ for all elements f of the stable subspace V defined by the projector $O(T)$ and—as we will see in the following—it plays a fundamental role in the theory as well as in the physical applications.

It would be very nice if one could find a sequence of elements $f_{\mathrm{I}}, f_{\mathrm{II}}, f_{\mathrm{III}}, \cdots$ in the subspace V so that the irreducible subspaces $V(f_{\mathrm{I}})$, $V(f_{\mathrm{II}})$, $V(f_{\mathrm{III}})$, $\cdots$ decompose V in an exhaustive way. The orders $g_{\mathrm{I}}, g_{\mathrm{II}}, g_{\mathrm{III}}, \cdots$ of these subspaces are known as the *Segré characteristics* of the subspace V associated with the degenerate eigenvalue λ, and we will now study the necessary and sufficient conditions for such a decomposition.

For this purpose it is convenient to study the sequence of spaces

$$V, V' = NV, \quad V'' = NV' = N^2 V, \cdots \tag{5.38}$$

having the orders $g, g', g'', \cdots$. Since the order of an irreducible subspace $V(f)$ is diminished by one unit every time the operator N is applied, it is now easy to find the connection between the Segré characteristics $g_{\mathrm{I}}, g_{\mathrm{II}}, g_{\mathrm{III}}, \cdots$ and the sequence $g, g', g'', \cdots$. This connection is most easily demonstrated by a couple of numerical examples.

Let us start by considering a stable subspace V of order 5 having the Segré characteristics (3,2). This gives, for their contributions to the orders of the subspaces V, V', V'',

Contributions to order of:	V,	V',	V''
$m = 3$:	3,	2,	1
$m = 2$:	2,	1	
	5,	3,	1

, (5.39)

i.e., one gets the sequence (5,3,1). Conversely, if one starts from the following sequence $g, g', g'', \ldots$:

$$(14,8,4,2,1,0) \tag{5.40}$$

for the orders of the spaces (5.38), one can now easily derive a *necessary* condition for the Segré characteristics—provided that the corresponding subspaces really exist and add up to V. There is a total of five nonvanishing numbers in the sequence (5.40) and, since the space $V^{(5)} = N^5 V$ has the order zero, the minimal index is $m = 5$. There should hence exist at least one irreducible subspace of order $m = 5$ and, since $g^{(4)} = 1$, there is apparently exactly *one* irreducible subspace of this order, which contributes the sequence (5,4,3,2,1) to the sequence (5.40). This gives the difference

Contributions to:		V,	V',	V'',	V''',	V''''
		14,	8,	4,	2,	1
$m = 5$	−	5,	4,	3,	2,	1
		9,	4,	1,	0,	0

. (5.41)

Here (9,4,1) is a new sequence of numbers corresponding to $m_1 = 3$ and the contributions (3,2,1). Subtracting these contributions, one gets

Contributions to:		V,	V',	V''
		9,	4,	1
$m_1 = 3$	−	3,	2,	1
		6,	2,	0

, (5.42)

which result indicates that there must be exactly two subspaces having $m_2 = 2$, each one with the contributions (2,1). After subtracting (4,2) one is left with a single number 2, which corresponds to two subspaces having $m_3 = 1$. Remembering that the index of each sequence equals the number of nonvanishing figures and that one should start each subtraction procedure from the right, one can now write this decomposition directly in the following way:

Contributions to:	V,	V',	V'',	V''',	V''''
$m\backslash g$	14,	8,	4,	2,	1
5	5,	4,	3,	2,	1
3	3,	2,	1		
2	2,	1			
2	2,	1			
1	1				
1	1				

(5.43)

which result shows that the only possible Segré characteristics are represented by the sequence (5,3,2,2,1,1). One can get the same results by considering the second-order differences:

g:	14	8	4	2	1	0
Δg:	6	4	2	1	1	
$\Delta^2 g$:	2	2	1	0	1'	
$m=$	1	2	3	4	5	

(5.44)

where the last line indicates that there are two subspaces of order $m = 1$, two subspaces of order $m = 2$, one subspace of order $m = 3$, and one subspace of order $m = 5$.

It is easily shown that this theorem about the Segré characteristics is generally true. Letting $s^{(p)}$ denote the number of irreducible subspaces of order p, one should prove that

$$s^{(p)} = \Delta^2 g^{(p-1)} = g^{(p-1)} - 2g^{(p)} + g^{(p+1)}. \tag{5.45}$$

Observing that $g^{(m)} = g^{(m+1)} = \cdots = 0$, one has

$$\begin{aligned} s^{(m)} &= g^{(m-1)} = \Delta^2 g^{(m-1)}, \\ s^{(m-1)} &= g^{(m-2)} - 2g^{(m-1)} = \Delta^2 g^{(m-2)}, \\ &\cdots \\ &\cdots \\ &\cdots \end{aligned} \tag{5.46}$$

The remaining part of the proof is provided by induction. According to the general construction of Segré characteristics, as examplified in the table (5.43), one has

$$\begin{aligned} s^{(k-1)} &= g^{(k-2)} - 2s^{(k)} - 3s^{(k+1)} \\ &\quad - 4s^{(k+2)} - \cdots - (m+2-k)s^{(m)} \\ &= g^{(k-2)} - 2\Delta^2 g^{(k-1)} - 3\Delta^2 g^{(k)} - 4\Delta^2 g^{(k+1)} - \cdots \\ &= \Delta^2 g^{(k-2)}. \end{aligned} \tag{5.47}$$

Here the last step is achieved by using the recursion formulas for the second-order differences. If E is the step operator, one has $1 = (1-E)^2(1-E)^{-2} = \Delta^2(1 + 2E + 3E^2 + \cdots)$, or the identity

$$\begin{aligned} g^{(k-2)} &= \Delta^2 g^{(k-2)} + 2\Delta^2 g^{(k-1)} \\ &\quad + 3\Delta^2 g^{(k)} + 4\Delta^2 g^{(k+1)} + \cdots, \end{aligned} \tag{5.48}$$

which is the formula needed. Hence the theorem (5.45) is proven.

Before going into the problem of decomposing the stable subspace V into irreducible subspaces corresponding to Jordan blocks, it may be convenient to go into a few more mathematical details as to the properties of the operators $N, N^2, N^3, \ldots, N^{m-1}$. For this purpose, we will consider a general linear operator M defined on a subspace V of order g, which is stable under M. If the subspace V is spanned by the linearly independent set $\mathbf{f}$, one has $M\mathbf{f} = \mathbf{fM}$, where $\mathbf{M} = \{\mathbf{M}_1, \ldots, \mathbf{M}_g\}$ is a quadratic matrix which consists of the column vectors $\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_g$. Considering the transformed set $\mathbf{f}' = M\mathbf{f} = \mathbf{fM}$, one has $f'_k = \mathbf{fM}_k$, and the number of linearly independent elements f'_k in the set $\mathbf{f}'$ hence equals the number of linearly independent column vectors $\mathbf{M}_k$ in the matrix $\mathbf{M}$. This number is also given by the *rank* r of the matrix $\mathbf{M}$.

In elementary matrix theory, one says that a singular matrix $\mathbf{M}$ of order g having a vanishing determinant, $|\mathbf{M}| = 0$, is of rank r if at least one minor to $\mathbf{M}$ of order r is different from zero, whereas all minors of order $(r+1)$ are vanishing. In such a case, one has the fundamental theorem that the equation system

$$\mathbf{Ma}_i = \mathbf{0} \tag{5.49}$$

has exactly $(g-r)$ linearly independent solutions $\mathbf{a}_i$ for $i = 1,2,\ldots,g-r$, which form a rectangular matrix $\mathbf{a} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{g-r}\}$ of order $g \times (g-r)$. Putting $\bar{f}_i = \mathbf{fa}_i$, one may construct a sequence $\bar{\mathbf{f}} = \{\bar{f}_i\} = \bar{\mathbf{f}}\mathbf{a}$ of $(g-r)$ linearly independent elements $\bar{f}_i$ which all have the property

$$M\bar{f}_i = M\mathbf{fa}_i = \mathbf{fMa}_i = 0. \tag{5.50}$$

If a basis for the subspace V is arranged such that it starts with these elements $\bar{f}_i$ for $i = 1,2,\ldots,g-r$, one then gets through the operator M automatically a basis for the space $V' = MV$ which contains r linearly independent elements. This theorem is of particular importance in contructing the basis for the subspace $V^{(m-1)} = N^{(m-1)}V$ of order $s^{(m)}$, which forms the starting point for the decomposition procedure.

In the following, we will apply some of these results to the operator sequence $M = N, N^2, \ldots, N^{m-1}$. Starting out from the relation $N\mathbf{f} = \mathbf{fN}$, it is evident that the sequence

$$g, g', g'', \ldots, g^{(m-1)}, \tag{5.51}$$

which is fundamental in determining the Segré characteristics, corresponds to the ranks of the sequence of matrices

$$\mathbf{1}, \mathbf{N}, \mathbf{N}^2, \ldots, \mathbf{N}^{m-1}, \tag{5.52}$$

i.e., to the numbers of linearly independent column vectors in each one of them. It should be observed that the evaluation of these numbers becomes particularly simple, if the matrix $\mathbf{N}$ has been brought to triangular form from the very beginning.

Construction of the Jordan projectors. In matrix theory, the decomposition of a nilpotent matrix $\mathbf{N}$ into Jordan blocks may be carried out by elementary algebraic methods involving only the handlings of vectors and matrices. Here we will try a slightly different approach, which adds one more aspect to the problem.

Let us assume that the stable subspace V defined by the product projection operator $O(T)$ has the minimal index m. This means that, for every element x of the full space $A = \{x\}$, one has $N^m O(T)x = 0$, i.e., the operator relation $N^m O(T) = 0$. Since further the operators N and O commute, one gets

$$N^m O = ON^m, \tag{5.53}$$

as well as the adjoint relation

$$(N^\dagger)^m O^\dagger = O^\dagger (N^\dagger)^m . \tag{5.54}$$

This implies that the subspace $V^\dagger$ defined by $O^\dagger$ has a minimal index $m^\dagger$ with respect to $O^\dagger$, having the property $m^\dagger \leqslant m$. It may be shown, however, that $m^\dagger = m$ and that the subspace $V^\dagger$ has indeed the same Segré characteristics as the subspace V. For this purpose, we may now span the space $V^\dagger = O^\dagger A$ by the set $\bar{\mathbf{g}} = O^\dagger \mathbf{f}$ introduced by (5.27) or by the reciprocal set $\bar{\mathbf{g}}_r = \bar{\mathbf{g}}\langle \mathbf{f}|\bar{\mathbf{g}}\rangle^{-1} = O^\dagger \mathbf{f}\langle \mathbf{f}|\mathbf{f}\rangle^{-1}$. According to (5.25), one has the relations

$$N\mathbf{f} = \mathbf{f}\mathbf{N}, \quad N^\dagger \bar{\mathbf{g}}_r = \bar{\mathbf{g}}_r \mathbf{N}^\dagger,$$

where $\mathbf{N}^\dagger$ is the adjoint of the matrix $\mathbf{N} = \langle \bar{\mathbf{g}}_r | N | f \rangle$. However, since the matrices

$$\mathbf{1}, \mathbf{N}^\dagger, (\mathbf{N}^\dagger)^2, \ldots, (\mathbf{N}^\dagger)^{m-1} \tag{5.55}$$

have the same ranks as the matrices (5.52), the Segré characteristics for the subspace $V^\dagger$ are the same as those for the subspace V.

Let us now consider an arbitrary element f_s of index m of the subspace V such that

$$N^m f_s = 0, \quad N^{m-1} f_s \neq 0, \tag{5.56}$$

and let us further introduce the sequence

$$f_1 = N^{m-1} f_s, \quad f_2 = N^{m-2} f_s, \ldots,$$
$$f_k = N^{m-k} f_s, \ldots, f_m = f_s, \tag{5.57}$$

in accordance with (5.32). The index s indicates that f_s is the "starting element" for the construction. We recall that the elements $\{f_1, f_2, \ldots, f_m\}$ are linearly independent and that they span the irreducible subspace $V(f_s)$ of order m.

By means of the adjoint product projector $O^\dagger$, one can now go from the subspace V to the subspace $V^\dagger$ through the formula $V^\dagger = O^\dagger V$ in accordance with the relation (5.27). Of particular interest is the projection

$$V^\dagger_{(m-1)} = O^\dagger V^{(m-1)}, \tag{5.58}$$

since it turns out $V^\dagger_{(m-1)}$ may be spanned by $s^{(m)}$ linearly independent elements all having the index m.

For the sake of simplicity, we will start by considering a single element $f_1 = N^{m-1} f_s$ and its projection:

$$g_s = O^\dagger f_1 . \tag{5.59}$$

Since $\langle g_s | f_1 \rangle = \langle O^\dagger f_1 | f_1 \rangle = \langle f_1 | O f_1 \rangle = \langle f_1 | f_1 \rangle \neq 0$, one has necessarily $g_s \neq 0$. In addition, we will now introduce the sequence

$$g_1 = g_s, \quad g_2 = N^\dagger g_1,$$
$$g_3 = N^\dagger g_2, \ldots, g_g = N^\dagger g_{g-1}, \tag{5.60}$$

i.e.,

$$g_k = (N^\dagger)^{k-1} g_s = (N^\dagger)^{k-1} O^\dagger f_1 . \tag{5.61}$$

In general, one has

$$\langle g_k | f_k \rangle = \langle (N^\dagger)^{k-1} O^\dagger f_1 | N^{m-k} f_s \rangle$$
$$= \langle f_1 | N^{m-1} f_s \rangle = \langle f_1 | f_1 \rangle \neq 0, \tag{5.62}$$

and this means that no one of the functions g_k for $k = 1,2,\ldots,m$ can be identically vanishing. In particular, one has $g_m = (N^\dagger)^{m-1} g_s \neq 0$, which means that the starting element g_s is of index m.

It is now easily shown that the matrix $\mathbf{\Delta} = \langle \bar{\mathbf{g}} | \bar{\mathbf{f}} \rangle$ formed by the sets $\bar{\mathbf{f}}$ and $\bar{\mathbf{g}}$ of order m is nonsingular. Observing that $O\bar{\mathbf{f}} = \bar{\mathbf{f}}$, one has

$$\Delta_{kl} = \langle g_k | f_l \rangle = \langle (N^\dagger)^{k-1} O^\dagger f_1 | N^{m-l} f_s \rangle$$
$$= \langle f_1 | N^{m+(k-l)-1} | f_s \rangle = t_{k-l}, \tag{5.63}$$

which shows that Δ_{kl} is a function of the difference $(k-l)$ only. Of course, one has $\Delta_{kl} = 0$, as soon as $m + (k-l) - 1 \geqslant m$, i.e., whenever $k - l \geqslant 1$. This means that $\mathbf{\Delta}$ is a triangular matrix with vanishing elements below the diagonal having the form

$$\mathbf{\Delta} = \begin{pmatrix} t_0, & t_{-1}, & t_{-2}, & t_{-3}, \ldots & \vdots \\ 0 & t_0, & t_{-1}, & t_{-2}, \ldots & \vdots \\ 0 & 0 & t_0, & t_{-1}, \ldots & \vdots \\ \vdots & \vdots & \vdots & \vdots & t_{-1} \\ 0 & 0 & 0 & \ldots & t_0 \end{pmatrix} \tag{5.64}$$

and that $\mathbf{\Delta}$ has the determinant

$$|\mathbf{\Delta}| = t_0^g = \langle f_1 | f_1 \rangle^g . \tag{5.65}$$

According to a well-known theorem—see, e.g., Appendix A—such a matrix has an inverse $\mathbf{d} = \mathbf{\Delta}^{-1}$, which is also triangular and has the property $d_{kl} = d_{k-l}$. Since the set $\bar{\mathbf{f}} = \{f_1, f_2, \ldots, f_m\}$ is stable under the operator N, and the set $\bar{\mathbf{g}} = \{g_1, g_2, \ldots, g_m\}$ is stable under the operator $N^\dagger$, one can now expect that the projector

$$Q = |\bar{\mathbf{f}}\rangle \langle \bar{\mathbf{g}}|\bar{\mathbf{f}}\rangle^{-1} \langle \bar{\mathbf{g}}| = \sum_{k,l} |f_k\rangle d_{kl} \langle g_l| \tag{5.66}$$

constructed according to (4.21) should decompose the operator N, i.e., $NQ = QN$. Observing that $Nf_k = f_{k-1}$, and that $\langle g_l | N = \langle N^\dagger g_l | = \langle g_{l+1} |$, and that $d_{k+1,l} = d_{k+1-l} = d_{k,l-1}$, one obtains that

$$NQ = \sum_{k,l} |f_{k-1}\rangle d_{k,l} \langle g_l| = \sum_{k,l} |f_k\rangle d_{k+1,l} \langle g_l|$$

$$= \sum_{k,l} |f_k\rangle d_{k,l-1} \langle g_l| = \sum_{k,l} |f_k\rangle d_{k,l} \langle g_{l+1}| = QN, \tag{5.67}$$

which proves the statement. It is interesting to observe that, if one introduces the reciprocal basis

$$\bar{\mathbf{g}}_r = \bar{\mathbf{g}}\langle \bar{\mathbf{f}}|\bar{\mathbf{g}}\rangle^{-1}, \tag{5.68}$$

one has not only the property $\langle \bar{\mathbf{g}}_r | \mathbf{f} \rangle = \mathbf{1}$ but also

$$N^\dagger g_{r,k} = g_{r,k+1} . \tag{5.69}$$

We note that the projector Q defined by (5.66) is essentially characterized by the starting element f_s, and that it is sometimes convenient to denote it by the symbol $Q(f_s)$.

In order to proceed, we note that one also has the relations $OQ = QO = Q$. That means that the operator

$$P = O - Q \tag{5.70}$$

has the property $P^2 = O^2 + Q^2 - OQ - QO = O - Q = P$, i.e., that P is a projector having the order $\mathrm{Tr}\, P = \mathrm{Tr}\, O - \mathrm{Tr}\, Q = g - m$. One has further $QP = PQ = 0$, which means that Q and P are mutually exclusive projectors.

In the stable subspace $V_P = PV$, the number of irreducible subspaces of order m has been decreased by one unit in comparison to V. If $s^{(m)} \geqslant 2$, one should now pick another

starting element f_s' of index m out of the subspace V_P and repeat the procedure leading to the construction of the new projector $Q(f_s')$. Since the starting element f_s' satisfied the relation $Pf_s' = f_s'$, and the operators P and N commute, one obtains

$$PQ(f_s') = Q(f_s')P = Q(f_s'), \tag{5.71}$$

which implies that the projectors $Q(f_s)$ and $Q(f_s')$ are mutually exclusive.

Proceeding in this way, one first exhausts all the linearly independent elements of index m, and then continues with the elements of index $(m-1)$, etc. In this way, one obtains a decomposition of the product projection operator $O(T)$ into mutually exclusive projectors $Q(f_s), Q(f_s'), Q(f_s''), \cdots$, which all commute with N, and we observe that the Segré characteristics given by the second-order difference (5.45) are of great help as guidance in this connection.

In concluding this subsection, it should be observed that the construction of the associated irreducible subspaces $V(f)$ and $V^\dagger(g)$ given above is not particularly elegant, but it gives at least the associated projector Q and $Q^\dagger$ without any further ado.

Reduced Cayley–Hamilton equation. One of the most important results in this section is the establishment of the existence of a *minimal index* m_k for each subspace V_k defined by a product projection operator $O_k(T)$. From our discussion, it is evident that m_k must be identical to the largest Segré characteristic associated with the space V_k. Instead of the fundamental relations (5.11), one has now

$$N_k^{m_k} O_k = O_k N_k^{m_k} = 0, \tag{5.72}$$

in accordance with (5.53). For self-adjoint and normal operators, one can actually prove the general theorem that all $m_k = 1$, but here we will treat the general case when $m_k \geqslant 1$.

It is now worthwhile to go back and re-examine the reasoning which formed the start of this section. Instead of the characteristic polynomial (5.1), we will here consider the reduced characteristic polynomial:

$$F_1(z) \equiv \prod_k (z - T_k)^{m_k}, \tag{5.73}$$

where the minimal indices m_k replace the degeneracies g_k in the previous expression. Expanding the inverse of $F_1(z)$ in terms of partial fractions,

$$\frac{1}{F_1(z)} \equiv \sum_k \frac{r_k(z)}{(z-\lambda_k)^{m_k}}, \tag{5.74}$$

where $r_k(z)$ is a specific polynomial in z of a degree equal to or less than $(m_k - 1)$, and introducing the notations

$$O_k^{(1)}(z) \equiv r_k(z) \prod_{l \neq k} (z - \lambda_l)^{m_l}, \tag{5.75}$$

one obtains the algebraic identity

$$1 \equiv \sum_k O_k^{(1)}(z). \tag{5.76}$$

The operators

$$O_k^{(1)}(T) = r_k(T) \prod_{l \neq k} (T - \lambda_l \cdot I)^{m_l} \tag{5.77}$$

are polynomials in the operator T, which apparently satisfy the identity

$$\sum_k O_k^{(1)}(T) = I. \tag{5.78}$$

Observing that the operator $O_k^{(1)}(T) = r_k(T)\Pi_{l \neq k} N_l^{m_l}$ contains the factor $N_l^{m_l}$, one gets immediately for $l \neq k$

$$O_k^{(1)}(T) O_l(T) = O_l(T) O_k^{(1)}(T) = 0, \tag{5.79}$$

in accordance with (5.72). Multiplying the resolution (5.9) to the left by $O_k^{(1)}(T)$, one hence obtains

$$\begin{aligned} O_k^{(1)}(T) &= O_k^{(1)}(T)\left(\sum_l O_l(T)\right) = \sum_l O_k^{(1)}(T) O_l(T) \\ &= O_k^{(1)}(T) O_k(T) = O_k(T) O_k^{(1)}(T). \end{aligned} \tag{5.80}$$

Using (5.72) once more, one gets further

$$N_k^{m_k} O_k^{(1)}(T) = O_k^{(1)}(T) N_k^{m_k} = 0, \tag{5.81}$$

as well as

$$O_k^{(1)}(T) O_l^{(1)}(T) = 0 \quad \text{for } k \neq l. \tag{5.82}$$

Multiplying (5.78) to the left by $O_k^{(1)}(T)$, one has also

$$\begin{aligned} O_l^{(1)}(T) &= O_l^{(1)}(T)\left[\sum_k O_k^{(1)}(T)\right] = \sum_k O_l^{(1)}(T) O_k^{(1)}(T) \\ &= O_l^{(1)}(T) O_l^{(1)}(T). \end{aligned} \tag{5.83}$$

The relations (5.82) and (5.83) indicate that the operators $O_k^{(1)}(T)$ for $k = 1,2,3,\cdots$ are mutually exclusive projectors which form a resolution of the identity (5.78). Since further

$$TO_k^{(1)} = O_k^{(1)} T, \tag{5.84}$$

it is evident that the projectors $O_k^{(1)}$ decompose the operator T in accordance with (4.37). These projectors provide a decomposition of the full space $A = \{x\}$ into subspace $V_k^{(1)} = O_k^{(1)} A$, which are stable under the operator T. Since further $O_k V_k^{(1)} = V_k^{(1)}$, it is clear that the subspaces $V_k^{(1)}$ are going to replace the previously used subspaces V_k in our discussions; in fact, they are identical.

Finally, we note that if the product operator

$$F_1(T) = \prod_k N_k^{m_k} \tag{5.85}$$

works on the identity operator as defined by (5.78), one gets the result zero, which implies $F_1(T)X = 0$ for all elements x in the space $A = \{x\}$. The relation

$$F_1(T) = \prod_k (T - \lambda_k \cdot I)^{m_k} = 0, \tag{5.86}$$

is known as the *reduced Cayley–Hamilton equation.*

It is interesting to observe that, even if the space $A = \{x\}$ is of *infinite order,* one may still apply most of the formulas in this subsection as long as the operator T has only a *finite* number of eigenvalues $\lambda_1, \lambda_2, \lambda_3, \cdots$ in the complex plane with *finite* minimal indices $m_1, m_2, m_3, \cdots$—even if the degeneracies themselves, $g_1, g_2, g_3, \cdots$, are infinite. In such a case, the starting point for the theory is the reduced Cayley–Hamilton equation (5.86), whereas fundamental projectors $O_k^{(1)}$ are again defined by the relation (5.77).

Many equations in physics may indeed be interpreted as reduced Cayley–Hamilton equations. If one considers the

exchange operator P_{12} having the property $P_{12}f(1,2) = f(2,1)$, it satisfies the relation $P_{12}^2 = I$ or

$$(P_{12} - I)(P_{12} + I) = 0, \tag{5.87}$$

i.e., P_{12} has the eigenvalue $\lambda = +1$ with $m = 1, g = \infty$, and the eigenvalue $\lambda = -1$ with $m = 1, g = \infty$. In many applications, this approach has been used successfully to treat the constants of motion[5] of many-particle systems in the quantum theory of matter.

Summary; the classical canonical form. Let us now try to summarize the results of this section and compare them with those previously obtained. We have been studying a linear space $A = \{x\}$ of finite order n, but the results are also applicable to the case of a finite order stable subspace of an infinite Hilbert space. Let us span the space $A = \{x\}$ by a linearly independent set $\mathbf{X} = \{X_1, X_2, X_3, \ldots, X_n\}$ which serves as a basis, so that one has the expansion theorem

$$x = \sum_k X_k a_k = \mathbf{Xa}. \tag{5.88}$$

One has the metric matrix $\mathbf{\Delta} = \langle \mathbf{X}|\mathbf{X}\rangle$ with the elements $\Delta_{kl} = \langle X_k | X_l \rangle$ having the property $\mathbf{\Delta}^\dagger = \mathbf{\Delta}$. The reciprocal basis $\mathbf{X}_r = \mathbf{X}\mathbf{\Delta}^{-1}$ satisfies the relations

$$\langle \mathbf{X}|\mathbf{X}_r\rangle = \langle \mathbf{X}_r|\mathbf{X}\rangle = \mathbf{1}, \tag{5.89}$$

and $\mathbf{X}$ and $\mathbf{X}_r$ are said to be biorthonormal. The basis $\boldsymbol{\varphi} = \mathbf{X}\mathbf{\Delta}^{-1/2}$ satisfies the relation $\langle \boldsymbol{\varphi}|\boldsymbol{\varphi}\rangle = \mathbf{1}$, and it is hence orthonormal in the ordinary sense; we note that the set $\boldsymbol{\varphi}$ is self-reciprocal. Multiplying the relation (5.88) to the left by $\langle \mathbf{X}_r|$, and solving for $\mathbf{a}$, one obtains

$$\mathbf{a} = \langle \mathbf{X}_r | x \rangle. \tag{5.90}$$

This means that the expansion theorem may be written in the form $x = \mathbf{X}\langle \mathbf{X}_r | x\rangle$ for all x, which gives the operator relations

$$\begin{aligned} I &= |\mathbf{X}\rangle\langle \mathbf{X}_r| = |\mathbf{X}\rangle\mathbf{\Delta}^{-1}\langle \mathbf{X}| \\ &= |\mathbf{X}_r\rangle\langle \mathbf{X}|, \end{aligned} \tag{5.91}$$

which may be considered as various types of "resolutions of the identity."

Let us now consider a pair of adjoint operators T and $T^\dagger$ characterized by the relations

$$T\mathbf{X} = \mathbf{XT}, \quad T^\dagger \mathbf{X} = \mathbf{XR}, \tag{5.92}$$

where $\mathbf{T}$ and $\mathbf{R}$ may be considered as the matrix representations of T and $T^\dagger$ with respect to the basis $\mathbf{X}$. IF $x = \mathbf{Xa}$, one gets directly $Tx = \mathbf{XTa}$ and $T^\dagger x = \mathbf{XRa}$, so the operators are fully described by their matrices. Multiplying (5.92) to the left by $\langle \mathbf{X}_r|$ and solving for $\mathbf{T}$ and $\mathbf{R}$, respectively, one has

$$\mathbf{T} = \langle \mathbf{X}_r | T\mathbf{X}\rangle, \quad \mathbf{R} = \langle \mathbf{X}_r | T^\dagger \mathbf{X}\rangle. \tag{5.93}$$

Using the definition (1.1), one obtains further

$$\begin{aligned} \mathbf{R} &= \langle \mathbf{X}_r | T^\dagger \mathbf{X}\rangle = \langle \mathbf{X}\mathbf{\Delta}^{-1} | T^\dagger \mathbf{X}\rangle = \mathbf{\Delta}^{-1}\langle \mathbf{X}|T^\dagger \mathbf{X}\rangle \\ &= \mathbf{\Delta}^{-1}\langle T\mathbf{X}|\mathbf{X}\rangle = \mathbf{\Delta}^{-1}\langle \mathbf{XT}|\mathbf{X}\rangle = \mathbf{\Delta}^{-1}\mathbf{T}^\dagger\mathbf{\Delta}. \end{aligned} \tag{5.94}$$

If the basis undergoes a linear transformation $\mathbf{X}' = \mathbf{X}\boldsymbol{\alpha}$, the matrix $\mathbf{T}$ undergoes the similarity transformation $\mathbf{T}' = \boldsymbol{\alpha}^{-1}\mathbf{T}\boldsymbol{\alpha}$. Putting $\mathbf{X}_r = \mathbf{X}\mathbf{\Delta}^{-1}$, one gets directly $\mathbf{R}_r = \mathbf{\Delta}\mathbf{R}\mathbf{\Delta}^{-1} = \mathbf{T}^\dagger$. Hence, one has

$$T\mathbf{X} = \mathbf{XT}, \quad T^\dagger \mathbf{X}_r = \mathbf{X}_r \mathbf{T}^\dagger, \tag{5.95}$$

where $\mathbf{T}^\dagger$ is the adjoint of the matrix $\mathbf{T}$.

In connection with relation (4.4), we discussed previously how a matrix $\mathbf{T}$ could be *block-diagonalized* as far as ever possible by means of a similarity transformation $\boldsymbol{\gamma}$. The answer obtained in this section is well known in matrix theory; every matrix $\mathbf{T}$ may first of all be block-diagonalized after its eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots$ and, if the eigenvalue λ is nondegenerate, the corresponding block is of order 1 and consists of the eigenvalue on the diagonal. However, if the eigenvalue λ is degenerate of order g, the corresponding diagonal block may be transformed into a sequence of Jordan blocks of type (4.5), the orders of which are given by the Segré characteristics. This form $\boldsymbol{\lambda}$ of the matrix is known as the *classical canonical form*. Hence one has

$$\boldsymbol{\gamma}^{-1}\mathbf{T}\boldsymbol{\gamma} = \boldsymbol{\lambda}, \tag{5.96}$$

as well as the adjoint relation

$$\boldsymbol{\gamma}^\dagger \mathbf{T}^\dagger (\boldsymbol{\gamma}^\dagger)^{-1} = \boldsymbol{\lambda}^\dagger, \tag{5.97}$$

where $\boldsymbol{\lambda}^\dagger$ has the complex conjugate eigenvalues λ_k^* on the diagonal and the 1's below instead of above this diagonal.

Many of the relations found in Sec. 2 for operators having only distinct eigenvalues may now be extended also to more general operators. Introducing the eigenbases $\mathbf{C}$ and $\mathbf{D}$ to the operators T and $T^\dagger$, respectively, through the relation

$$\mathbf{C} = \mathbf{X}\boldsymbol{\gamma}, \quad \mathbf{D} = \mathbf{X}_r(\boldsymbol{\gamma}^\dagger)^{-1}, \tag{5.98}$$

one gets immediately

$$T\mathbf{C} = \mathbf{C}\boldsymbol{\lambda}, \quad T^\dagger \mathbf{D} = \mathbf{D}\boldsymbol{\lambda}^\dagger. \tag{5.99}$$

One has further

$$\langle \mathbf{D}|\mathbf{C}\rangle = \langle \mathbf{X}_r(\boldsymbol{\gamma}^\dagger)^{-1}|\mathbf{X}\boldsymbol{\gamma}\rangle = \boldsymbol{\gamma}^{-1}\langle \mathbf{X}_r|\mathbf{X}\rangle\boldsymbol{\gamma} = \mathbf{1}, \tag{5.100}$$

which is a generalization of the *biorthonormality* theorem. It is interesting to observe that the set $\mathbf{D}$ may be evaluated from the set $\mathbf{C}$, since $\mathbf{D}$ is identical to the reciprocal basis $\mathbf{C}_r = \mathbf{C}\langle \mathbf{C}|\mathbf{C}\rangle^{-1}$:

$$\begin{aligned} \mathbf{C}_r &= \mathbf{C}\langle \mathbf{C}|\mathbf{C}\rangle^{-1} = \mathbf{X}\boldsymbol{\gamma}[\langle \mathbf{X}\boldsymbol{\gamma}|\mathbf{X}\boldsymbol{\gamma}\rangle]^{-1} \\ &= \mathbf{X}\boldsymbol{\gamma}[\boldsymbol{\gamma}^\dagger\langle \mathbf{X}|\mathbf{X}\rangle\boldsymbol{\gamma}]^{-1} \\ &= \mathbf{X}\boldsymbol{\gamma}\boldsymbol{\gamma}^{-1}\langle \mathbf{X}|\mathbf{X}\rangle^{-1}(\boldsymbol{\gamma}^\dagger)^{-1} \\ &= \mathbf{X}_r(\boldsymbol{\gamma}^\dagger)^{-1} = \mathbf{D}. \end{aligned} \tag{5.101}$$

In Sec. 2, we had further been able to express the resolution of the identity in form of the second relation (2.22). Here one gets similarly

$$\begin{aligned} |\mathbf{C}\rangle\langle \mathbf{D}| &= |\mathbf{X}\boldsymbol{\gamma}\rangle\langle \mathbf{X}_r(\boldsymbol{\gamma}^\dagger)^{-1}| \\ &= |\mathbf{X}\rangle\boldsymbol{\gamma}\boldsymbol{\gamma}^{-1}\langle \mathbf{X}_r| \\ &= |\mathbf{X}\rangle\langle \mathbf{X}_r| = 1, \end{aligned} \tag{5.102}$$

as well as the adjoint relation

$$|\mathbf{D}\rangle\langle \mathbf{C}| = 1. \tag{5.103}$$

Letting the operators T and $T^\dagger$ work on (5.101) and (5.103), respectively, one obtains

$$T = |\mathbf{C}\rangle\boldsymbol{\lambda}\langle \mathbf{D}|, \quad T^\dagger = |\mathbf{D}\rangle\boldsymbol{\lambda}^\dagger\langle \mathbf{C}|, \tag{5.104}$$

which are the analogs of the spectral resolutions in the distinct case. However, because of the existence of degenerate eigenvalues and Jordan blocks, they are now slightly more complicated than before.

In concluding this subsection, we note that by replacing the eigenvalue problem with the stability problem—or, which is the same, the diagonalization procedure with the block-diagonalization procedure—one may be able to generalize most of the fundamental theorems found in Sec. 2 also to more general operators.

Bivariational principle and the stability problem. So far, little has been said in the summary about any possible extensions of the bivariational principle of Sec. 3 to the stability problem formulated in terms of subspaces and projectors. In order to proceed, we observe[6] that if a physical system is described by a system operator Γ having the property $\Gamma^\dagger = \Gamma$, then the expectation value of a physical observable $F = F^\dagger$ is given by the expression

$$\langle F \rangle_{\rm av} = \mathrm{Tr}\, F\Gamma / \mathrm{Tr}\, \Gamma. \tag{5.105}$$

In the case of a homogeneous ensemble, one has the auxiliary condition $\Gamma^2 = \Gamma\, \mathrm{Tr}\, \Gamma$.

In studying a general linear operator T, we will in analogy with (5.105) consider the quantities

$$I_1 = \mathrm{Tr}\, T\Gamma / \mathrm{Tr}\, \Gamma, \quad I_2 = \mathrm{Tr}\, T^\dagger \Gamma^\dagger / \mathrm{Tr}\, \Gamma^\dagger = I_1^*, \tag{5.106}$$

where $\Gamma \neq \Gamma^\dagger$. These expressions are obvious generalizations of (3.6) for $\Gamma = |x_1\rangle\langle x_2|x_1\rangle^{-1}\langle x_2|$. In order to connect the quantities (5.106) with the stability problem, we will now assume that the general operator is a *projector* satisfying the relations

$$\Gamma^2 = \Gamma, \quad \mathrm{Tr}\, \Gamma = g. \tag{5.107}$$

For any variations $\delta\Gamma$ of the operator Γ, one gets then the auxiliary conditions

$$\delta\Gamma = \Gamma\delta\Gamma + \delta\Gamma\cdot\Gamma, \quad \Gamma\delta\Gamma\Gamma = 0, \quad \mathrm{Tr}\, \delta\Gamma = 0, \tag{5.108}$$

where the second relation is obtained from the first by multiplying to the left or right by Γ. For the variation of I_1, one obtains

$$\begin{aligned}\delta I_1 &= (1/\mathrm{Tr}\, \Gamma)[\mathrm{Tr}\, T\delta\Gamma - I_1\, \mathrm{Tr}\, \delta\Gamma] \\ &= (1/g)\mathrm{Tr}(T\Gamma + \Gamma T)\delta\Gamma.\end{aligned} \tag{5.109}$$

One would hence perhaps expect that a sufficient condition that $\delta I_1 = 0$ is that $T\Gamma + \Gamma T = 0$. In reality, this condition is never fulfilled, since it leads to a contradiction. Instead, one has the simple condition $T\Gamma = \Gamma T$ since, in such a case, one has

$$\begin{aligned}\delta I_1 &= (2/g)\, \mathrm{Tr}\, T\Gamma\delta T = (2/g)\, \mathrm{Tr}\, T\Gamma^2\delta T \\ &= (2/g)\, \mathrm{Tr}\, \Gamma T\Gamma\delta\Gamma = (2/g)\, \mathrm{Tr}\, T(\Gamma\delta\Gamma\Gamma) = 0.\end{aligned} \tag{5.110}$$

In order to study the necessary conditions that $\delta I_1 = 0$ when $\Gamma^2 = \Gamma$, we will multiply the second relation (5.108) by an operator Λ—corresponding to the ordinary Lagrangian multipliers—and take the trace, which gives the auxiliary condition

$$\mathrm{Tr}\, \Lambda\Gamma\delta\Gamma\Gamma = \mathrm{Tr}(\Gamma\Lambda\Gamma)\delta\Lambda = 0. \tag{5.111}$$

Combining (5.109) and (5.111), one obtains the relation

$$\delta I_1 = (1/g)\mathrm{Tr}(T\Gamma + \Gamma T - \Gamma\Lambda\Gamma)\delta\Gamma = 0 \tag{5.112}$$

for arbitrary variations $\delta\Gamma$, which leads to the condition

$$T\Gamma + \Gamma T - \Gamma\Lambda\Gamma = 0. \tag{5.113}$$

Multiplying this relation to the left by Γ and to the right by Γ, respectively, one obtains

$$\Gamma T\Gamma + \Gamma T = \Gamma\Lambda\Gamma = T\Gamma + \Gamma T\Gamma, \tag{5.114}$$

i.e.,

$$T\Gamma = \Gamma T, \tag{5.115}$$

which is hence a necessary and sufficient condition that $\delta I_1 = 0$ when $\Gamma^2 = \Gamma$. This result implies that Γ must be a projector O, which *decomposes* the operator T.

The relation (5.115) connects the bivariational principle $\delta I_1 = 0$ with the stability problem for the operator T, and we will now look for approximate solutions. Let us assume that, as in Sec. 3, we have two truncated sets $\phi = \{\phi_1, \phi_2, \ldots, \phi_m\}$ and $\psi = \{\psi_1, \psi_2, \ldots, \psi_m\}$ of order m at our disposal, and that the matrix $\langle\psi|\phi\rangle$ is nonsingular. According to (4.21), one can then construct a pair of adjoint projectors

$$\bar{O} = |\phi\rangle\langle\psi|\phi\rangle^{-1}\langle\psi|, \tag{5.116}$$

$$\bar{O}^\dagger = |\psi\rangle\langle\phi|\psi\rangle^{-1}\langle\phi|, \tag{5.117}$$

where the bar indicates that we are dealing with approximate quantities. In order to study the degree of approximation, it may be convenient to introduce the difference

$$\omega = T\bar{O} - \bar{O}T, \tag{5.118}$$

since the quantity

$$\gamma = \mathrm{Tr}\, \omega^\dagger\omega \geqslant 0 \tag{5.119}$$

is then non-negative and zero only when the relation $T\bar{O} = \bar{O}T$ is exactly fulfilled. We note that (5.119) is a generalization of the ordinary concept of the "width" of an operator T.

As before, we will now introduce the reciprocal set $\psi_r = \psi\langle\phi|\psi\rangle^{-1}$ having the property $\langle\psi_r|\phi\rangle = \langle\phi|\psi_r\rangle = \mathbf{1}$, which gives

$$\bar{O} = |\phi\rangle\langle\psi_r|, \quad \bar{O}^\dagger = |\psi_r\rangle\langle\phi|. \tag{5.120}$$

Assuming that the relations $T\bar{O} = \bar{O}T$ and $T^\dagger\bar{O}^\dagger = \bar{O}^\dagger T^\dagger$ are exactly valid, one gets directly

$$T|\phi\rangle\langle\psi_r| = |\phi\rangle\langle\psi_r|T, \tag{5.121}$$

$$T^\dagger|\psi_r\rangle\langle\phi| = |\psi_r\rangle\langle\phi|T^\dagger. \tag{5.122}$$

Multiplying the first to the right by $|\phi\rangle$ and the second to the right by $|\psi_r\rangle$, respectively, and introducing the notation $\bar{\mathbf{T}} = \langle\psi_r|T|\phi\rangle$, one obtains

$$T\phi = \phi\bar{\mathbf{T}}, \quad T^\dagger\psi_r = \psi_r\bar{\mathbf{R}}. \tag{5.123}$$

Here, $\bar{\mathbf{R}} = \langle\phi|T^\dagger\psi_r\rangle = \langle T\phi|\psi_r\rangle = \bar{\mathbf{T}}^\dagger$, as before. Introducing the matrix $\bar{\gamma}$ which brings $\bar{\mathbf{T}}$ to classical canonical form $\bar{\lambda}$, one obtains

$$\bar{\gamma}^{-1}\bar{\mathbf{T}}\bar{\gamma} = \bar{\lambda}, \quad \bar{\gamma}^\dagger\bar{\mathbf{T}}^\dagger(\bar{\gamma}^\dagger)^{-1} = \bar{\lambda}^\dagger, \tag{5.124}$$

and introducing the approximate eigenbases $\bar{\mathbf{C}} = \phi\bar{\gamma}$ and $\bar{\mathbf{D}} = \psi_r(\bar{\gamma}^\dagger)^{-1}$, one obtains the approximate relations

$$T\bar{\mathbf{C}} = \bar{\mathbf{C}}\bar{\lambda}, \quad T^\dagger\bar{\mathbf{D}} = \bar{\mathbf{D}}\bar{\lambda}^\dagger. \tag{5.125}$$

It is then easily checked that the other fundamental relations in Sec. 3 are also valid.

We observe finally that the transformation of $\bar{\mathbf{T}}$ and $\bar{\mathbf{T}}^\dagger$ to classical canonical form corresponds to the decomposition of the projectors $\bar{O}$ and $\bar{O}^\dagger$ into irreducible projectors

$\overline{O}_k$ decomposing the two operators, so that

$$\overline{O} = \sum_k \overline{O}_k, \quad \overline{O}^\dagger = \sum_k \overline{O}_k^\dagger. \tag{5.126}$$

At this stage, one can check the accuracy of each projector $\overline{O}_k$ by forming the commutators

$$\omega_k = T\overline{O}_k - \overline{O}_k T, \tag{5.127}$$

and by evaluating the quantities

$$\gamma_k = \mathrm{Tr}\, \omega_k^\dagger \omega_k \geqslant 0. \tag{5.128}$$

It should be observed, however, that—even if this scheme is also approximate—it has the great advantage that the approximate solutions have essentially the same properties as the exact ones. In certain connections, it may perhaps be more appropriate to consider the approximate operators

$$\overline{T} = \overline{O}T\overline{O} = |\phi\rangle\langle\psi_r|T|\phi\rangle\langle\psi_r| = |\phi\rangle\overline{\mathbf{T}}\langle\psi_r|, \tag{5.129}$$

$$\overline{T}^\dagger = \overline{O}^\dagger T^\dagger \overline{O}^\dagger = |\psi_r\rangle\langle\phi|T^\dagger|\psi_r\rangle\langle\phi| = |\psi_r\rangle\mathbf{T}^\dagger\langle\phi|, \tag{5.130}$$

which satisfy the relations

$$\overline{O}\overline{T} = \overline{T}\overline{O}, \quad \overline{O}^\dagger\overline{T}^\dagger = \overline{T}^\dagger\overline{O}^\dagger \tag{5.131}$$

exactly. This means that the approximate solutions for the operators T and $T^\dagger$ previously obtained may be considered as exact solutions associated with the operators $\overline{T}$ and $\overline{T}^\dagger$.

6. COMPLEX CONJUGATE OPERATORS

In this section, we will consider certain aspects which may be important in the numerical treatment of general linear operators T. Let us consider a linear space $A = \{x\}$, where the elements x represent certain complex functions of some specified variables, and let the symbol x^* represent the complex conjugate function. Such a space is said to be *stable under complex conjugation* if both x and x^* belong to the space.

Let us assume that $A = \{x\}$ has a basis $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ of finite order n, and that every element x may be expressed in the form

$$x = \sum_k X_k a_k = \mathbf{X}\mathbf{a}. \tag{6.1}$$

This gives directly

$$x^* = \mathbf{X}^*\mathbf{a}^* = \mathbf{X}\mathbf{a}', \tag{6.2}$$

where $\mathbf{X}^* = (X_1^*, X_2^*, \ldots, X_m^*)$ is the complex conjugate basis, which may be expressed in terms of the original basis $\mathbf{X}$, so that

$$\mathbf{X}^* = \mathbf{X}\boldsymbol{\alpha}, \quad \mathbf{X} = \mathbf{X}^*\boldsymbol{\alpha}^{-1}. \tag{6.3}$$

This is a rather special type of linear transformation, since one gets directly $\mathbf{X}^* = \mathbf{X}(\boldsymbol{\alpha}^{-1})$, i.e.,

$$\boldsymbol{\alpha} = (\boldsymbol{\alpha}^*)^{-1}, \quad \boldsymbol{\alpha}^*\boldsymbol{\alpha} = \mathbf{1}. \tag{6.4}$$

Since $|\boldsymbol{\alpha}|^*|\boldsymbol{\alpha}| = 1$, the absolute value of the determinant $|\boldsymbol{\alpha}|$ is hence 1. Combining (6.2) and (6.4), one gets further

$$\mathbf{a}' = \boldsymbol{\alpha}\mathbf{a}^*. \tag{6.5}$$

A somewhat different way of approaching this problem is to use the identity

$$\mathbf{X} = \tfrac{1}{2}(\mathbf{X} + \mathbf{X}^*) + \tfrac{1}{2}(\mathbf{X} - \mathbf{X}^*) = \boldsymbol{\Phi}_1 + i\boldsymbol{\Phi}_2, \tag{6.6}$$

and to consider the set $\{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2\}$ having the elements

$$\boldsymbol{\Phi}_1 = \tfrac{1}{2}(\mathbf{X} + \mathbf{X}^*), \quad \boldsymbol{\Phi}_2 = (1/2i)(\mathbf{X} - \mathbf{X}^*), \tag{6.7}$$

with the property $\boldsymbol{\Phi}_1^* = \boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2^* = \boldsymbol{\Phi}_2$. Starting with the first element in the set $\{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2\}$ and leaving out all elements which are either vanishing or linear combinations of the preceding elements, one arrives at a sequence $\boldsymbol{\Phi}$ of linearly independent elements which may serve as a basis. Since $\boldsymbol{\Phi} = \boldsymbol{\Phi}^*$, one speaks of a real basis.

It is evident that, if the original space $A = \{x\}$ is stable under complex conjugation, this construction will not change the order of the space. On the other hand, if this is not the case, the sequence $\{\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2\}$ contains twice as many elements as the original set $\mathbf{X}$ and, by introducing the new basis $\boldsymbol{\Phi}$, one may hence have increased the order of the space and extended the original space $A = \{x\}$, so that the new space becomes stable under complex conjugation. Instead of (6.1) and (6.2), one gets the simpler relations

$$x = \boldsymbol{\Phi}\mathbf{a}, \quad x^* = \boldsymbol{\Phi}\mathbf{a}^*, \tag{6.8}$$

where the column vector $\mathbf{a}$ is different from the one occurring in (6.1). Because of the property expressed by (6.8), the operation of "complex conjugation" is often described as *antilinear*.

Irrespective of the properties of the basis, the *complex conjugate operator* T^* is defined by the relation

$$T^*x = (Tx^*)^*. \tag{6.9}$$

Observing that, for the domain of T^*, one has $D(T^*) = \{D(T)\}^*$, one gets directly the formulas

$$(T_1\alpha_1 + T_2\alpha_2)^* = T_1^*\alpha_1^* + T_2^*\alpha_2^*, \tag{6.10}$$

$$(T_1T_2)^* = T_1^*T_2^*, \quad (T^*)^* = T. \tag{6.11}$$

We note that, because of the special form of the first relation (6.11), the complex conjugation is *not* an involution. In the case when the operator T is expressed analytically in terms of real and imaginary quantities, one obtains the expression for T^* simply by replacing the imaginary unit i by $i^* = -i$. For instance, for the momentum operator in quantum mechanics

$$p = \frac{h}{2\pi i}\frac{\partial}{\partial x}, \tag{6.12}$$

one has $p^* = -p$ as well as $p^\dagger = p$.

For the matrix representation of T^*, one gets directly, by using the definitions,

$$T^*\boldsymbol{\Phi} = (T\boldsymbol{\Phi})^* = (\boldsymbol{\Phi}\mathbf{T})^* = \boldsymbol{\Phi}\mathbf{T}^*. \tag{6.13}$$

From the relation $\boldsymbol{\gamma}^{-1}\mathbf{T}\boldsymbol{\gamma} = \boldsymbol{\lambda}$, it follows also that $(\boldsymbol{\gamma}^*)^{-1}\mathbf{T}^*\boldsymbol{\gamma}^* = \boldsymbol{\lambda}^*$, where $\boldsymbol{\lambda}^*$ has the same classical canonical form as $\boldsymbol{\lambda}$ with the eigenvalues λ_k replaced by λ_k^* but with the 1's of the Jordan block still above the diagonal. Since $\lambda^\dagger$ may be obtained from $\boldsymbol{\lambda}^*$ simply by permutating the basis elements of each Jordan block, the operators $T^\dagger$ and T^* are apparently connected by a similarity transformation. We note finally that, if $OT = TO$, one has also $O^*T^* = T^*O^*$, i.e., the projectors O^* decompose the operator T^*.

In relativistic quantum theory, the complex conjugation is closely associated with the fundamental operation of

"charge conjugation" and time reversal, but we will study it here in a more elementary connection.

Special case when $T^\dagger = T^*$. In the partitioning technique as well as in the complex-scaling method, there are many examples of operators having the special property

$$T^\dagger = T^*, \tag{6.14}$$

which means that the similarity transformation in the general case is replaced by an identity. From the eigenvalue problems $TC = \lambda C$ and $T^\dagger D = \mu D$, one gets directly $T^*C^* = \lambda^* C^*$, which gives $D = C^*$ for $\mu = \lambda^*$. In general, one has the property

$$\mathbf{D} = \mathbf{C}^* = \mathbf{C}\langle \mathbf{C}|\mathbf{C}\rangle^{-1}, \tag{6.15}$$

which is a special case of (6.3). The result implies that

$$\langle \mathbf{C}|\mathbf{C}\rangle^*\langle \mathbf{C}|\mathbf{C}\rangle = \mathbf{1}, \tag{6.16}$$

in accordance with (6.4), and the absolute value of the determinant $|\langle \mathbf{C}|\mathbf{C}\rangle|$ is then equal to 1.

It should be observed that, if one has introduced a real basis $\boldsymbol{\phi}$ as well as its reciprocal basis $\boldsymbol{\phi}_r$ through the relation

$$\boldsymbol{\phi}_r = \boldsymbol{\phi}\langle \boldsymbol{\phi}|\boldsymbol{\phi}\rangle^{-1}, \tag{6.17}$$

with the properties $\langle \boldsymbol{\phi}|\boldsymbol{\phi}_r\rangle = \langle \boldsymbol{\phi}_r|\boldsymbol{\phi}\rangle = \mathbf{1}$, then also the reciprocal basis $\boldsymbol{\phi}_r$ is real. One has then the matrix representations

$$T\boldsymbol{\phi} = \boldsymbol{\phi}\mathbf{T}, \quad T^*\boldsymbol{\phi} = \boldsymbol{\phi}\mathbf{T}^*, \quad T^\dagger\boldsymbol{\phi}_r = \boldsymbol{\phi}_r\mathbf{T}^\dagger, \tag{6.18}$$

where $\mathbf{T} = \langle \boldsymbol{\phi}_r|T\boldsymbol{\phi}]$ and $\mathbf{T}^\dagger$ is its adjoint matrix. In the special case when $T^\dagger = T^*$, it is convenient to introduce an orthonormal real basis, e.g., $\boldsymbol{\varphi} = \boldsymbol{\phi}\langle \boldsymbol{\phi}|\boldsymbol{\phi}\rangle^{-1/2}$, satisfying the relations $\langle \boldsymbol{\varphi}|\boldsymbol{\varphi}\rangle = \mathbf{1}$ and $\boldsymbol{\varphi}_r = \boldsymbol{\varphi}$. Using (6.18), one obtains $\mathbf{T}^\dagger = \mathbf{T}^*$, i.e.,

$$\tilde{\mathbf{T}} = \mathbf{T}, \quad T_{kl} = T_{lk}, \tag{6.19}$$

and the matrix $\mathbf{T}$ is hence a *symmetric* matrix with complex elements. From the numerical point of view, this may be a simplification since one may have to store only the part of the matrix $\{T_{kl}\}$ having $k \leqslant l$. A symmetric matrix with real elements is self-adjoint, of course, and it may hence always be diagonalized. The same is true for a symmetric matrix with complex elements, if the eigenvalues λ_k are distinct. It should be observed, however, that—in the case of degenerate eigenvalues—it is necessary to apply the full theory of the linear operators and the description of the degeneracy in terms of Jordan blocks and Segré characteristics.

7. CONCLUDING REMARKS

In this paper we have concentrated our interest in the study of the stability problem for a pair of adjoint operators on such problems which are simple to hand and which may still lead to a deeper understanding of the more complicated general ones. It should be remembered, however, that even in these simple cases there is still a great deal of work to be done from the point of view of numerical analysis and actual computations. Most of our attention has, so far, been devoted to the study of the classical canonical form of operators defined of finite linear spaces and the associated projectors.

The treatment of general linear operators on an *infinite* space is a difficult problem and, even if many important results have been obtained by the mathematicians,[7] most of them are rather technical in nature and are hard to handle in the practical applications carried out by theoretical physicists and quantum chemists. Hence it seems desirable to go over these problems also from the practical point of view.

It should be observed that there is at least one series of results in this paper, which are easily generalized also to infinite spaces and which have already been successfully applied, in the treatment of the "constants of motion."[5] This depends on the fact that some of the results in Sec. 5, which were based on the reduced Cayley–Hamilton equation (5.86) to construct a resolution of the identity (5.78) in terms of product projection operators (5.79), may be applied also to *infinite spaces* as long as the number of eigenvalues λ_k stays finite and each one of them has a finite minimal index m_k corresponding to the largest Segré characteristic. In such a case, the product projection operators $O_k^{(1)}$ split the Hilbert space into a finite number of stable subspaces V_k—each one of infinite order $g_k = \infty$.

If the number of eigenvalues λ_k in the complex plane becomes infinite, the spectrum may become very complicated and may contain several continuous portions. However, as long as the minimal indices m_k remain finite, the problem can be handled at least in principle. Of course, one has convergence problems in treating the infinite products associated with $F_1(z)$ and $O_k^{(1)}(z)$, but these can be overcome by introducing the same convergence factors as occur in Weierstrass' and Mittag–Leffler's theorems about integer analytic functions. So far, little research has been done in this area.

In constructing the meromorphic function $1/F_1(z)$ having poles of order m_k in the points $z = \lambda_k$, one may again obtain some guidance from the case of a finite space. Starting from the generalized spectral resolution $T = |\mathbf{C}\rangle\boldsymbol{\lambda}\langle\mathbf{D}|$ expressed by the first relation (5.104), one obtains the following resolution of the resolvent:

$$R = (z\cdot\mathbf{1} - T)^{-1} = |\mathbf{C}\rangle(z\cdot\mathbf{1} - \boldsymbol{\lambda})^{-1}\langle\mathbf{D}|, \tag{7.1}$$

where $\boldsymbol{\lambda}$ is the classical canonical form which is block-diagonalized in terms of Jordan blocks of the type (5.36). Considering a specific diagonal block of the matrix $(z\cdot\mathbf{1} - \boldsymbol{\lambda})$ of order m_k associated with the eigenvalue λ_k, one obtains by using the matrices in Appendix A, Eq. (A2)

$$\begin{aligned}(z\cdot\mathbf{1} - \lambda_k\cdot\mathbf{1} - \mathbf{j}_1)^{-1} &= [(z-\lambda_k)\cdot\mathbf{1} - \mathbf{j}_1]^{-1} \\ &= (z-\lambda_k)^{-1}\cdot\mathbf{1} + (z-\lambda_k)^{-2}\mathbf{j}_1 \\ &+ (z-\lambda_k)^{-3}\mathbf{j}_2 + \cdots + (z-\lambda_k)^{-m_k}\mathbf{j}_{m_k-1},\end{aligned} \tag{7.2}$$

i.e., a finite expression containing exactly m_k terms. Taking two arbitrary elements φ_1 and φ_2 of the Hilbert space, one may instead consider the auxiliary function

$$\begin{aligned}\langle\varphi_1|R|\varphi_2\rangle &= \langle\varphi_1|(z\cdot\mathbf{1} - T)^{-1}|\varphi_2\rangle \\ &= \langle\varphi_1|\mathbf{C}\rangle(z\cdot\mathbf{1} - \boldsymbol{\lambda})^{-1}\langle\mathbf{D}|\varphi_2\rangle\end{aligned} \tag{7.3}$$

in the form of a binary product. In general this is a meromorphic function of the complex variable z with poles of the order m_k in the points $z = \lambda_k$, which relates it to the function $1/F_1(z)$. In such a case the coefficients of $(z-\lambda_k)^{-p}$ may be determined by carrying out the limiting procedure $z \to \lambda_k$ properly. In practice, it may be easier to handle this problem by means of the partitioning technique,[8] which

from many points of view seems well suited also for handling the case of infinite spaces. This problem will be further treated in a forthcoming paper.

ACKNOWLEDGMENTS

The author would like to thank the present members of the Uppsala and Florida groups for many valuable discussions. During the early 1960's, the author also had the privilege of discussing some of the problems treated in this paper with some visiting scientists; he is particularly indebted to Professor Harold McIntosh, Univ. Autonoma de Puebla, Puebla, Mexico, for valuable discussions about the stability problem in general, and to Professor John A. Coleman, Mathematics Department, Queens University, Kingston, Ontario, Canada, for a fine introduction into the theory of nilpotent operators and classical canonical forms.

APPENDIX A: DERIVATION OF THE INVERSE OF A TRIANGULAR MATRIX OF A CERTAIN TYPE

In studying the properties of a specific triangular matrix Δ of the form (5.64) having the properties

$$\Delta_{kl} = 0, \quad \text{for } k \geqslant l+1,$$
$$\Delta_{kl} = t_{k-l}, \quad \text{for } k \leqslant l, \tag{A1}$$

it is convenient to introduce the sequence of matrices of order m:

$$\mathbf{j}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdot \\ 0 & 0 & 1 & 0 & 0 & \cdot \\ 0 & 0 & 0 & 1 & 0 & \cdot \\ 0 & 0 & 0 & 0 & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \quad \mathbf{j}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & \cdot \\ 0 & 0 & 0 & 1 & 0 & \cdot \\ 0 & 0 & 0 & 0 & 1 & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$
$$\mathbf{j}_3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & \cdot \\ 0 & 0 & 0 & 0 & 1 & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \text{etc.} \tag{A2}$$

where $\mathbf{j}_p$ consists of a diagonal of 1's p steps above the main diagonal. One has directly the connections

$$\mathbf{j}_1^p = \mathbf{j}_p, \quad \text{for } p \leqslant m-1, \quad \mathbf{j}_1^m = \mathbf{0}. \tag{A3}$$

Starting from (5.64), one obtains the expansion

$$\Delta = t_0 \cdot \mathbf{1} + t_{-1} \cdot \mathbf{j}_1 + t_{-2}\mathbf{j}_2 + \cdots + t_{-(m-1)}\mathbf{j}_{m-1}, \tag{A4}$$

and this gives for the inverse

$$\mathbf{d} = \Delta^{-1} = t_0^{-1}\left\{\mathbf{1} + \sum_{k=1}^{m-1} (t_{-k}/t_0)\mathbf{j}_1{}^k\right\}^{-1}$$
$$= t_0^{-1}\left\{\mathbf{1} + \sum_{l=1}^{m-1} (-1)^l \left[\sum_{k=1}^{m-1} (t_{-k}/t_0)\mathbf{j}_1{}^k\right]^l\right\}. \tag{A5}$$

It is evident that the last expression may be rearranged into powers of the matrix $\mathbf{j}_1$, and hence this gives an expansion of the form

$$\mathbf{d} = d_0 \cdot \mathbf{1} + d_{-1} \cdot \mathbf{j}_1 + d_{-2}\mathbf{j}_2 + \cdots + d_{-(m-1)}\mathbf{j}_{m-1}. \tag{A6}$$

This means that $\mathbf{d}$ is also a triangular matrix having the properties

$$d_{kl} = 0, \quad \text{for } k \geqslant l+1,$$
$$d_{kl} = d_{k-l}, \quad \text{for } k \leqslant l. \tag{A7}$$

Once this result is established, it may be simpler to get the coefficients d_{-p} recursively by using the relation $\mathbf{d}\Delta = \mathbf{1}$, i.e.,

$$\mathbf{1} = \mathbf{d}\Delta = d_0 t_0 \cdot \mathbf{1} + (d_{-1}t_0 + d_0 t_{-1})\mathbf{j}_1$$
$$+ (d_{-2}t_0 + d_{-1}t_{-1} + d_0 t_{-2})\mathbf{j}_2 + \cdots, \tag{A8}$$

which gives

$$d_0 t_0 = 1,$$
$$d_{-1}t_0 + d_0 t_{-1} = 0,$$
$$d_{-2}t_0 + d_{-1}t_{-1} + d_0 t_{-2} = O. \tag{A9}$$

The essential property of the inverse $\mathbf{d} = \Delta^{-1}$ used in Sec. 5 is given by relation (A7).

[1]See, e.g., P. O. Löwdin, Int. J. Quantum Chem. **12**, 197 (1977), particularly p. 235.
[2]See, e.g., P. O. Löwdin, *Advances in Quantum Chemistry* (Academic, New York, 1980), Vol. 12, p. 263 (particularly pp. 298–306); Phys. Scripta **21**, 229 (1980).
[3]For general references on "complex scaling," see special workshop issue in Int. J. Quantum Chem. **14**, 343–542 (1978).
[4]E. A. Hylleraas and B. Undheim, Z. Phys. **65**, 759 (1930).
[5]P. O. Löwdin, Phys. Rev. **27**, 1509 (1955); Rev. Mod. Phys. **34**, 80 (1962); **34**, 520 (1962); **36**, 966 (1964).
[6]P. O. Löwdin, Int. J. Quantum Chem. **12**, 197 (1977); **21**, 275 (1982).
[7]See, e.g., N. Dunford and J. T. Schwartz, *Linear Operators, Part III* (Wiley-Interscience, New York, 1971).
[8]See Ref. 2.

On the Hartree–Fock scheme for a pair of adjoint operators

Piotr Froelich and Per-Olov Löwdin
Department of Quantum Chemistry, Uppsala University, Box 518, S-751 20, Uppsala, Sweden and Quantum Theory Project, University of Florida, Gainesville, Florida 32611

(Received 21 December 1981; accepted for publication 21 April 1982)

A generalization of the Hartree–Fock scheme for an arbitrary linear operator—and its adjoint—is derived by using the bivariational principle. It is shown that, if the system operator in the transition value is approximated by two Slater determinants, it is determined by a projector ρ, which corresponds to a generalization of the conventional Fock–Dirac density matrix, but which is no longer self-adjoint. The effective one-particle operator then takes the same form as in the conventional theory. The solution of the stability problem for a pair of adjoint effective operators is finally discussed. Numerical applications are performed elsewhere.

PACS numbers: 03.65. – w

1. INTRODUCTION

In the quantum theory of matter, the Hartree–Fock scheme corresponding to the independent-particle model is one of the standard methods for deriving approximate eigenfunctions and eigenvalues to self-adjoint operators—particularly to the Hamiltonian. During the last decades, there has also been an increasing interest in *general linear operators* T—defined on a Hilbert space $\mathfrak{H} = \{f\}$ having a binary product $\langle f|g\rangle$—which are neither self-adjoint nor normal. They are of interest as a mathematical tool in the theory as well as in the physical applications, for instance, in the partioning technique[1] or in the complex scaling method.[2]

Such an operator T having the domain $D(T)$ has an adjoint $T^\dagger$—with the domain $D(T^\dagger)$—defined through the relation

$$\langle f|Tg\rangle = \langle T^\dagger f|g\rangle \tag{1.1}$$

for every pair (f, g) belonging to the proper domains $D(T^\dagger)$ and $D(T)$, respectively. The stability problem for such a pair of adjoint operators—T and $T^\dagger$—has been discussed in another paper.[3] For the sake of simplicity, we will here concentrate our interest on the eigenvalue problem, which takes the form

$$TC_k = \lambda_k C_k, \quad T^\dagger D_l = \mu_l D_l, \tag{1.2}$$

where $\mu_l = \lambda_l^*$. According to the general theory, one has further the biorthogonality theorem: $\langle D_l|C_k\rangle = 0$ for $\mu_l \neq \lambda_k^*$, which may be combined with a normalization condition $\langle D_k|C_k\rangle = 1$ to the *biorthonormality* relation

$$\langle D_l|C_k\rangle = \delta_{lk}. \tag{1.3}$$

The question is now whether the Hartree–Fock method may be extended also to a general linear many-particle operator T of the form

$$T = T_{(0)} + \sum_i T_i + \sum_{i<j} T_{ij} + \cdots, \tag{1.4}$$

i.e., whether one can approximate the eigenfunctions C_k and D_l in (1.2) by *single Slater determinants* built up from one-particle functions or spin orbitals in a meaningful way. In connection with the complex-scaling method, one of us (P.F.) worked out a spin-orbital treatment of this problem based on the bivariational principle. However, once the solution was found, it turned out that a much simpler and more transparent derivation could be obtained in terms of density matrices. Only this derivation will be described here.

2. OPERATORS T WITH DISTINCT PURE POINT SPECTRA

The bivariational principle

In the study of general linear operators T, the bivariation principle[3] plays a fundamental role. One starts by considering a so-called "transition value":

$$\langle T\rangle_{12} = \frac{\langle \Phi_2|T|\Phi_1\rangle}{\langle \Phi_2|\Phi_1\rangle} = \mathrm{Tr}\, T\Gamma, \tag{2.1}$$

where Φ_1 and Φ_2 are arbitrary trial wave functions having the property $\langle \Phi_2|\Phi_1\rangle \neq 0$. In Dirac's nomenclature,[4] the binary product or the bracket $\langle f|g\rangle$ is the product of a bra-vector $\langle f|$ and a ket-vector $|g\rangle$. In the following, we will often use ket–bra operators $\omega = |g\rangle\langle f|$ defined through the relations

$$\omega = |g\rangle\langle f|, \quad \omega h = g\langle f|h\rangle, \tag{2.2}$$

for all elements h in $\mathfrak{H}$. They satisfy the reduced characteristic equation $\omega[\omega - \langle f|g\rangle \cdot 1] = 0$, and they have the properties

$$\omega^\dagger = |f\rangle\langle g|, \ \mathrm{Tr}\,\omega = \langle f|g\rangle. \tag{2.3}$$

For the system operator Γ in (2.1), one obtains in this formalism the explicit expression

$$\Gamma = \frac{|\Phi_1\rangle\langle\Phi_2|}{\langle\Phi_2|\Phi_1\rangle}. \tag{2.4}$$

It is then evident that the operator Γ has the properties

$$\Gamma^2 = \Gamma, \ \mathrm{Tr}\,\Gamma = 1, \ \Gamma \neq \Gamma^\dagger, \tag{2.5}$$

i.e., Γ is a one-dimensional projector which is not self-adjoint unless Φ_1 and Φ_2 are proportional, i.e., are related by a complex constant. The ranges of Γ and $\Gamma^\dagger$ are the one-dimensional linear manifolds $\{\Phi_1 \cdot \alpha\}$ and $\{\Phi_2 \cdot \beta\}$, respectively.

The *bivariational principle*[3] says that the first-order variation of the transition value (2.1) is vanishing:

 0022-2488/83/010088-09$02.50

$$\delta\langle T\rangle_{12} = 0, \tag{2.6}$$

if Φ_1 and Φ_2 are varied around the eigenfunctions C_k and D_k, and vice versa, which means that the relation (2.6) is equivalent to the two eigenvalue relations (1.2). In the more general case—also including degenerate eigenvalues—the bivariational principle (2.6) is equivalent to the commutation relation

$$T\Gamma = \Gamma T, \tag{2.7}$$

where Γ is a projector of the same order as the degeneracy. For the sake of simplicity, we will consider here only nondegenerate eigenvalues.

Construction of the system operator Γ from Slater determinants

In the Hartree–Fock scheme, the trial functions Φ_1 and Φ_2 are assumed to be single Slater determinants built up from one-particle functions, i.e.,

$$\Phi_1(X) = (N!)^{-1/2}|\psi_k(x_i)|,$$
$$\Phi_2(X) = (N!)^{-1/2}|\varphi_l(x_j)|, \tag{2.8}$$

where $X = (x_1, x_2, \ldots, x_N)$ and $x_i = (r_i, \zeta_i)$ is the combined space-spin coordinate of the ith particle, whereas the indices i and j go from 1 to N. The determinants are built from linearly independent one-particle functions or spinorbitals:

$$\psi = \{\psi_k\} = \{\psi_1, \psi_2, \ldots, \psi_N\},$$
$$\varphi = \{\varphi_l\} = \{\varphi_1, \varphi_2, \ldots, \varphi_N\}, \tag{2.9}$$

where the indices k and l go from 1 to N. The two sets ψ and φ span the linear manifolds $\mathcal{M}_\psi$ and $\mathcal{M}_\varphi$, respectively. We note that, if the sets ψ and φ undergo nonsingular linear transformations:

$$\psi' = \psi\alpha, \quad \varphi' = \varphi\beta, \tag{2.10}$$

the determinants (2.8) are changed only by the constant factors $|\alpha|$ and $|\beta|$, respectively, as in the ordinary Hartree–Fock scheme.

The derivations in the following follow closely similar derivations given by one of the authors in a study of the ordinary Hartree-Fock scheme formulated in terms of density matrices.[5] One of the key quantities in the bivariational principle is the overlap integral $\langle\Phi_2|\Phi_1\rangle$ between the Slater determinants (2.8). Using the antisymmetric projector

$$O_{AS} = \frac{1}{N!}\sum_P (-1)^p P, \tag{2.11}$$

having the properties $O_{AS}^\dagger = O_{AS}$ and $O_{AS}^2 = O_{AS}$, and introducing the *overlap integral*

$$d_{lk} = \langle\varphi_l|\psi_k\rangle = \int \varphi_l^*(x_1)\psi_k(x_1)dx_1, \tag{2.12}$$

one obtains in the standard way

$$\begin{aligned}\langle\Phi_2|\Phi_1\rangle &= (N!)\langle O_{AS}\varphi_1(x_1)\cdots\varphi_N(x_N)|O_{AS}\psi_1(x_1)\cdots\psi_N(x_N)\rangle \\ &= (N!)\langle\varphi_1(x_1)\cdots\varphi_N(x_N)|O_{AS}|\psi_1(x_1)\cdots\psi_N(x_N)\rangle \\ &= \sum_P(-1)^p\int\varphi_1^*(x_1)\cdots \\ &\quad\times\varphi_N^*(x_N)P_x\psi_1(x_1)\cdots\psi_N(x_N)dx_1\cdots dx_N \\ &= \sum_P(-1)^p\int\varphi_1^*(x_1)\cdots \\ &\quad\times\varphi_N^*(x_N)P_k^{-1}\psi_{k_1}(x_1)\cdots\psi_{k_N}(x_N)dx_1\cdots dx_N \\ &= \sum_P(-1)^p P_k^{-1}\langle\varphi_1|\psi_{k_1}\rangle\langle\varphi_2|\psi_{k_2}\rangle\cdots\langle\varphi_N|\psi_{k_N}\rangle \\ &= \sum_P(-1)^p P_k^{-1}d_{1k_1}d_{2k_2}\cdots d_{Nk_N} = |d_{lk}| = |\mathbf{d}|,\end{aligned} \tag{2.13}$$

where $|\mathbf{d}|$ is the determinant of the overlap matrix $\mathbf{d} = \{d_{lk}\}$. Here and in the following, we will use the same notation $\langle\ |\ \rangle$ for the binary product in the N-particle space and in the one-particle space, if there is no risk of misunderstanding. One should only remember that the symbol $\langle\ |\ \rangle$ indicates that one should integrate over the space coordinates and sum over the spin coordinates.

Since the bivariational principle requests the condition $\langle\Phi_2|\Phi_1\rangle = |\mathbf{d}| \neq 0$, the matrix $\mathbf{d}$ must be nonsingular. We note that this condition is fulfilled provided that there is no element in M_φ (except the zero element) which is orthogonal to the entire space M_ψ.[3]

Using the law of determinant multiplication, $|\mathbf{A}\cdot\mathbf{B}| = |\mathbf{A}|\cdot|\mathbf{B}|$ and $|\mathbf{C}|^{-1} = |\mathbf{C}^{-1}|$, one can now easily find an expression for the system operator Γ defined by (2.4) and (2.8):

$$\begin{aligned}\Gamma &= \frac{|\Phi_1\rangle\langle\Phi_2|}{\langle\Phi_2|\Phi_1\rangle} = \frac{1}{N!}\frac{||\psi\rangle|\cdot|\langle\varphi||}{|\mathbf{d}|} \\ &= \frac{1}{N!}||\psi\rangle|\cdot|\mathbf{d}^{-1}||\langle\varphi|| \\ &= \frac{1}{N!}||\psi\rangle\mathbf{d}^{-1}\langle\varphi|| = \frac{1}{N!}|\rho|,\end{aligned} \tag{2.14}$$

where

$$\rho = |\psi\rangle\mathbf{d}^{-1}\langle\varphi|. \tag{2.15}$$

This operator derivation in terms of "bold symbols" is very short and condensed—perhaps too condensed. If one instead considers the associated kernel or "density matrix" $\Gamma(X|X')$ one obtains similarly

$$\begin{aligned}\Gamma(X|X') &= \frac{\Phi_1(X)\Phi_2^*(X')}{\langle\Phi_2|\Phi_1\rangle} \\ &= \frac{1}{N!}\frac{|\psi_k(x_i)|\cdot|\varphi_l^*(x_j')|}{|\mathbf{d}|} \\ &= \frac{1}{N!}|\psi_k(x_i)|\cdot|(\mathbf{d}^{-1})_{kl}|\cdot|\varphi_l^*(x_j')| \\ &= \frac{1}{N!}\left|\sum_{k,l=1}^N \psi_k(x_i)(\mathbf{d}^{-1})_{kl}\varphi_l^*(x_j')\right| \\ &= \frac{1}{N!}|\rho(x_i, x_j')|,\end{aligned} \tag{2.16}$$

where

$$\rho(x_i, x_j') = \sum_{k,l=1}^N \psi_k(x_i)(\mathbf{d}^{-1})_{kl}\varphi_l^*(x_j') \tag{2.17}$$

is the kernel of the one-particle operator ρ defined by (2.15). Using (2.15) and the fact that $\langle\varphi|\psi\rangle = \mathbf{d}$, one gets immedi-

ately the relations

$$\rho^2 = \rho, \quad \mathrm{Tr}\,\rho = N, \quad \rho \neq \rho^\dagger. \tag{2.18}$$

It is evident that ρ and $\rho^\dagger$ are projectors defined on the one-particle Hilbert space $\mathscr{H}_1$ having the ranges M_ψ and M_φ, respectively. Writing (2.15) and its adjoint in the form

$$\rho = |\psi\rangle\langle\varphi|\psi\rangle^{-1}\langle\varphi|, \tag{2.19}$$

$$\rho^\dagger = |\varphi\rangle\langle\psi|\varphi\rangle^{-1}\langle\psi|, \tag{2.20}$$

it is clear that the operators ρ and $\rho^\dagger$ are N-dimensional projectors of a general type discussed in a previous paper.[3] If the basic sets ψ and φ undergo nonsingular linear transformations according to (2.10), these projectors stay invariant:

$$\rho' = \rho, \quad (\rho^\dagger)' = \rho^\dagger. \tag{2.21}$$

Everything is hence essentially the same as in the ordinary Hartree–Fock scheme, except that the projector ρ is no longer self-adjoint.

As in the conventional theory,[5] it is now possible to derive the reduced density matrices $\Gamma(x_1, x_2, \ldots, x_p | x_1', x_2', \ldots, x_p')$ of lower order by successive trace formation:

$$\Gamma(x_1, x_2, \ldots, x_p | x_1', x_2', \ldots, x_p') = \binom{N}{p} \int \Gamma(x_1, x_2, \ldots, x_p, x_{p+1}, \ldots, x_N | x_1', x_2', \ldots, x_p', x_{p+1}, \ldots, x_N) dx_{p+1} \ldots dx_N. \tag{2.22}$$

Starting from the density matrix of order N given by (2.16), i.e.,

$$\Gamma(x_1, x_2, \ldots, x_N | x_1', x_2', \cdots, x_N') = (1/N!)|\rho(x_i, x_j')|, \tag{2.23}$$

expanding the determinant of order N in the right-hand member after its last column, putting $x_N' = x_N$, integrating over x_N, using (2.18), and multiplying by the factor N according to (2.22), one obtains the reduced density matrix of order $(N-1)$. Repeating this procedure according to formula (2.22), one gets for the reduced density matrix of order p

$$\Gamma(x_1, \ldots, x_p | x_1', \ldots, x_p') = \frac{1}{p!}|\rho(x_i, x_j')|, \tag{2.24}$$

where the determinant is of order p and the indices i and j go from 1 to p. For $p = 2$, one gets particularly

$$\Gamma(x_1, x_2 | x_1' x_2') = \frac{1}{2}\begin{vmatrix} \rho(x_1, x_1') & \rho(x_1, x_2') \\ \rho(x_2, x_1') & \rho(x_2, x_2') \end{vmatrix} = \tfrac{1}{2}(1 - P_{12})\rho(x_1, x_1')\rho(x_2, x_2'), \tag{2.25}$$

where P_{12} is the exchange operator which changes x_1 into x_2 and x_2 into x_1, so that $P_{12} f(x_1, x_2) = f(x_2, x_1)$. For $p = 1$, one gets finally

$$\Gamma(x_1 | x_1') = \rho(x_1, x_1'). \tag{2.26}$$

All the reduced density matrices are hence determined by the one-particle projector ρ, and this fact renders a great simplification of the structure of the theory.

Using (1.4) and the theory of reduced density matrices,[5] one gets further for the transition value (2.1) of the operator Γ

$$\langle T \rangle_{12} = T_{(0)} + \int T_1 \Gamma(x_1 | x_1')\, dx_1 + \int T_{12} \Gamma(x_1 x_2 | x_1' x_2')\, dx_1\, dx_2 = T_{(0)} + \int T_1 \rho(x_1, x_1')\, dx_1 + \frac{1}{2}\int \overline{T}_{12} \rho(x_1, x_1')\rho(x_2, x_2')\, dx_1\, dx_2, \tag{2.27}$$

where

$$\overline{T}_{12} = T_{12}(1 - P_{12}). \tag{2.28}$$

We have further used the standard convention that the operators T_1, T_{12}, $\overline{T}_{12}$, $\cdots$ work only on the unprimed coordinates and that one puts $x_1' = x_1$, $x_2' = x_2$ before the integration.

Derivation of the bivariational Hartree–Fock equations

In applying the bivariational principle (2.6), one should now vary the trial wave functions Φ_1, and Φ_2 defined by (2.8). This may be accomplished by varying the sets $\psi = \{\psi_k\}$ and $\varphi = \{\varphi_l\}$ subject to the condition $|\mathbf{d}| = |\langle\varphi|\psi\rangle| \neq 0$, or simply by varying the associated density operator ρ given by (2.15) which satisfies automatically the conditions $\rho^2 = \rho$, $\mathrm{Tr}\,\rho = N$ according to (2.18). For the variations $\delta\rho$ this gives the auxiliary conditions

$$\delta\rho = \rho\,\delta\rho + \delta\rho\cdot\rho; \quad \mathrm{Tr}\,(\delta\rho) = 0. \tag{2.29}$$

Multiplying the first relation on the left (or on the right) by ρ, one obtains directly

$$\rho\cdot\delta\rho\cdot\rho = 0, \tag{2.30}$$

which means that the projection of $\delta\rho$ within the subspace of ρ must necessarily be vanishing. This implies also that the second relation (2.29) follows from the first, since one has

$$\mathrm{Tr}\,\delta\rho = \mathrm{Tr}\,(\rho\cdot\delta\rho + \delta\rho\cdot\rho) = \mathrm{Tr}(\rho^2\cdot\delta\rho + \delta\rho\cdot\rho^2) = 2\,\mathrm{Tr}(\rho\cdot\delta\rho\cdot\rho) = 0. \tag{2.31}$$

This could be expected, since the relation $\rho = \rho^2$ implies that $\mathrm{Tr}\,\rho$ must be an integer which cannot be continuously changed from one value to another. Using (2.27), one now obtains

$$\delta\langle T\rangle_{12} = \int T_1\,\delta\rho(x_1, x_1')\, dx_1 + \frac{1}{2}\int \overline{T}_{12}[\delta\rho(x_1, x_1')\rho(x_2, x_2') + \rho(x_1, x_1')\delta\rho(x_2, x_2')]\, dx_1\, dx_2 = \int\{T_1 + \int \overline{T}_{12}\rho(x_2, x_2')\, dx_2\}\delta\rho(x_1, x_1')\, dx_1 = \int T_{\mathrm{eff}}(1)\delta\rho(x_1, x_1'), dx_1 = \mathrm{Tr}\,T_{\mathrm{eff}}\delta\rho = 0, \tag{2.32}$$

where

$$T_{\mathrm{eff}}(1) = T_1 + \int dx_2 \overline{T}_{12}\rho(x_2, x_2'). \tag{2.33}$$

In the derivation, we have used the fact that the integration variables x_1 and x_2 are "dummy variables", the names of

which may be interchanged. We note that $T_{\rm eff}(1)$ is an *effective one-particle operator* which depends only on ρ and which is hence invariant under the linear transformations (2.10). So far, everything is analogous with the conventional Hartree–Fock theory,[5] except that—since $\rho \neq \rho^\dagger$—the effective one-particle operator $T_{\rm eff}(1)$ is no longer self-adjoint:

$$T_{\rm eff}(1) \neq T_{\rm eff}^\dagger(1). \tag{2.34}$$

Let us now study the meaning of the relation

$$\delta\langle T\rangle_{12} = \mathrm{Tr}\, T_{\rm eff}\delta\rho = 0. \tag{2.35}$$

Using (2.29) and (2.30), one obtains

$$\rho\cdot\delta\rho\cdot\rho = 0, \quad (1-\rho)\cdot\delta\rho\cdot(1-\rho) = 0, \tag{2.36}$$

which means that $\delta\rho$ has vanishing components in the subspaces defined by the projectors ρ and $\rho' = 1-\rho$, respectively. It is hence convenient to study a variation of the form

$$\delta\rho = (1-\rho)\Lambda_1\rho + \rho\Lambda_2(1-\rho), \tag{2.37}$$

where Λ_1 and Λ_2 are two general linear variational operators, which can be made arbitrarily small. It is interesting to observe that, in this case, the two relations (2.29) are automatically fulfilled, i.e., one has

$$\rho\,\delta\rho + \delta\rho\cdot\rho = \delta\rho, \quad \mathrm{Tr}\,\delta\rho = 0, \tag{2.38}$$

and that (2.37) is obviously the most general form one can give the variation $\delta\rho$. Substituting (2.37) into (2.35), one obtains directly

$$\begin{aligned}\mathrm{Tr}\, T_{\rm eff}\delta\rho &= \mathrm{Tr}\{T_{\rm eff}(1-\rho)\Lambda_1\rho \\ &\quad + T_{\rm eff}\rho\Lambda_2(1-\rho)\} \\ &= \mathrm{Tr}\{\rho T_{\rm eff}(1-\rho)\Lambda_1 + (1-\rho)T_{\rm eff}\rho\Lambda_2\} = 0,\end{aligned} \tag{2.39}$$

for arbitrary variational operators Λ_1 and Λ_2. Such a relation is valid if and only if

$$\rho T_{\rm eff}(1-\rho) = 0, \quad (1-\rho)T_{\rm eff}\,\rho = 0, \tag{2.40}$$

i.e.,

$$\rho T_{\rm eff} = \rho T_{\rm eff}\rho, \quad T_{\rm eff}\rho = \rho T_{\rm eff}\rho, \tag{2.41}$$

which implies that

$$T_{\rm eff}\rho = \rho T_{\rm eff}. \tag{2.42}$$

This relation is hence the necessary and sufficient condition for the fulfillment of the variational principle in determinantal approximation. This condition implies that the projector ρ decomposes the effective operator $T_{\rm eff}$ but, since this operator according to (2.33) depends on ρ, one is evidently faced with a nonlinear problem. Taking the adjoint of the operator relation (2.42), one obtains further

$$\rho^\dagger T_{\rm eff}^\dagger = T_{\rm eff}^\dagger\rho^\dagger, \tag{2.43}$$

which means that the adjoint projector $\rho^\dagger$ reduces the adjoint effective operator $T_{\rm eff}^\dagger$.

For the projectors ρ and $\rho^\dagger$, we will now use the expressions (2.19) and (2.20). Observing that $\rho\psi = \psi$ and $\rho^\dagger\varphi = \varphi$, one may now write the relations (2.42) and (2.43) in the form

$$T_{\rm eff}\psi = \rho T_{\rm eff}\psi = |\psi\rangle\langle\varphi|\psi\rangle^{-1}\langle\varphi|T_{\rm eff}|\psi\rangle, \tag{2.44}$$

$$T_{\rm eff}^\dagger\varphi = \rho^\dagger T_{\rm eff}^\dagger\varphi = |\varphi\rangle\langle\psi|\varphi\rangle^{-1}\langle\psi|T_{\rm eff}^\dagger|\varphi\rangle. \tag{2.45}$$

Introducing the notations

$$\mathbf{h}_1 = \langle\varphi|\psi\rangle^{-1}\langle\varphi|T_{\rm eff}|\psi\rangle, \tag{2.46}$$

$$\mathbf{h}_2 = \langle\psi|\varphi\rangle^{-1}\langle\psi|T_{\rm eff}^\dagger|\varphi\rangle, \tag{2.47}$$

one may hence write the conditions for the fulfillment of the bivariational principle (2.6) in the special form

$$T_{\rm eff}\psi = \psi\mathbf{h}_1, \quad T_{\rm eff}^\dagger\varphi = \varphi\mathbf{h}_2. \tag{2.48}$$

They are generalizations of the conventional Hartree–Fock equations, where $\mathbf{h}_1$ and $\mathbf{h}_2$ are the matrices formed by the "Lagrangian multipliers". They represent, of course, the stability relations indicating that the sets ψ and φ are stable under the operators $T_{\rm eff}$ and $T_{\rm eff}^\dagger$, respectively.

Structure of the Hartree–Fock equations

It is evident from the derivation that, if the two relations (2.48) are to have any meaning, it is necessary to treat the two operators $T_{\rm eff}$ and $T_{\rm eff}^\dagger$ independently of each other. In this subsection we will show, however, that there is still a certain coupling between the two equations and their solutions.

Let us first start by considering the first relation (2.48). According to elementary matrix theory, there always exists a similarity transformation $\mathbf{s}_1$ which will bring the matrix $\mathbf{h}_1$ to *classical canonical form* $\boldsymbol{\lambda}_1$:

$$\mathbf{s}_1^{-1}\mathbf{h}_1\,\mathbf{s}_1 = \boldsymbol{\lambda}_1, \quad \mathbf{h}_1 = \mathbf{s}_1\,\boldsymbol{\lambda}_1\,\mathbf{s}_1^{-1}. \tag{2.49}$$

Introducing the *canonical spin orbitals* ψ' through the transformation

$$\psi' = \psi\,\mathbf{s}_1 \tag{2.50}$$

one gets directly $T_{\rm eff}\psi' = T_{\rm eff}\psi\,\mathbf{s}_1 = \psi\,\mathbf{h}_1\,\mathbf{s}_1 = \psi'\,\mathbf{s}_1^{-1}\mathbf{h}_1\,\mathbf{s}_1 = \psi'\,\boldsymbol{\lambda}_1$, i.e.,

$$T_{\rm eff}\psi' = \psi'\,\boldsymbol{\lambda}_1, \tag{2.51}$$

which is a generalization of the canonical Hartree-Fock equations. It is important to observe that $T_{\rm eff}$ is invariant under the transformation (2.50). If the diagonal elements in $\boldsymbol{\lambda}_1$ are distinct, the classical canonical form is necessarily diagonal, but—if they are degenerate—there may be Jordan blocks of order $m = 2$ and higher which are ultimately described by the so-called Segré characteristics, depending on the fact that the operator $T_{\rm eff}$ is neither self-adjoint nor normal.

According to (2.46)–(2.49), one gets for the matrix $\mathbf{h}_2$

$$\begin{aligned}\mathbf{h}_2 &= \langle\psi|\varphi\rangle^{-1}\langle\psi|T_{\rm eff}^\dagger|\varphi\rangle \\ &= \langle\psi|\varphi\rangle^{-1}\langle T_{\rm eff}\psi|\varphi\rangle \\ &= \langle\psi|\varphi\rangle^{-1}\langle\psi\,\mathbf{h}_1|\varphi\rangle \\ &= \langle\psi|\varphi\rangle^{-1}\mathbf{h}_1^\dagger\langle\psi|\varphi\rangle \\ &= \langle\psi|\varphi\rangle^{-1}(\mathbf{s}_1^\dagger)^{-1}\,\boldsymbol{\lambda}_1^\dagger\,\mathbf{s}_1^\dagger\langle\psi|\varphi\rangle,\end{aligned} \tag{2.52}$$

and this implies that $\mathbf{h}_2$ may be brought to the special form $\boldsymbol{\lambda}^\dagger$ by means of a similarity transformation:

$$\mathbf{s}_2 = \langle\psi|\varphi\rangle^{-1}(\mathbf{s}_1^\dagger)^{-1} = \{\mathbf{s}_1^\dagger\langle\psi|\varphi\rangle\}^{-1} = \langle\psi'|\varphi\rangle^{-1}. \tag{2.53}$$

It should be observed that $\boldsymbol{\lambda}_1^\dagger$ is not a proper classical canonical form—since it consists of Jordan blocks of so-called second type—but that it can be brought to such a form by permuting the basic elements associated with each Jordan

block. However, for our purpose, the special form $\lambda^\dagger$ is more convenient. Introducing the functions

$$\varphi' = \varphi \mathbf{s}_2 = \varphi \langle \psi | \varphi \rangle^{-1} (\mathbf{s}_1^\dagger)^{-1}, \tag{2.54}$$

the second relation (2.48) takes the form

$$T_{\text{eff}}^\dagger \varphi' = \varphi' \lambda_1^\dagger. \tag{2.55}$$

This also leads to the relation

$$\begin{aligned} \langle \varphi' | \psi' \rangle &= \mathbf{s}_2^\dagger \langle \varphi | \psi \rangle \mathbf{s}_1 \\ &= \mathbf{s}_1^{-1} \langle \varphi | \psi \rangle^{-1} \langle \varphi | \psi \rangle \mathbf{s}_1 = \mathbf{1}, \end{aligned} \tag{2.56}$$

which shows that the "canonical"solutions φ' and ψ' are automatically going to be *biorthonormal* once $\mathbf{s}_1$ has been evaluated and $\mathbf{s}_2$ has been determined through the relation (2.53).

3. SOLUTION OF THE HARTREE–FOCK EQUATIONS BY MEANS OF EXPANSION METHODS

In the self-adjoint case, the Hartree–Fock equations are often conveniently solved by expansion methods.[6] For the treatment of a pair of adjoint operators, T_{eff} and $T_{\text{eff}}^\dagger$, we will here use a generalization of a special technique developed previously for the conventional case.[5]

Writing the Hartree–Fock equation (2.51) and (2.55) in the form

$$T_{\text{eff}} \psi = \psi \lambda_1, \quad T_{\text{eff}}^\dagger \varphi = \varphi \lambda_1^\dagger, \tag{3.1}$$

we observe that—even in the case of degenerate eigenvalues—they represent stability problems of the type treated in a previous paper.[3] According to the general theory, such stability problems are solved by looking for a pair of projectors, Q and $Q^\dagger$, which decompose the operators T_{eff} and $T_{\text{eff}}^\dagger$, respectively, so that

$$T_{\text{eff}} Q = Q T_{\text{eff}}, \quad T_{\text{eff}}^\dagger Q^\dagger = Q^\dagger T_{\text{eff}}^\dagger, \tag{3.2}$$

where the second relation is the adjoint of the first. Once such projectors are determined, they should further be decomposed into *irreducible components*, which leads to a complete solution of the stability problem (3.1). The relations (3.2) are again equivalent with the bivariational principle (2.6) for the effective operators, and this fact is particularly useful in deriving approximate solutions.

In order to treat this problem in greater detail, we will now introduce two sets $\Phi = \{\phi_1, \phi_2, \ldots, \phi_M\}$ and $\Psi = \{\psi_1, \psi_2, \ldots, \psi_M\}$, each consisting of M linearly independent elements, where $N \ll M$. We will further assume that there is no element in the subspace spanned by Ψ which is orthogonal to all the elements in Φ, which means that the matrix

$$\Delta = \langle \Phi | \Psi \rangle \tag{3.3}$$

should be nonsingular. In principle, it should be possible to let $M \to \infty$, and to make the two sets Φ and Ψ complete. Under such assumptions, there exists[3] a pair of adjoint projectors

$$Q = |\Psi\rangle \langle \Phi | \Psi \rangle^{-1} \langle \Phi |, \tag{3.4}$$

$$Q^\dagger = |\Phi\rangle \langle \Psi | \Phi \rangle^{-1} \langle \Psi |, \tag{3.5}$$

having the property $\text{Tr}\, Q = \text{Tr}\, Q^\dagger = M$. Forming the difference

$$\omega = T_{\text{eff}} Q - Q T_{\text{eff}}, \tag{3.6}$$

we realize that the quantity,

$$\gamma = \text{Tr}\, \omega^\dagger \omega \geqslant 0, \tag{3.7}$$

is nonnegative and zero only when $\omega = 0$, i.e., when Q decomposes the operator T_{eff}. The quantity (3.7), which is a generalization of the concept of the "width" of an operator, is hence a convenient measure of the degree of approximation introduced into the theory by the choice of the two sets Φ and Ψ.

The two projectors (3.4) and (3.5) are, of course, invariant under linear transformations of the type

$$\Phi' = \Phi \alpha, \quad \Psi' = \Psi \beta, \tag{3.8}$$

where α and β are nonsingular matrices of order $M \times M$. Introducing the reciprocal set

$$\Phi_r = \Phi \langle \Psi | \Phi \rangle^{-1}, \tag{3.9}$$

which has the biorthonormality property

$$\langle \Psi | \Phi_r \rangle = \langle \Phi_r | \Psi \rangle = \mathbf{1}, \tag{3.10}$$

one gets particularly

$$Q = |\Psi\rangle \langle \Phi_r |, \quad Q^\dagger = |\Phi_r\rangle \langle \Psi |. \tag{3.11}$$

We will now assume that the relations (3.2) are approximately valid for the projectors Q and $Q^\dagger$, so that

$$T_{\text{eff}} Q \approx Q T_{\text{eff}}, \quad T_{\text{eff}}^\dagger Q^\dagger \approx Q^\dagger T_{\text{eff}}^\dagger \tag{3.12}$$

or

$$T_{\text{eff}} |\Psi\rangle \langle \Phi_r | \approx |\Psi\rangle \langle \Phi_r | T_{\text{eff}}, \tag{3.13}$$

$$T_{\text{eff}}^\dagger |\Phi_r\rangle \langle \Psi | \approx |\Phi_r\rangle \langle \Psi | T_{\text{eff}}^\dagger. \tag{3.14}$$

Multiplying the first relation on the right by $|\Psi\rangle$ and the second on the right by $|\Phi_r|\rangle$, one obtains

$$T_{\text{eff}} \Psi \approx \Psi \langle \Phi_r | T_{\text{eff}} | \Psi \rangle = \Psi \mathbf{t}_1, \tag{3.15}$$

$$T_{\text{eff}}^\dagger \Phi_r \approx \Phi_r \langle \Psi | T_{\text{eff}}^\dagger | \Phi_r \rangle = \Phi_r \mathbf{t}_2, \tag{3.16}$$

where $\mathbf{t}_1 = \langle \Phi_r | T_{\text{eff}} | \Psi \rangle$ and $\mathbf{t}_2 = \langle \Psi | T_{\text{eff}}^\dagger | \Phi_r \rangle$ $= \langle T_{\text{eff}} \Psi | \Phi_r \rangle = \langle \Phi_r | T_{\text{eff}} \Psi \rangle^\dagger = \mathbf{t}_1^\dagger$.

Hence, one has the approximate relations

$$T_{\text{eff}} \Psi \approx \Psi \mathbf{t}_1, \quad T_{\text{eff}} \Phi_r \approx \Phi_r \mathbf{t}_2, \tag{3.17}$$

which become exact only when the subspaces spanned by Ψ and Φ_r are stable under the operators T_{eff} and $T_{\text{eff}}^\dagger$, respectively. A more detailed investigation shows, however, that these approximations may be justified by the bivariational principle, and we will return to this problem in a later subsection.

In order to decompose the projector Q into its irreducible components, we will now transform the matrix $\mathbf{t}_1$ of order $M \times M$ to classical canonical form Λ_1 by means of the similarity transformation $\mathbf{S}_1$:

$$\mathbf{S}_1^{-1} \mathbf{t}_1 \mathbf{S}_1 = \Lambda_1, \quad \mathbf{t}_1 = \mathbf{S}_1 \Lambda_1 \mathbf{S}_1^{-1}. \tag{3.18}$$

Taking the adjoint relations, one obtains

$$\mathbf{S}_1^\dagger \mathbf{t}_1^\dagger (\mathbf{S}_1^\dagger)^{-1} = \Lambda_1^\dagger, \mathbf{t}_1^\dagger = (\mathbf{S}_1^\dagger)^{-1} \Lambda_1^\dagger \mathbf{S}_1^\dagger, \tag{3.19}$$

which show that $\mathbf{t}_2 = \mathbf{t}_1^\dagger$ is transformed to pseudoclassical-canonical form $\Lambda_1^\dagger$ by the similarity transformation $\mathbf{S}_2 = (\mathbf{S}_1^\dagger)^{-1}$. Introducing the *canonical orbitals*

$$\Psi' = \Psi \mathbf{S}_1, \quad \Phi' = \Phi_r \mathbf{S}_2 = \Phi_r (\mathbf{S}_1^\dagger)^{-1}, \tag{3.20}$$

one obtains

$$T_{\text{eff}}\Psi' = \Psi'\Lambda_1, \quad T_{\text{eff}}^{\dagger}\Phi' = \Phi'\Lambda_1^{\dagger}. \tag{3.21}$$

We note also that the canonical orbitals have the biorthonormality property

$$\langle\Psi'|\Phi'\rangle = \mathbf{1}, \quad \langle\Phi'|\Psi'\rangle = \mathbf{1}. \tag{3.22}$$

This means that, for the projectors Q and $Q^{\dagger}$, one obtains the decompositions

$$Q = |\Psi'\rangle\langle\Phi'| = \sum_{k=1}^{M} |\Psi'_k\rangle\langle\Phi'_k| = \sum_{k=1}^{M} O_k, \tag{3.23}$$

$$Q^{\dagger} = |\Phi'\rangle\langle\Psi'| = \sum_{k=1}^{M} |\Phi'_k\rangle\langle\Psi'_k| = \sum_{k=1}^{M} O_k^{\dagger}, \tag{3.24}$$

where the operators

$$O_k = |\Psi'_k\rangle\langle\Phi'_k|, \quad O_k^{\dagger} = |\Phi'_k\rangle\langle\Psi'_k| \tag{3.25}$$

are one-dimensional projectors, which are (exactly or approximately) stable under the effective operators T_{eff} and $T_{\text{eff}}^{\dagger}$, respectively, only when they are associated with distinct eigenvalues or with degenerate eigenvalues having simple Jordan blocks of order $m = 1$. If a Jordan block (J_K) is of higher order, there are several indices k associated with the block, and one obtains instead for the irreducible projectors Q_K and $Q_K^{\dagger}$ corresponding to the block

$$Q_K = \sum_{k}^{(J_k)} O_k, \quad Q_K^{\dagger} = \sum_{k}^{(J_k)} O_k^{\dagger}, \tag{3.26}$$

where one sums over all indices coupled to J_K. Instead of (3.23) and (3.24), one obtains the following decompositions into irreducible projectors

$$Q = \sum_{K} Q_K, \quad Q^{\dagger} = \sum_{K} Q_K^{\dagger}, \tag{3.27}$$

where the summation goes over all Jordan blocks.

Selection of the essential solutions

So far, we have neglected the fact that the effective operator T_{eff} according to (2.33) depends on the projector ρ:

$$T_{\text{eff}}(1) = T_1 + \int dx_2\, T_{12}(1 - P_{12})\rho(x_2, x_2'), \tag{3.28}$$

which means that, in reality, one is dealing with a *nonlinear problem*. In conventional Hartree–Fock theory, this difficulty is circumvented by the well-known self-consistent-field (SCF) procedure which is an iterative procedure of the first order. Extending this approach to the more general case treated here, one would start from an initial approximation $\rho^{(0)}$ to the projector ρ. Solving the Hartree–Fock equations (3.1) for the associated effective operator $T_{\text{eff}}^{(0)}$ formed according to (3.28), one then obtains a new approximation $\rho^{(1)}$ to the projector ρ which may serve as a new starting point. In this way, one obtains a series of approximations $\rho^{(0)}, \rho^{(1)}, \rho^{(2)}, \rho^{(3)}, \cdots$ which are defined through the cycle

$$\rho^{(n)} \rightarrow T_{\text{eff}}^{(n)} \rightarrow \rho^{(n+1)}. \tag{3.29}$$

Under favorable circumstances, this interaction procedure is convergent and leads to self-consistent solutions. It should be observed, however, that even if the process is divergent, it is often possible to use it to construct modified iteration procedures which are convergent. Similarly, one may often convert slowly convergent procedures into rapidly convergent ones.

There is one problem in this procedure which should now be discussed in greater detail. In solving the Hartree–Fock equations (3.21), one obtains a total of M canonical solutions, Ψ' and Φ', which then should be divided into two groups: the N *essential* spin orbitals Ψ'_{ess} and Φ'_{ess} entering into the next approximation of the projectors ρ and $\rho^{\dagger}$ to be denoted by q and q', respectively,

$$q = |\Psi'_{\text{ess}}\rangle\langle\Phi'_{\text{ess}}|, \quad q^{\dagger} = |\Phi'_{\text{ess}}\rangle\langle\Psi'_{\text{ess}}|, \tag{3.30}$$

and the $(M\text{-}N)$ *virtual* spin orbitals, which have no direct physical significance but which may still be mathematically useful in the treatment of the original eigenvalue problem (1.2).

In the conventional Hartree–Fock method for studying the ground state of an atomic, molecular, or solid-state system, one intuitively selects the essential spin orbitals by taking the canonical solutions associated with the lowest one-particle energies. Even if this is physically reasonable and corresponds to the famous "Aufbau-principle," there is—as far as we know—no mathematical proof that this is the correct way to carry out the iteration procedure.

In the general case, the situation is more complicated—partly due to the fact that the one-particle eigenvalues may now be complex and without special ordering. In this case, however, our problem is more mathematical than physical, and we may concentrate our interest on the iteration cycle (3.29).

Starting from the projector $\rho^{(n)}$ and its adjoint, we will now consider the nth step of (3.29) where the solution of the Hartree-Fock equations (3.21) leads to the canonical solutions

$$\Psi'^{(n+1)} = \{\psi_{\mu}'^{(n+1)}\}, \quad \Phi'^{(n+1)} = \{\Phi_{\nu}'^{(n+1)}\} \tag{3.31}$$

for $\mu, \nu = 1, 2, 3, \ldots, M$. For the sake of brevity, we will in the following leave out the superscripts n and $(n+1)$, if there is no risk for misunderstanding. It is now convenient to consider a canonical solution Ψ'_k as *essential* if it is mainly situated within the subspace of ρ so that—at least approximately—one has

$$\rho\Psi'_k = \Psi'_k. \tag{3.32}$$

If, on the other hand, the solution Ψ'_{μ} is mainly situated outside the subspace of ρ, so that $\rho\Psi'_{\mu} \approx 0$, the solution Ψ'_{μ} is considered as *virtual*. In the physical interpretation of this scheme, the essential spin orbitals are considered as occupied by particles, whereas the virtual spin orbitals are unoccupied. In order to treat this classification systematically, one may study the numbers

$$m_{\mu} = \|(1-\rho)\Psi'_{\mu}\|^2 = \langle\Psi'_{\mu}|1 - \rho - \rho^{\dagger} + \rho^{\dagger}\rho\,|\Psi'_{\mu}\rangle \geqslant 0, \tag{3.33}$$

for $\mu = 1, 2, 3, \ldots, M$ and normalized solutions Ψ'_{μ}. The in-

dices k for the essential solutions are then found by selecting the N smallest numbers out of the sequence $m = \{m_1, m_2,...,m_M\}$, where one should also remember the rule that solutions associated with the same Jordan block should always belong together. If anyone of these numbers is vanishing, the relation (3.32) is—of course—exactly fulfilled. Once an essential solution Ψ'_k has been determined, the associated solution Φ'_k is automatically given by the pairing condition contained in the biorthonormality relation (2.56), or

$$\langle \Psi'_k | \Phi'_l \rangle = \delta_{kl}, \tag{3.34}$$

since there is only one solution Φ'_l which is not orthogonal to Ψ'_k. Using the essential solutions, one can now form the new projectors q and $q^\dagger$ of order $(n+1)$ according to (3.30), and repeat the entire procedure. The iteration process becomes *self-consistent* whenever the new projectors of order $(n+1)$ agree within the accuracy desired with the old projectors of order n, and the process has become convergent whenever $q = \rho$. In such a case, the trivial relations

$$q\Psi'_{\rm ess} = \Psi'_{\rm ess}, \quad q^\dagger \Phi'_{\rm ess} = \Phi'_{\rm ess} \tag{3.35}$$

will imply the existence of the relations $\rho\Psi'_{\rm ess} = \Psi'_{\rm ess}$ and $\rho^\dagger \Phi'_{\rm ess} = \Phi'_{\rm ess}$, or

$$\rho\Psi'_k = \Psi'_k, \quad \rho^\dagger \Phi'_k = \Phi'_k \tag{3.36}$$

for all the essential solutions. Conversely, the existence of the two relations (3.36) indicate that $q = \rho$ and that the iteration process has converged.

The proof for this statement is based on the fact that by combining (3.30), (3.35), and (3.36) one gets directly the operator relations

$$\rho q = q, \quad \rho^\dagger q^\dagger = q^\dagger, \quad q\rho = q. \tag{3.37}$$

Considering the difference $\Delta = \rho - q$, one gets further $\Delta^2 = \rho^2 + q^2 - \rho q - q\rho = \rho - q = \Delta$, i.e., Δ is an idempotent which may be diagonalized having only the eigenvalues 0 and 1. However, since $\mathrm{Tr}\,\Delta = 0$, this implies that $\Delta = 0$, i.e., that $q = \rho$.

It is evident that it would be interesting to study the convergence properties of the iteration cycle (3.29) and particularly its connection with the numbers m_μ defined by (3.33) and with other similar quantities in greater detail, but such studies are considered outside the framework of the present paper and will be reserved for later communications.

Reformulation of the theory by means of the charge- and bond-order matrix

For the practical computer applications, it is convenient to reformulate the theory in a slightly different form. As before, we will start from two linearly independent sets Ψ and Φ of order M, and we will further use the reciprocal set $\Phi_r = \Phi\langle\Psi|\Phi\rangle^{-1}$ characterized by the relations (3.9) and (3.10). Let us now introduce the two sets of *essential canonical solutions* to the Hartree–Fock equations (3.21) through the formulas

$$\psi' = \Psi\mathbf{c}, \quad \varphi' = \Phi_r\mathbf{d}, \tag{3.38}$$

where $\mathbf{c}$ and $\mathbf{d}$ are rectangular matrices of order $M \times N$. Because of the biorthonormality property (3.22), one has

$$\langle\varphi'|\psi'\rangle = \mathbf{d}^\dagger\langle\Phi_r|\Psi\rangle\mathbf{c} = \mathbf{d}^\dagger\mathbf{c} = \mathbf{1}, \tag{3.39}$$

where the right-hand member is a unit matrix of order $N \times N$. For the projector ρ, one gets further

$$\begin{aligned}\rho &= |\psi'\rangle\langle\varphi'|\psi'\rangle^{-1}\langle\varphi'| = |\psi'\rangle\langle\varphi'| \\ &= |\Psi\rangle\mathbf{c}\,\mathbf{d}^\dagger\langle\Phi_r| = |\Psi\rangle\mathbf{R}\langle\Phi_r|,\end{aligned} \tag{3.40}$$

where

$$\mathbf{R} = \mathbf{c}\,\mathbf{d}^\dagger \tag{3.41}$$

is a matrix of order $N \times N$, which is a generalization of the well-known charge- and bond-order matrix.[7] Since $\mathbf{d}^\dagger\mathbf{c} = \mathbf{1}$, it is evident that it satisfies the relations

$$\mathbf{R}^2 = \mathbf{R}, \quad \mathrm{Tr}\,\mathbf{R} = N, \quad \mathbf{R} \neq \mathbf{R}^\dagger, \tag{3.42}$$

i.e., that $\mathbf{R}$ is an idempotent matrix of order N. Since $\mathbf{R}$ is the representation of ρ in the bases chosen, it will sometimes be referred to as the *projector matrix*.

It is now easy to express the effective one-particle operator $T_{\rm eff}(1)$ defined by (2.33) in terms of the matrix $\mathbf{R}$. Using (3.40), one gets directly

$$\begin{aligned}T_{\rm eff}(1) &= T_1 + \int dx_2 \overline{T}_{12}\rho(x_2, x'_2) \\ &= T_1 + \int dx_2 \overline{T}_{12}|\Psi(x_2)\rangle\mathbf{R}\langle\Phi_r(x'_2)| \\ &= T_1 + \sum_{\kappa\lambda} R_{\kappa\lambda}\, dx_2 \Phi^*_{r,\lambda}(x_2)\overline{T}_{12}\Psi_\kappa(x_2),\end{aligned} \tag{3.43}$$

where $\overline{T}_{12} = T_{12}(1 - P_{12})$. In accordance with (3.17) the matrix $\mathbf{t}$ which forms the starting point for the block-diagonalization procedure (3.18) is then given by the expression

$$\mathbf{t} = \langle\Phi_r(1)|T_{\rm eff}(1)|\Psi(1)\rangle, \tag{3.44}$$

and hence it has the matrix elements

$$\begin{aligned}t_{\nu\mu} &= \langle\nu|T_{\rm eff}(1)|\mu\rangle, \\ &= \int \Phi^*_{r,\nu}(x_1)T_{\rm eff}(1)\Phi_\mu(x_1)dx_1 \\ &= \langle\nu|T_1|\mu\rangle + \sum_{\nu,\lambda} R_{\kappa\lambda}\langle\nu\lambda|\overline{T}_{12}|\mu\kappa\rangle \\ &= \langle\nu|T_1|\mu\rangle + \sum_{\kappa,\mu} R_{\kappa\lambda}[\langle\nu\lambda|T_{12}|\mu\kappa\rangle \\ &\quad - \langle\nu\lambda|T_{12}|\kappa\mu\rangle],\end{aligned} \tag{3.45}$$

where the indices ν and λ refer to functions out of the set Φ_r, and the indices μ and κ to functions out of the set Ψ. When one block-diagonalizes the matrix $\mathbf{t}$ by means of the similarity transformation $\mathbf{S}$, so that

$$\mathbf{S}^{-1}\mathbf{t}\,\mathbf{S} = \mathbf{\Lambda}, \quad \mathbf{t}\,\mathbf{S} = \mathbf{S}\,\mathbf{\Lambda}, \tag{3.46}$$

one gets, of course, M solutions to the stability problem expressed by the second relation (3.46) and they are represented by the M column vectors $\mathbf{c}_k$ of the matrix $\mathbf{S}$. The essential solutions are then given by the vectors $\mathbf{c}_k$ satisfying—exactly or approximately—the relation

$$\mathbf{R}\mathbf{c}_k = \mathbf{c}_k, \tag{3.47}$$

analogous to (3.32), and the N essential solutions $\mathbf{c}_k$ form then the rectangular matrix $\mathbf{c}$ of order $M \times N$. According to (3.19), the matrix $\mathbf{t}^\dagger$ is further block-diagonalized by the similarity transformation $\mathbf{S}_2 = (\mathbf{S}_1^{-1})^\dagger$, and the M solutions $\mathbf{d}_l$ to the stability problem of $\mathbf{t}^\dagger$ are then given by the column vectors of $\mathbf{S}_2$. To every vector $\mathbf{c}_k$ there is one and only one

vector $\mathbf{d}_k$ having the property

$$\mathbf{d}_k^\dagger \mathbf{c}_k = 1, \tag{3.48}$$

since all the other vectors $\mathbf{d}_l$ are automatically orthogonal to $\mathbf{c}_k$. The N vectors $\mathbf{d}_k$ associated with the essential solutions $\mathbf{c}_k$ then form the rectangular matrix $\mathbf{d}$ of order $M \times N$, and one can then recompute the matrix $\mathbf{R} = \mathbf{c}\mathbf{d}^\dagger$. This means that, instead of (3.29), one obtains the iteration cycle

$$\mathbf{R} \rightarrow \mathbf{t} \rightarrow \mathbf{S} \text{ and } \mathbf{S}^{-1} \rightarrow \mathbf{c} \text{ and } \mathbf{d} \rightarrow \mathbf{R} = \mathbf{c}\mathbf{d}^\dagger, \tag{3.49}$$

which forms the basis for the computational procedure.

In this scheme, the matrix $\mathbf{t}$ is evaluated from $\mathbf{R}$ by using formula (3.45) for the matrix elements, the similarity transformation $\mathbf{S}$ may be found by using standard algebraic procedures for diagonalizing and block-diagonalizing matrices, whereas the essential solutions $\mathbf{c}_k$ and $\mathbf{d}_k$ are selected according to (3.47)—which is at least approximately fulfilled—and (3.48), respectively.

It should be observed that the transformation matrix $\mathbf{S}$ is by no means unique, and that any nonsingular matrix $\mathbf{S}$ of the type

$$\overline{\mathbf{S}} = f(\mathbf{t})\mathbf{S} = \mathbf{S} f(\Lambda) \tag{3.50}$$

will also perform the block-diagonalization. For practical purposes, one may remember the rule that all the column vectors of $\mathbf{S}$ associated with one and the same Jordan block may be multiplied by the same constant, and that it may be convenient to choose this constant so that the first vector of the block becomes normalized to unity. The formula $\mathbf{R} = \mathbf{c}\mathbf{d}^\dagger$ is, of course, invariant under such a transformation.

Evaluation of the transition value

It now remains to evaluate the transition value $\langle T \rangle_{12}$, which forms the basis for the bivariational principle, in terms of the matrix $\mathbf{R}$. In applying formula (2.27), one should remember that the operators T_1 and $\overline{T}_{12} = T_{12}(1 - P_{12})$ work only on the unprimed coordinates x_1 and x_2 and that one has to put $x_1' = x_1$ and $x_2' = x_2$ before the integrations—consisting of integrations over the space coordinates and summations over the spin coordinates—are carried out. Introducing the expression (3.40) for the projector ρ into (2.27), one obtains

$$\begin{aligned}
\langle T \rangle_{12} &= T_{(0)} + \int T_1 \rho(x_1, x_1')\, dx_1 \\
&\quad + \frac{1}{2} \int \overline{T}_{12} \rho(x_1, x_1') \rho(x_2, x_2')\, dx_1\, dx_2 \\
&= T_{(0)} + \int T_1 \Psi(x_1) \mathbf{R} \Phi_r^*(x_1')\, dx_1 \\
&\quad + \frac{1}{2} \int \overline{T}_{12} \{ \Psi(x_1) \mathbf{R} \Phi_r^*(x_1') \} \\
&\quad \times \{ \Psi(x_2) \mathbf{R} \Phi_r^*(x_2') \}\, dx_1\, dx_2 \\
&= T_{(0)} + \sum_{\mu,\nu} R_{\mu,\nu} \langle \nu | T_1 | \mu \rangle \\
&\quad + \frac{1}{2} \sum_{\mu,\nu} \sum_{\kappa,\lambda} R_{\kappa\lambda} R_{\mu\nu} \langle \nu\lambda \,|\overline{T}_{12}| \mu\kappa \rangle \\
&= T_{(0)} + \sum_{\mu,\nu} R_{\mu,\nu} \langle \nu | T_1 | \mu \rangle \\
&\quad + \frac{1}{2} \sum_{\mu,\nu} \sum_{\kappa,\lambda} R_{\kappa\lambda} R_{\mu\nu} \{ \langle \nu\lambda \,|T_{12}| \mu\kappa \rangle \\
&\quad - \langle \nu\lambda \,|T_{12}| \kappa\mu \rangle \},
\end{aligned} \tag{3.51}$$

where we have used the same matrix notations as in (3.45). Using the expression (3.45), one obtains further

$$\begin{aligned}
\langle T \rangle_{12} &= T_{(0)} + \sum_{\mu\nu} R_{\mu\nu} t_{\nu\mu} \\
&\quad - \frac{1}{2} \sum_{\mu\nu} \sum_{\kappa\lambda} R_{\mu\nu} R_{\kappa\lambda} \langle \nu\lambda \,|\overline{T}_{12}| \mu\kappa \rangle,
\end{aligned} \tag{3.52}$$

where the last term occurs to compensate for the fact that all the two-particle interactions would otherwise be counted twice. For the second term one obtains, according to (3.41), (3.18), and (3.39), that

$$\begin{aligned}
\sum_{\mu\nu} R_{\mu\nu} t_{\nu\mu} &= \mathrm{Tr}(\mathbf{t}\mathbf{R}) = \mathrm{Tr}(\mathbf{t}\mathbf{c}\mathbf{d}^\dagger) \\
&= \mathrm{Tr}(\mathbf{c}\,\lambda\mathbf{d}^\dagger) = \mathrm{Tr}(\lambda\,\mathbf{d}^\dagger\mathbf{c}) = \mathrm{Tr}\,\lambda = \sum_k^{(n)} \lambda_k,
\end{aligned} \tag{3.53}$$

where one sums over the eigenvalues λ_k of all the essential solutions. This gives the expression

$$\begin{aligned}
\langle T \rangle_{12} &= T_{(0)} + \sum_{k=1}^{N} \lambda_k \\
&\quad - \frac{1}{2} \sum_{\mu\nu} \sum_{\kappa\lambda} R_{\mu\nu} R_{\kappa\lambda} \{ \langle \nu\lambda \,|T_{12}| \mu\kappa \rangle \\
&\quad - \langle \nu\lambda \,|T_{12}| \kappa\mu \rangle \}.
\end{aligned} \tag{3.54}$$

Finally, one may take the average of the relations (3.51) and (3.54), which gives the comparatively simple formula

$$\begin{aligned}
\langle T \rangle_{12} &= T_{(0)} + \frac{1}{2} \sum_{k=1}^{N} \lambda_k \\
&\quad + \frac{1}{2} \sum_{\mu\nu} R_{\mu\nu} \langle \nu | T_1 | \mu \rangle.
\end{aligned} \tag{3.55}$$

It should be observed, however, that—since the stability equation $\mathbf{t}\mathbf{c} = \mathbf{c}\,\lambda$ is very seldom exactly fulfilled—the expression (3.51) for $\langle T \rangle_{12}$ is more fundamental than the relations (3.54) and (3.55).

In a previous subsection, it was pointed out that the relations (3.12)—(3.17) were only approximately valid and that these steps in our derivation may be justified by a second application of the bivariational principle. In connection with the expansion methods, however, a simple alternative approach is also possible, which will be discussed below.

Let us assume that the original trial wave functions Φ_1 and Φ_2 are approximated by Slater determinants of one-particle functions $\psi = \{\psi_k\}$ and $\varphi = \{\varphi_l\}$—where $k, l = 1, 2, 3, \ldots, N$—which in turn are built up by expansion methods from two one-particle sets Ψ and Φ of order M, respectively, so that

$$\psi = \Psi\mathbf{c}, \quad \varphi = \Phi_r \mathbf{d}, \tag{3.56}$$

where $\Phi_r = \Phi\langle\Psi|\Phi\rangle^{-1}$. If M is finite, our starting point is hence more limited than before. In such a case, the fundamental projector ρ defined by (2.15) takes the special form

$$\begin{aligned}
\rho &= |\psi\rangle\langle\varphi|\psi\rangle^{-1}\langle\varphi| \\
&= |\Psi\rangle\mathbf{c}(\mathbf{d}^\dagger\mathbf{c})^{-1}\mathbf{d}^\dagger\langle\Phi_r| \\
&= |\Psi\rangle\mathbf{R}\langle\Phi_r|,
\end{aligned} \tag{3.57}$$

where

$$\mathbf{R} = \mathbf{c}(\mathbf{d}^\dagger \mathbf{c})^{-1}\mathbf{d}^\dagger \tag{3.58}$$

is a matrix of order $M \times M$ which automatically satisfies the relations

$$\mathbf{R}^2 = \mathbf{R}, \quad \mathrm{Tr}\,\mathbf{R} = N, \quad \mathbf{R} \neq \mathbf{R}^\dagger, \tag{3.59}$$

which are identical to (3.42). The form (3.58) reduces to the form (3.41) if the sets ψ and φ are chosen biorthonormal, which is always possible by using the transformed set $\varphi' = \varphi\langle\psi|\varphi\rangle^{-1}$. For the moment, it may be preferable to use the more general form (3.58).

Substituting the expression (3.57) for ρ into the relation (2.27), one obtains the formula (3.51), which now forms the starting point for the bivariational principle $\delta\langle T\rangle_{12} = 0$. In this case, one may vary the rectangular matrices $\mathbf{c}$ and $\mathbf{d}$ of order $M \times N$ or even better the charge- and bond-order matrix $\mathbf{R}$ of order $M \times M$ subject to the constraints (3.59). In carrying out the details of this procedure it is not surprising that one recovers the formulas of the previous subsections, but now in a more exact form. In this approach it is hence sufficient to use the bivariational principle only once.

In varying the coefficients $R_{\mu\nu}$ and $R_{\kappa\lambda}$ in formula (3.54), one should observe that, since the last term is quadratic in these quantities, the same contribution will be obtained twice from this term. Using (3.45), one gets directly

$$\delta\langle T\rangle_{12} = \sum_{\mu\nu}\delta R_{\mu\nu}\cdot t_{\nu\mu} = \mathrm{Tr}(\delta\mathbf{R}\cdot\mathbf{t}). \tag{3.60}$$

Using the same reasoning as in the relations (2.32)–(2.42), one finds that the necessary and sufficient condition for the fulfillment of the bivariational principle is that the projector matrix $\mathbf{R}$ decomposes the matrix $\mathbf{t}$, i.e.,

$$\mathbf{t}\,\mathbf{R} = \mathbf{R}\,\mathbf{t} \tag{3.61}$$

This problem is then solved by finding the similarity transformation $\mathbf{S}$ which brings the matrix $\mathbf{t}$ to classical canonical form $\mathbf{\Lambda}$:

$$\mathbf{S}^{-1}\,\mathbf{t}\,\mathbf{S} = \mathbf{\Lambda}, \quad \mathbf{t}\,\mathbf{S} = \mathbf{S}\,\mathbf{\Lambda}. \tag{3.62}$$

The rectangular matrices $\mathbf{c}$ and $\mathbf{d}$ of order $M \times N$ are then found by selecting the N essential solutions $\mathbf{c}_k$ and $\mathbf{d}_k$ out of the column vectors of $\mathbf{S}$ and $(\mathbf{S}^{-1})^\dagger$, which means that the relation $\mathbf{d}^\dagger\mathbf{c} = 1_n$ is automatically fulfilled. The new matrix $\mathbf{R}$ is then found according to the simple formula $\mathbf{R} = \mathbf{c}\,\mathbf{d}^\dagger$. The matrix $\mathbf{t}$ and the projector matrix $\mathbf{R}$ are hence the essential tools for solving the Hartree–Fock equations occurring in this type of problem. Numerical applications of this scheme are carried out in other publications.[8]

ACKNOWLEDGMENTS

The authors would like to thank the various members of the Uppsala and Florida groups for valuable discussions of this problem.

[1]For a survey of the partitioning technique see, e.g., P.O. Löwdin, Int. J. Quantum Chem. **2**, 867 (1968).
[2]For a survey of the complex-scaling method see, e.g., the proceedings from the 1978 Sanibel workshop published in Int. J. Quantum Chem. **14**, 343–542 (1978).
[3]P. O. Löwdin, "On the Stability Problem of a Pair of Adjoint Operators," Uppsala Technical Note Number 650 (1981).
[4]P. A. M. Dirac, *The Principle of Quantum Mechanics* (Clarendon, Oxford, 1958).
[5]P. O. Löwdin, Phys. Rev. **97**, 1474, 1490 (1955).
[6]C. A. Coulson, Proc. Cambridge Philos. Soc. **34**, 204 (1938); C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951); G. G. Hall, Proc. R. Soc. London Ser. A **208**, 328 (1951).
[7]C. A. Coulson and H. C. Longuet-Higgins, Proc. R. Soc. London Ser. A **191**, 39 (1947); **192**, 16 (1947); **193**, 447, 456, (1948); **195**, 188 (1948); see also P. O. Löwdin, Phys. Rev. **97**, 1490 (1955), particularly p. 1498.
[8]M. Mishra, Y. Öhrn, and P. Froelich, Phys. Lett. A **84**, 4 (1981).

When is the Wigner function of multi-dimensional systems nonnegative?

Francisco Soto[a] and Pierre Claverie
Laboratoire de Chimie Quantique, Institut de Biologie Physico-Chimique, 13, rue Pierre et Marie Curie, 75005 Paris, France

(Received 24 February 1981, accepted for publication 18 September 1981)

It is shown that, for systems with an arbitrary number of degrees of freedom, a necessary and sufficient condition for the Wigner function to be nonnegative is that the corresponding state wavefunction is the exponential of a quadratic form. This result generalizes the one obtained by Hudson [Rep. Math. Phys. **6**, 249 (1974)] for one-dimensional systems.

PACS numbers: 03.65.Bz, 03.65.Ca, 05.30.Ch

I. INTRODUCTION

The Wigner function of a system in a pure state with wavefunction $\Psi(\mathbf{q},t)$ is given by[1-3]

$$F(\mathbf{p},\mathbf{q},t)=\frac{1}{\hbar^{3n}}\int_{\mathbf{R}^n} d\mathbf{v}\,\Psi^*(\mathbf{q}+\tfrac{1}{2}\mathbf{v},t)e^{(i/\hbar)\mathbf{p}\cdot\mathbf{v}}\Psi(\mathbf{q}-\tfrac{1}{2}\mathbf{v},t), \tag{1}$$

where n is the dimensionality of the configuration space. It is well known[1-3] that this function has the properties of a probability distribution, with the exception that for some state wavefunctions it is not nonnegative.

Thus it is pertinent to ask the question, when is the Wigner function nonnegative? In the case of pure states of one-dimensional systems an answer was already given by Hudson[4] (a similar, but partly erroneous result was also published by Piquet[5]; for the sake of completeness this work is discussed in the appendix); he showed that the Wigner function is nonnegative if and only if the state wavefunction is a gaussian function:

$$\Psi(q)=\exp[-\tfrac{1}{2}(aq^2+2bq+c)], \tag{2}$$

where a, b are complex numbers with $\mathrm{Re}\,a>0$, and c is a normalization constant that can be taken as real[4]; in other words, the Wigner function is nonnegative if and only if the system is in a coherent state (see Ref. 6 for the coherent states).

In this paper we generalize this result to pure states of multidimensional systems. We show that a necessary and sufficient condition for the Wigner function to be nonnegative is that the state wavefunction is of the form

$$\Psi_{\mathbf{A},\mathbf{b}}(\mathbf{q})=\exp[-\tfrac{1}{2}(\mathbf{q}^+\mathbf{A}\mathbf{q}+2\mathbf{b}\cdot\mathbf{q}+c)], \tag{3}$$

where $\mathbb{A}$ is a symmetric complex matrix with $|\mathrm{Re}\,\mathbb{A}|>0$, $\mathbf{b}$ is a complex n-dimensional vector, c a real normalization constant, and $q=(q_1,\ldots,q_n)$.

Our proof follows the one given by Hudson[4] for the one-dimensional case. In actual fact, in Hudson's proof, only one step is not directly generalizable to several (complex) variables, namely that step where he utilizes the Hadamard factorization theorem that, to our knowledge, does not have a several variables version. However, as we shall see below, only a restricted version of this theorem is needed, and this restricted version can be generalized to the case of several variables.

[a]Research fellow from the UNAM (Universidad Nacional Autónoma de Mexico).

Thus the purpose of the present work is to prove this restricted version of Hadamard's theorem for several complex variables, and to substitute this theorem for the genuine Hadamard theorem in Hudson's proof, which is thus generalized to an arbitrary number of dimensions.

The structure of the paper is as follows: In Sec. II we give the proof of the sufficiency; this is trivial and it is given only for the sake of completeness. In Sec. III we reproduce for several variables the first part of Hudson's proof (i.e., the part which precedes the use of Hadamard's theorem). We give, in Sec. IV, the proof of the restricted Hadamard theorem for several complex variables. We finish the proof of the main theorem (about the Wigner function) in Sec. V.

II. THE SUFFICIENCY CONDITION

In order to find the Wigner function associated with the state wavefunction (3), we substitute it in (1) and we utilize the following result:

$$\int_{\mathbf{R}^n}\exp[-\tfrac{1}{2}\mathbf{x}^+\mathbb{B}\mathbf{x}+\boldsymbol{\tau}\cdot\mathbf{x}]\,d\mathbf{x}=\frac{(2\pi)^{n/2}}{|\mathbb{B}|^{1/2}}\exp\left[\tfrac{1}{2}\sum_{j=1}^{n}\frac{(\mathbf{f}_j\cdot\boldsymbol{\tau})^2}{\mu_j}\right], \tag{4}$$

where $|\mathbb{B}|>0$, $\mathbf{f}_j$ $(j=1,\ldots,n)$ are the eigenvectors of $\mathbb{B}$ and μ_j $(j=1,\ldots,n)$ the corresponding eigenvalues. We then find that the Wigner function associated with the wavefunction (3) is

$$F(\mathbf{p},\mathbf{q})=\frac{1}{\pi^{n/2}\hbar^n|\mathrm{Re}\,\mathbb{A}|^{1/2}}\exp\left\{-[\mathbf{q}^+\mathrm{Re}\,\mathbf{A}\mathbf{q}+2\mathrm{Re}\,\mathbf{b}\cdot\mathbf{q}+c]-\sum_{j=1}^{n}\frac{[\mathbf{e}_j\cdot(\mathbf{q}^+\mathrm{Im}\,\mathbb{A}+\mathrm{Im}\,\mathbf{b}+(1/\hbar)\mathbf{p})]^2}{\lambda_j}\right\}, \tag{5}$$

where $\mathbf{e}_j$ $(j=1,\ldots,n)$ are the eigenvectors of $\mathrm{Re}\,\mathbb{A}$ and λ_j $(j=1,\ldots,n)$ its eigenvalues.

Thus the Wigner function of a multidimensional Gaussian wavefunction is a multivariate Gaussian distribution, which is always nonnegative.

III. THE NECESSITY. FIRST PART

We want to find which is the set of wavefunctions that give a nonnegative Wigner function. Let us denote by Ω this set, and let us take an arbitrary element Ψ of it; we define in the complex space $\mathbb{C}^n$ the complex function $J(\mathbf{z})$ as

$$J(\mathbf{z})=e^{(1/2)c}\langle\Psi\,|\,\Psi_{1,\mathbf{z}}\rangle, \tag{6}$$

where $\Psi_{1,\mathbf{z}}$ corresponds to definition (3) with $\mathbb{A}=\mathbb{1}$, $\mathbf{b}=\mathbf{z}$,

 0022-2488/83/010097-04$02.50

and

$$\langle \Psi_1 | \Psi_2 \rangle = \int_{\mathbb{R}^n} d\mathbf{q} \Psi_1^*(\mathbf{q}) \Psi_2(\mathbf{q}). \tag{7}$$

This function $J(\mathbf{z})$ has the following properties:

(1) It is an entire function. This is evident from its definition (6) and from (3) and (7).

(2) It does not have zeros in $\mathbb{C}^n$.

To prove this we utilize the following property (see Refs. 4 and 7):

$$\int_{\mathbb{R}^{2n}} d\mathbf{p}\, d\mathbf{q}\, F^*_{\Psi_1}(\mathbf{p},\mathbf{q}) F_{\Psi_2}(\mathbf{p},\mathbf{q}) = \frac{1}{(2\pi)^n} |\langle \Psi_1 | \Psi_2 \rangle|^2. \tag{8}$$

Using it in the definition (6) we find

$$|J(\mathbf{z})|^2 = e^c (2\pi)^n \int_{\mathbb{R}^{2n}} d\mathbf{p}\, d\mathbf{q}\, F^*_{\Psi}(\mathbf{p},\mathbf{q}) F_{\Psi_{1,\mathbf{z}}}(\mathbf{p},\mathbf{q}), \tag{9}$$

and since $F_{\Psi_{1,\mathbf{z}}}(\mathbf{p},\mathbf{q}) > 0$ $\forall \mathbf{z} \in \mathbb{C}^n$ and $\Psi \in \Omega$ [therefore $F_{\Psi}(\mathbf{p},\mathbf{q}) \geqslant 0$], we have

$$|J(\mathbf{z})|^2 > 0 \quad \forall \mathbf{z} \in \mathbb{C}^n \tag{10}$$

which gives the desired conclusion.

(3) Its order of growth ρ (see Sec. 26 in Ref. 8) is at most two ($\rho \leqslant 2$). [Let us recall that the order of growth of a function $f(\mathbf{z})$ may be defined as $\rho = \lim_{R \to \infty} \ln \ln M(R) / \ln R$, with $M(R) = \sup_{|\mathbf{z}| = R} |f(\mathbf{z})|$.]

From (6) we have

$$|J(\mathbf{z})|^2 \leqslant e^c \|\Psi\|^2 \|\Psi_{1,\mathbf{z}}\|^2, \tag{11}$$

and using (4),

$$\|\Psi_{1,\mathbf{z}}\|^2 = e^{-c} (\pi)^{n/2} \exp[(\mathrm{Re}\, \mathbf{z})^2]. \tag{12}$$

Thus,

$$|J(\mathbf{z})|^2 \leqslant \pi^{n/2} \|\Psi\|^2 \exp[(\mathrm{Re}\, \mathbf{z})^2]. \tag{13}$$

Since the order of growth of $\exp[(\mathrm{Re}\, \mathbf{z})^2]$ is obviously two, we conclude that $\rho \leqslant 2$ for $[J(\mathbf{z})]^2$, and therefore also for $J(\mathbf{z})$.

We remark that this property implies that the order of growth of J as a function of only one of its variables (the other being fixed) is also at most two.

From these three properties we would like to find the explicit form of the function $J(\mathbf{z})$, which in fact, for $\mathbf{z} = i\mathbf{y}$, is the Fourier transform of $e^{-(1/2)\mathbf{q}^2} \Psi^*(\mathbf{q})$. At this point the problem of the generalization arises (until now we have followed Hudson's proof); in the one-dimensional case Hudson utilizes the

Hadamard factorization theorem[9]: If $f(z)$ is an entire function of order ρ with an m-fold zero at the origin, we have

$$f(z) = z^m e^{Q(z)} P(z), \tag{14}$$

where $Q(z)$ is a polynomial of degree $r \leqslant \rho$ and $P(z)$ is the canonical product (of genus s) formed with the zeros (other than $z = 0$) of $f(z)$.

With this theorem and Properties (1)–(3) we conclude (in the one-dimensional case) that $J(z)$ is of the form (14) *with z^m and $P(z)$ absent*,

$$J(z) = \exp[\alpha z^2 + \beta z + \gamma], \tag{15}$$

and thus that the Fourier transform of $e^{-(1/2)q^2} \Psi^*(q)$ is Gaussian, which is possible only if $\Psi(q)$ itself is Gaussian.

To our knowledge, a generalization of Hadamard's theorem to several variables does not exist, and one of the reasons is that, in this case, both zeros and poles are not isolated as in one dimension. But in actual fact, the essential part of Hudson's proof is that an entire function without zeros is the exponential of a polynomial; thus, when there is a function $f(z)$ with zeros, they are eliminated by dividing it by $z^m P(z)$, and then, using this fact, the expression (14) is obtained. As we said, in several variables it is the product $z^m P(z)$ which is not easily generalized.

Nevertheless, in order to find the expression of $J(z)$ a restricted version of the Hadamard theorem is sufficient, namely the version corresponding to a function without zeros. This restricted version may be more easily generalized, and this is done in the following section.

IV. THE RESTRICTED VERSION OF HADAMARD'S THEOREM

We are going to prove the following

Theorem: If $f(\mathbf{z})$ is an entire function, in the space of n complex variables $\mathbb{C}^n$, with order of growth ρ and without zeros, we have

$$f(\mathbf{z}) = e^{Q(\mathbf{z})}, \tag{16}$$

where $Q(\mathbf{z})$ is a polynomial of degree $r \leqslant \rho$.

Proof: We are going to give the explicit proof of this theorem for two variables only, the case of n variables being more cumbersome and without anything especially new.

Let us fix z_2 in $f(z_1,z_2)$; we thus have an entire function of z_1, with order of growth at most ρ and without zeros; by the Hadamard factorization theorem we can write it as $\exp[\sum_{j=1}^{r} \alpha_j z_1^j]$ with $r \leqslant \rho$ and where $\alpha_j (j = 1, \cdots)$ depends on z_2, and this is valid for all z_2 finite. Thus we have

$$f(z_1,z_2) = \exp\left[\sum_{j=1}^{r} \alpha_j(z_2) z_1^j\right] \tag{17}$$

in the set

$$\mathbb{D}_1 = \{(z_1,z_2): z_1 \in \mathbb{C},\ 0 \leqslant |z_2| \leqslant M_1\}. \tag{18}$$

Keeping now z_1 fixed, and proceeding in the same way, we conclude that

$$f(z_1,z_2) = \exp\left[\sum_{k=1}^{s} \beta_k(z_1) z_2^k\right] \tag{19}$$

in the set

$$\mathbb{D}_2 = \{(z_1,z_2): 0 \leqslant |z_1| \leqslant M_2,\ z_2 \in \mathbb{C}\} \tag{20}$$

and with $s \leqslant \rho$.

Since $f(z_1,z_2)$ is an entire function without zeros, we may define its logarithm $Q(z_1,z_2)$ as an entire one-valued function over $\mathbb{C}^2$, and, from (17) and (19) we deduce that

$$\sum_{j=1}^{r} \alpha_j(z_2) z_1^j = \sum_{k=1}^{s} \beta_k(z_1) z_2^k \tag{21}$$

in $\mathbb{D}_1 \cap \mathbb{D}_2$, because these two functions are just two expressions in $\mathbb{D}_1 \cap \mathbb{D}_2$ of one and the same function $Q(z_1,z_2) = \ln f(z_1,z_2)$.

Now, we differentiate (21) n times ($n \leqslant s$) with respect to z_2 [this is possible, because $Q(z_1,z_2)$ is an entire function], thus

obtaining

$$\sum_{j=1}^{r} \alpha_j^{(n)}(z_2) z_1^j = \sum_{k=n}^{s} \frac{k!}{(k-n)!} \beta_k(z_1) z_2^{k-n} \quad (n = 1,2,...,s), \tag{22}$$

and, taking $z_2 = 0$, we get

$$\beta_n(z_1) = \sum_{j=1}^{r} \frac{\alpha_j^{(n)}(0)}{n!} z_1^j \quad (n = 1,2,...,s) \tag{23}$$

for all z, such that $0 < |z_1| < M_2$. Thus,

$$Q(z_1,z_2) = \sum_{j=1}^{r} \sum_{k=1}^{s} \frac{\alpha_j^{(k)}(0)}{k!} z_1^j z_2^k \tag{24}$$

in $\mathbb{D}_1 \cap \mathbb{D}_2$. Denoting for brevity $\alpha_j^k(0)/k! = \gamma_{jk}$, we finally get the expression

$$\ln f(z_1,z_2) = Q(z_1,z_2) = \sum_{j=1}^{r} \sum_{k=1}^{s} \gamma_{jk} z_1^j z_2^k \tag{25}$$

in $\mathbb{D}_1 \cap \mathbb{D}_2$. In fact (25) is valid in every bounded set of $\mathbb{C}^2$ because M_1 and M_2 in (18) and (20) are arbitrary but finite.

Since $Q(z_1,z_2)$ is an entire function over $\mathbb{C}^2$, it is easily deduced by analytic continuation[10] that (25) and hence

$$f(z_1,z_2) = \exp\left[\sum_{j=1}^{r} \sum_{k=1}^{s} \gamma_{jk} z_1^j z_2^k\right] \tag{26}$$

are valid in the whole complex space $\mathbb{C}^2$.

It only remains to show that the degree of $Q(z_1,z_2)$ is at most ρ, i.e., that in (26) $\gamma_{jk} = 0$ for $j + k > \rho$. Supposing that this is not the case and taking $z_1 = z_2 = R$ (real) we easily conclude that the order of growth of $f(z_1,z_2)$ would be greater than ρ, in contradiction with the assumption made; thus $\gamma_{jk} = 0$ for $j + k > \rho$ and the proof is completed.

V. THE NECESSITY. SECOND PART

We start from the results obtained in the first part (Sec. III). We defined the function $J = \mathbb{C}^n \to \mathbb{C}$ such that

$$J(\mathbf{z}) = e^{(1/2)c} \langle \Psi \mid \Psi_{1,\mathbf{z}} \rangle \tag{27}$$

with $\Psi \in \Omega$, and we found that it has the following properties:

(1) It is an entire function.

(2) It does not have zeros in $\mathbb{C}^n$.

(3) Its order of growth is at most two.

Thus by the restricted version of Hadamard's theorem in several dimensions, we have

$$J(\mathbf{z}) = \exp\left\{ \sum_{\substack{j_1,...,j_n=1 \\ (j_1+\cdots+j_n=2)}}^{2} \alpha_{j_1,...,j_n} z_1^{j_1} \cdots z_n^{j_n} + \sum_{j=1}^{n} \beta_j z_j + \gamma \right\}. \tag{28}$$

Now

$$J(i\mathbf{y}) = \exp\left\{ - \sum_{\substack{j_1,...,j_n=1 \\ (j_1+\cdots+j_n=2)}}^{2} \alpha_{j_1,...,j_n} y_1^{j_1} \cdots y_n^{j_n} + i \sum_{j=1}^{n} \beta_j y_j + \gamma \right\}, \tag{29}$$

but from its definition (27) we also have

$$J(i\mathbf{y}) = \int_{\mathbf{R}^n} d\mathbf{q}\, \Psi^*(\mathbf{q})\, e^{-(1/2)q^2}\, e^{-i\mathbf{y}\cdot\mathbf{q}}, \tag{30}$$

i.e., the Fourier transform of $e^{-(1/2)q^2}\Psi^*(\mathbf{q})$ is a multidimensional Gaussian function, which implies that $e^{-(1/2)q^2}\Psi^*(\mathbf{q})$ itself is a Gaussian function, and Ω is the set of multivariate Gaussian functions, which is the desired result.

VI. CONCLUSION

The question concerning the nonnegative character of the Wigner function is therefore now completely settled for pure states: whatever the number of variables, only the Gaussian wavefunctions give rise to a nonnegative Wigner distribution (which is itself Gaussian, too). Of course, as mentioned by Hudson,[4] the study of this question for mixed (instead of pure) states remains an open problem.

ACKNOWLEDGMENTS

We express our thanks to L. Pesquera for helpful discussions concerning this work, to Dr. Hudson and the referee for valuable comments, and to Dr. Piquet for communicating to us his modified proof.[11]

APPENDIX: DISCUSSION OF PIQUET'S TREATMENT CONCERNING THE ONE-DIMENSIONAL CASE[5]

Piquet's proof is correct for a real wavefunction Ψ, except that the restricted form $\Psi = \exp[-(ax^2 + b)]$ with $a > 0$ and $b > 0$ is unnecessary[11] (see before Theorem 3.3 in Ref. 5; this paper is brief, but all the argument can be recasted in a detailed form[11]). However in the case of a complex wavefunction, his proof is incorrect. He considers $\Psi\Psi^*$ and applies to it the result for a real function; then he utilizes the Levy–Cramer Theorem[12] which states that if the product of two characteristic functions is Gaussian they are also Guassian, in order to conclude that Ψ is of the form (2). Nevertheless the application of the Levy–Cramer theorem is not legitimate because we do not know *a priori* whether Ψ and Ψ^* are characteristic functions, and in actual fact, it is easy to see that in general (2) cannot be considered as a characteristic function (we would like here to thank Dr. Hudson and the referee for drawing our attention to this point).

Finally, we want to mention that Dr. Piquet has modified his proof,[11] which is now correct and actually gives exactly the same conclusion as Hudson's proof. This modified proof remains essentially different from Hudson's one, in the sense that this proof does not make use of the Hadamard theorem but uses instead some properties of convex functions (by the way, in his modified proof, Piquet does no longer distinguishes the peculiar case where Ψ is real: the proof holds directly for complex Ψ).

[1] E. Wigner, Phys. Rev. **40**, 749–759 (1932).

[2] J. E. Moyal, Proc. Cambridge Philos. Soc. **45**, 99–124 (1949).

[3] S. De Groot, *La transformation de Weyl et la fonction de Wigner: une forme alternative de la mécanique quantique* (Les Presses de l'Université de Montréal, 1974).

[4] R. L. Hudson, Rep. Math. Phys. **6**, 249–252 (1974).

[5] C. Piquet, C. R. Acad. Sci. Paris A **279**, 107–109 (1974).

[6] R. J. Glauber, Phys. Rev. **131**, 2766–2788 (1963).

[7] J. C. T. Pool, J. Math. Phys. **7**, 66–76 (1966).

[8] B. A. Fuks, *Introduction to the Theory of Analytic Functions of Several*

Complex Variables (American Mathematical Society, Providence, R.I., 1963).
[9]R. P. Boas, *Entire Functions* (Academic, New York, 1954).
[10]H. Cartan, *Théories Elémentaire des Fonctions Analytiques d'une ou plusieurs variables complexes* (Hermann, Paris, 1963). English translation: *Elementary Theory of Analytic Functions of one or several Variables* (Addison Wesley, Reading, Mass.), See Sec. IV 2.3.
[11]C. Piquet (private communication).
[12]W. Feller, *Introduction to probability theory and its applications* (Wiley, New York, 1966), Chap. XV, Sec. 8, pp. 498–499.

A characteristic function approach to the discrete spectrum of electrically charged particles

Metin Demiralp
Applied Mathematics Department, Marmara Scientific and Industrial Research Institute, P. O. Box 141 Kadıköy, Istanbul, Turkey

(Received 20 January 1981; accepted for publication 11 September 1981)

The purpose of this article is to obtain a scalar function the zeros of which give the discrete energy values for a system of electrically charged particles. The relation between the serial expansion of the characteristic function in powers of the system's eigenvalue and the Stieltjes series has been revealed not only for the electrically charged particles but also generally for any positive (or negative) definite Hermitian operator which has only a discrete spectrum. The use of Padé approximants to express the characteristic function has offered a rapidly convergent scheme to evaluate the system's eigenvalues. The first few elements of the Padé Table for the reciprocal of the characteristic function of certain systems have been given to verify the presented idea numerically. The determination of these elements needs the values of certain complicated integrals which we name "Zeroth Order Hyperspherical Spectral Coefficients" [HSC(φ_0)]. The first two of these coefficients are investigated and their evaluation is realized analytically.

PACS numbers: 03.65.Ge

I. INTRODUCTION

Several authors have employed the space folding method together with some perturbational schemes in quantum mechanical calculations.[1,2] The main idea of this procedure is to convert the original eigenvalue equation into a scalar one. To this end, the solution space of the original equation is divided into a conveniently chosen space and complementary with the aid of some projection operations. After some intermediate steps one can arrive at a scalar equation for the determination of the original equations' eigenvalue. Homogenization of this equation gives a function, zeros of which are the desired eigenvalues of the operator under consideration. The serial representation of this function in powers of the eigenvalue parameter is needed to obtain an explicit structure. This expansion does, however, converge only in a restricted domain of the complex plane of eigenvalue parameter which does not cover all the eigenvalues of the operator under investigation. Fortunately we have a possibility of obtaining such a representation which is valid on the domain of all desired eigenvalues and offers a rapidly converging computational scheme. Indeed, rational approximations and the Padé Table built for them have this property. Recent years of science bear an increasing tendency to use Padé approximants in several problems of physics and chemistry.[3-7] The effectiveness of such rational approximations is that in most cases only a few approximants are sufficient to obtain a reasonable accuracy.

In the following sections we shall obtain the characteristic function for a positive definite operator having only a discrete spectrum and investigate its serial expansion in the sense of Stieltjes series. Some analytical evaluations and numerical calculations for certain systems will complete the present work.

II. DERIVATION OF THE CHARACTERISTIC FUNCTION

Consider a system which can be described by the following equation:

$$\Lambda f = \lambda W f, \tag{2.1}$$

where the Hermitian operators Λ, W, the scalar λ, and the function f characterize the structure and the behavior of this system. Λ and W are two operators on a Hilbert space, and may be matrices, integral operators, or differential operators with some compatible boundary conditions. As in most of the quantum mechanical problems we can assume that at least one of the operators Λ and W is positive definite and hence is invertible. We can also additionally assume that the Eq. (2.1) has only a discrete spectrum and its eigenfunctions form a complete basis set for the Hilbert space under consideration. On the other hand another assumption which states the boundedness of the operator $\Lambda^{-1/2} W \Lambda^{-1/2}$ is needed to prove the theorem of the next section.

Let us choose a normalized function φ_0 in the Hilbert space spanned by the eigenfunctions of Eq. (2.1). Decomposition of f into two orthogonal components, one of which is proportional to φ_0, gives the equalities below,

$$f = A\varphi_0 + g, \tag{2.2a}$$

$$(\varphi_0, g) = 0, \tag{2.2b}$$

where A is constant and the left side of the last equality denotes the inner product of φ_0 with g. By using Eqs. (2.2a) and (2.2b) in Eq. (2.1) and taking the inner product of the resulting equation with φ_0 one can get the following relation:

$$(\varphi_0, \Lambda\varphi_0)A + (\varphi_0, \Lambda g) = \lambda(\varphi_0, W\varphi_0)A + \lambda(\varphi_0, Wg). \tag{2.3}$$

If we define some projection operators P_0, P_c in the following equations where I denotes the unit operator on the aforementioned Hilbert space and h represents an arbitrary function in the same space,

$$P_0 h = (\varphi_0, h)\varphi_0, \tag{2.4}$$

$$P_c = I - P_0, \tag{2.5}$$

 0022-2488/83/010101-06$02.50

we can obtain another relation between g and A by using the fact that P_c is the unit operator on the complementary space of φ_0; hence $g = P_c\, g$. Following some algebraic steps we can get the formal result given below,

$$g = -[P_c(\Lambda - \lambda W)P_c]^{-1}P_c(\Lambda - \lambda W)\varphi_0 A. \quad (2.6)$$

The elimination of g using Eqs. (2.6) and (2.3) leads to the following algebraic equation:

$$(\varphi_0, [\Lambda - \lambda W]\varphi_0)A = (\varphi_0, [\Lambda - \lambda W]P_c[P_c(\Lambda - \lambda W) \times P_c]^{-1}P_c[\Lambda - \lambda W]\varphi_0)A. \quad (2.7)$$

Recalling the relation $\varphi_0 = P_0\varphi_0$, we can immediately notice that the operators in the expectation values above can be obtained from the matrix representation of $\Lambda - \lambda W$ via some reduction operators. We can therefore conclude the following equality after some intermediate steps:

$$(\varphi_0[\Lambda - \lambda W]\varphi_0) - (\varphi_0,[\Lambda - \lambda W]P_c[P_c(\Lambda - \lambda W)P_c]^{-1} \times P_c[\Lambda - \lambda W]\varphi_0) = (\varphi_0[\Lambda - \lambda W]^{-1}\varphi_0)^{-1}. \quad (2.8)$$

This result together with Eq. (2.7) implies the fact which is in a certain sense apparent that the zeros of the following function are the values of λ:

$$\Delta(\varphi_0|\lambda) = (\varphi_0, \Lambda^{-1}\varphi_0)(\varphi_0, [\Lambda - \lambda W]^{-1}\varphi_0)^{-1}. \quad (2.9)$$

We name this function "Characteristic Function of Eq. (2.1) with respect to the basis function φ_0" or briefly "Characteristic Function." For the characteristic function to exist, Λ must be assumed to be positive definite and this does not contradict with our assumptions about the operators Λ, W. Indeed, in the case where only W is positive definite, reformulation of Eq. (2.1) with a new eigenvalue parameter λ^{-1} instead of λ makes it possible to interchange W with Λ.

Let us now investigate the case $A = 0$ which satisfies Eq. (2.7). In this situation one has to find a nonzero element for g in Hilbert space to obtain a nontrivial solution of Eq. (2.1). However, a nonzero g with a vanishing A implies that $P_c(\Lambda - \lambda W)P_c$ must be singular and finally φ_0 must be an eigenfunction of Eq. (2.1), as can be deduced from Eq. (2.6). But this means φ_0 has coincidentally been chosen as the exact solution of Eq. (2.1); therefore the case $A = 0$ can be interpreted as trivial.

The selection of φ_0 may affect the number of zeros of the characteristic function. Indeed, in the case where φ_0 can be expressed as a finite linear combination of the eigenfunctions of Eq. (2.1) the characteristic function produces a finite number of eigenvalues. The possibility of selecting φ_0 as a finite linear combination of the eigenfunctions by chance decreases when the structures Λ and W are complicated.

The invariances of Λ and W under certain transformations (for example, the exchange of the particles coordinates) give the possibility of separating the solution space for Eq. (2.1) into two subspaces, one of which contains the symmetric functions and the other the antisymmetric ones, under one of the transformations mentioned above. If, however, all these transformations are commutative each of the subspaces can be further separated into similar subspaces. If φ_0 has been selected in one of these subspaces, the characteristic function will definitely not give any eigenvalue corresponding to other subspaces.

III. PADÉ SCHEME FOR THE CHARACTERISTIC FUNCTION AND ITS CONVERGENCE

The characteristic function defined by Eq. (2.9) has an explicit form. To find the zeros, what one needs is its explicit expression. One of the possible ways to this end is to expand the characteristic function in powers of λ. However, this type of expansion (Taylor series) does not cover all the complex domain of λ. This fact can be seen by using an expansion of φ_0 in terms of the true eigenfunctions of Eq. (2.1). Indeed, such an expansion creates the reciprocal of an infinite sum over simple fractions and the convergence domain of the Taylor series for this type of functions is restricted. In spite of its restricted convergence domain, the Taylor expansion of the characteristic function is not all that incovenient. Using some analytic continuation methods one can obtain the expression of the characteristic function over the whole complex domain of λ. The Padé scheme[8] which has recently been widely used, is a powerful example amongst such methods. We shall also employ this scheme to obtain the approximate spectrum of the system under consideration. For this purpose, first of all, we shall try to make a bridge between the characteristic function and Stieltjes series.[8] We shall then have the possibility of learning about the convergence of the Padé scheme and some of its important properties. Towards this goal, we can begin with the following theorem.

Theorem 3.1: If λ is a bounded positive definite Hermitian operator the function $\tilde{\Delta}$ defined as

$$\tilde{\Delta}(\phi|\lambda) = (\phi,[I + \lambda\tilde{\Lambda}]^{-1}\phi), \quad (3.1)$$

when expanded into the powers of λ, produces a Stieltjes series.

Proof: The serial representation of $\tilde{\Delta}$ can be written as

$$\tilde{\Delta}(\phi|\lambda) = \sum_{j=0}^{\infty}(\phi, \tilde{\Lambda}^j\phi)(-\lambda)^j. \quad (3.2)$$

If we construct an $n+1$ dimensional square matrix with its elements defined in the following manner,

$$\Omega_{jk}^{mn} = (\phi, \tilde{\Lambda}^{m+j+k}\phi), \quad m,n = 0,1,2,..., \quad j,k = 0,1,2,...,n, \quad (3.3)$$

all we have to do is to show the validity of the following set of inequalities known as one of the definitions of the Stieltjes series:

$$\det \Omega^{0n} > 0, \quad n = 0,1,2,..., \quad (3.4a)$$

$$\det \Omega^{1n} > 0, \quad n = 0,1,2,.... \quad (3.4b)$$

Now, consider the following quadratic form (C_j's are arbitrary constants):

$$\sum_{j,k=0}^{n} C_j C_k \Omega_{jk}^{mn} = (\phi,\tilde{\Lambda}^m|\sum_{j=0}^{n} C_j\tilde{\Lambda}^j|^2\phi). \quad (3.5)$$

Since $\tilde{\Lambda}$ is Hermitian and positive definite $\tilde{\Lambda}^{1/2}$ can be defined easily and this fact gives the possibility of writing the following equalities:

$$\tilde{\phi} = \sum_{j=0}^{n} C_j\tilde{\Lambda}^{j+m/2}\phi, \quad (3.6a)$$

$$\sum_{j,k=0}^{n} C_j^* C_k \Omega_{jk}^{mn} = (\tilde{\phi}, \tilde{\phi}) = ||\tilde{\phi}||^2. \tag{3.6b}$$

Using the positive definiteness of $\tilde{\Lambda}$, one can conclude that the quadratic form given by Eq. (3.5) is always positive for any nonzero basis function ϕ. However, a careful use of the matrix theory shows that Ω_{jk}^{mn} must be a positive definite matrix and this implies the validification of Eqs. (3.4a), and (3.4b) and therefore the correctness of the Theorem 3.1.

From this theorem the following corollaries can be written by recalling some theorems about Stieltjes series.[8]

Corollary 3.1: The reciprocal of the characteristic function can be expresed as a Stieltjes series by expanding into powers of λ. [Indeed the use of the scalar λ, the function $\phi = (\varphi_0, \Lambda^{-1}\varphi_0)\Lambda^{1/2}\varphi_0$, and the operator $\tilde{\Lambda} = \Lambda^{-1/2} W \Lambda^{-1/2}$ instead of $-\lambda, \varphi_0$ and W, respectively, brings us to this conclusion.]

Corollary 3.2: Any sequence of $[L + M/L]$ Padé approximants to the Stieltjes series for the reciprocal of the characteristic function converges to an analytic function in the cut complex plane $0 < \lambda < \infty$ as L increases unboundedly. The index M is restricted as $M > -1$ and the definition of $[L/M]$ Padé approximants can be written as follows:

$$[L/M] = P_L(\lambda)/Q_M(\lambda), \quad Q_M(0) = 1, \tag{3.7}$$

where P_L and Q_M are the L th and M th order polynomials of λ, respectively.

Corollary 3.3: The λ values obtained by using the $[L + M/L]$ $(M > -1)$ Padé approximants to the reciprocal of the characteristic function are on the positive real axis and the λ values (approximate eigenvalues) corresponding to the successive approximants interlace.

This theorem and its corollaries show how to obtain an approximate spectrum for a positive definite operator, furthermore they guarantee the convergence of the presented scheme. Therefore, to obtain the spectrum of a positive definite operator one has to evaluate the terms like $(\phi, \tilde{\Lambda}^j \phi)$, to construct the Padé table and then to arrive at the approximate spectrum by tracing some diagonals starting from one of the approximants like $[L/1]$, $L = 1,2,\ldots$ or $[1/2]$. Sections to follow will cover this type of work for electrically charged particles.

IV. ELECTRICALLY CHARGED PARTICLES AND HYPERSPHERICAL SPECTRAL COEFFICIENTS

Quantum chemical systems are composed of electrons and nuclei. After the separation of mass center coordinates and a suitable diagonalization procedure, the spin-free Schrödinger equation can be put into the following form without taking care of relativistic contributions[9,10]:

$$r(\tfrac{1}{4} - \nabla^2)\psi = \omega v(\theta)\psi, \quad \omega > 0; \quad \psi(r) = 0, \tag{4.1}$$

where r, θ, ∇^2, and $v(\theta)$ stand for the hyper-radial coordinate, the set of hyperangles, $3N$-dimensional Laplacian ($N + 1$ = the number of the particles in the system), and the hyperangular interaction potential,[10] respectively. The accompanying boundary conditions for Eq. (4.1) are the usual continuity conditions, at the singular points of the system's Hamiltonian. The eigenvalue parameter φ is related to the system's dimensionless energy parameter E as follows:

$$E = -1/2\omega^2. \tag{4.2}$$

Equation (4.1) is different than the original Schrödinger's equation for a system of electrically charged particles. Indeed in the original Schrödinger's equation the potential term $v(\theta)$ forms a part of the operator, the eigenvalues of which are investigated. In the present case, however, $v(\theta)$ is in a different position. Since the structures of the operators change when we transform from the original Schrödinger's equation into Eq. (4.1) we can expect the spectral behavior of the problem to change also. Due to the fact that the characteristic values of Eq. (4.1) correspond to the negative—bound state—energy values of system and this part of the energy spectrum is discrete we conjecture that the ω spectrum is also discrete. As a matter of fact the ω spectrum of the simplest system—hydrogen atom—is discrete, although its energy spectrum has discrete (some negative values) and continuous (all positive real axis) spectra.

Although some or all of the characteristic values of Eq. (4.1) may be negative depending on the nature of the interaction characterized by $v(\theta)$, the constraint $\omega > 0$ in Eq. (4.1) which appeared while transforming the original Schrödinger's equation into Eq. (4.1), eliminates these negative ω values. Therefore negative ω values do not correspond to any physical state; however, their existence may help us to classify the bounded states. To this end we can name five possible cases in the following manner with respect to the nature of the ω spectrum: (i) only positive ω values, "completely bounded system," (ii) including zero, a finite number of negative ω values in addition to positive ω values, "incompletely bounded system with a finite deficiency," (iii) infinitely many positive and negative, ω values, "incompletely bounded system with infinite deficiency," (iv) a finite number of positive ω values in addition to negative ω values, "highly deficient system," (v) only negative ω values, "unbounded system."

Now, if we define Λ and W as follows,

$$\Lambda = r(\tfrac{1}{4} - \nabla^2), \tag{4.3}$$

$$W = v(\theta), \tag{4.4}$$

to evaluate the characteristic ω values we can use the characteristic function approach presented in previous sections. Toward this end the following expansion can be employed:

$$\tilde{\Delta}(\phi \mid -\omega) = \sum_{j=0}^{\infty} \Delta_j \omega^j, \tag{4.5}$$

where Δ_j, the "hyperspherical spectral coefficient with respect to the basis function φ_0" or briefly HSC(φ_0), can be explicitly expressed by using an orthonormal expansion in terms of hyperspherical harmonics and their addition theorem[11] to put the inverse of Λ given by (4.3) into an integral operator form as follows:

$$\Delta_j = \left(\frac{\Gamma(\alpha)}{2\pi^{\alpha+1}}\right)^{j-1} \sum_{k_1=0}^{\infty} \cdots \sum_{k_j=0}^{\infty} \int_0^{\infty} \int_{S_1} \cdots \int_{S_j} \phi(r, \xi_1) \times \left\{ \prod_{l=1}^{j-1} [\rho + k_l(k_l + 2\alpha)]^{-1} r \right\} v(\xi_1)$$

$$\times\left[\prod_{l=1}^{j-1}(k_l+\alpha)C_{k_l}^{\alpha}(\xi_l^T\xi_{l+1})\,v(\xi_{l+1})\right]$$
$$\times\phi(r,\xi_j)dS_j\cdots dS_1dr, \tag{4.6}$$

where $v(\xi)$ is used instead of $v(\theta)$ and ξ_l and $C_k^{\alpha}(x)$ stand for the unit position vector which only depends on hyperangles and αth kind k th order Gegenbauer polynomial,[11] respectively. The parameter α is defined as $(3N-2)/2$ and the function ϕ is as defined in Corollary 3.1. ρ denotes the hyper-radial part of the operator $r\Lambda$.

By defining a new operator $\hat{\rho}$ as follows,

$$\rho r^{-\alpha}=r^{-\alpha}(\hat{\rho}+\alpha^2), \tag{4.7}$$

and using the Lebedev transform[12] technique which enables us to express the effect of the operator $r\hat{\rho}[(k+\alpha)^2]^{-1}$ on a Lebedev transformable function of $r, h_1(r)$ as follows,

$$h_2(r)=r[\hat{\rho}+(k+\alpha)^2]^{-1}h_1(r), \tag{4.8}$$

$$H_2(y)=\int_0^{\infty}[x^2+(k+\alpha)^2]^{-1}\times(\cosh\pi x+\cosh\pi y)^{-1}2y\sinh\pi yH_1(x)\,dx, \tag{4.9}$$

we can convert the hyper-radial parts of Eq. (4.6) into $(j+1)$-dimensional integrals without using any differential operator included in their kernels. H_1 and H_2 appearing in Eq. (4.9) are Lebedev transforms of h_1 and h_2, respectively, and the argument of modified Bessel function[12] in the kernel of the Lebedev transform is multiplied by $\frac{1}{2}$ for the sake of convenience.

On the other hand the following relation shows that the sums on Gegenbauer polynomials in Eq. (4.6) can be expressed in terms of hypergeometric functions.[12] Using some properties of hypergeometric functions

$$\sum_{k=0}^{\infty}\frac{k+\alpha}{(k+\alpha)^2+x^2}C_k^{\alpha}(\xi^T\eta)=\frac{1}{2}\int_0^{\infty}\frac{t^{\alpha-1+ix}+t^{\alpha-1-ix}}{[1-2t(\xi^T\eta)+t^2]^{\alpha}}dt. \tag{4.10}$$

This last relation can be verified by using the method of separation into partial fractions, integral representation[12] of these partial fractions, and the summation formula[12] for Gegenbauer polynomials.

Therefore by defining a new function $\beta_{\alpha}(x,y)$,

$$\beta_{\alpha}(x,y)=x\sinh\pi x\left[\Gamma\left(\frac{\alpha+ix}{2}\right)\Gamma\left(\frac{\alpha-ix}{2}\right)\times{}_2F_1\left(\begin{matrix}\frac{\alpha+ix}{2},\frac{\alpha-ix}{2}\\ \frac{1}{2}\end{matrix}\middle|y^2\right)+2y\Gamma\left(\frac{\alpha+1+ix}{2}\right)\Gamma\left(\frac{\alpha+1-ix}{2}\right)\times{}_2F_1\left(\begin{matrix}\frac{\alpha+1+ix}{2},\frac{\alpha+1-ix}{2}\\ \frac{3}{2}\end{matrix}\middle|y^2\right)\right], \tag{4.11}$$

one can arrive at the following expression for Δ_j after some intermediate steps:

$$\Delta_j=2^{2-2j}\pi^{(1-j)(\alpha+1)-2}\times\int_0^{\infty}\cdots\int_0^{\infty}\int_{S_j}\cdots\int_{S_1}\phi(r,\xi_1)r^{\alpha}K_{ix_1}\left(\frac{r}{2}\right)K_{ix_{j-1}}\left(\frac{\bar{r}}{2}\right)\times\phi(\bar{r},\xi_j)\bar{r}^{\alpha}\prod_{l=1}^{j}v(\xi_l)\prod_{l=1}^{j-1}\beta_{\alpha}(x_l,\xi_l^T\xi_{l+1})\times\prod_{l=1}^{j-2}(\cosh\pi x_l+\cosh\pi x_{l+1})^{-1}\times dS_1\cdots dS_j dx_1\cdots dx_{j-1}\,dr\,d\bar{r}. \tag{4.12}$$

In the case where the ω spectrum has negative together with positive spectra the convergence of the Padé scheme for the reciprocal of the characteristic function can be proved depending on the locations of poles for some domains of ω. However, this subject will be held outside of this paper, since some redefinitions of the operators and parameters make it possible to study with positive definite operators.

V. ANALYTICAL DETERMINATION OF THE FIRST AND SECOND HYPERSPHERICAL SPECTRAL COEFFICIENT

Since the evaluation of Δ_0 is trivial ($\Delta_0=1$) we can start with Δ_1. For its determination, one can select as the simplest basis function $\varphi_0, e^{-r/2}$, which is the ground state eigenfunction of $[r(\frac{1}{4}-\nabla^2)]^{-1}$ and this yields an integral on hyperangles. Using the explicit structure of potential function[10] $v(\theta)$ and rotating the hyperaxes in such a manner that the matrices appearing in the potential function are diagonalized, we can arrive at the following value for Δ_1:

$$\Delta_1=\frac{2\Gamma(\alpha+1)}{[\pi\Gamma(\alpha+\frac{3}{2})]^{1/2}}\left[-\sum_{j=1}^{N}\sum_{k=j+1}^{N+1}Z_jZ_k(\bar{m}_j+\bar{m}_k)^{-1/2}\right], \tag{5.1}$$

where Z_j and $\bar{m}_j$ denote "electrical charge parameter" and "mass parameter" of the jth particle, respectively.[10]

Performing the integration over r and $\bar{r}$ in Eq. (4.2) and recalling the evenness of the potential function $v(\xi)$ with respect to its argument, Δ_2 can be brought to a finite sum of the following $(6N-1)$ dimensional integrals:

$$S_{j_1k_1}^{j_2k_2}=\int_0^{\infty}\int_{S_{\eta}}\int_{S_{\xi}}(\xi^TV_{j_1k_1}\xi)^{-1/2}(\eta^TV_{j_2k_2}\eta)^{1/2}x\times\sinh\pi x\Gamma^2(\alpha+1+ix)\times\Gamma^2(\alpha+1-ix)\Gamma\left(\frac{\alpha+ix}{2}\right)\Gamma\left(\frac{\alpha-ix}{2}\right)\times{}_2F_1\left(\begin{matrix}\frac{\alpha+ix}{2},\frac{\alpha-ix}{2}\\ \frac{1}{2}\end{matrix}\middle|(\xi^T\eta)^2\right)dS_{\xi}d_{\eta}dx, \tag{5.2}$$

where V_{jk}'s denote potential matrices.[10] The integration over ξ and η can be analytically handled[13] and $S^{j_2k_2}_{j_1k_1}$ can be expressed in terms of some one-dimensional integrals as follows:

$$S^{j_2k_2}_{j_1k_1} = \frac{64\pi^{2\alpha+1}}{\Gamma(\alpha)}\left[\mathscr{S}_1(\gamma) + \mathscr{S}_2(\gamma) - \frac{\sqrt{\pi}\Gamma(\alpha)}{2\Gamma(\alpha-\frac{1}{2})}\mathscr{S}_3(\gamma)\right], \tag{5.3}$$

where γ has the meaning given in a previous paper[13] and the integrals denoted by $\mathscr{S}$'s have kernels which contain a hyperbolic sine function, some complex argumented gamma functions and their conjugates, and certain hypergeometric functions ${}_3F_2$, ${}_2F_1$ with complex parameters and their conjugates.

Using the explicit expression of the modulus of the complex argumented gamma function[12] for integer α values ($\alpha = n+1$) and the explicit structures of the generalized hypergeometric function[14] ${}_3F_2$ and Gaussian hypergeometric function[12,14] ${}_2F_1$ in addition to some properties of the digamma function ψ (logarithmic derivative of the gamma function), one can summarize the following results after detailed and tedious intermediate steps:

$$\mathscr{S}_1(\gamma) = \left\{\mathscr{D}_n\mathscr{D}_{n+1}\frac{1}{8\pi^3}\frac{d^{3\,n-1}}{dt^{3\,n-1}}\prod_{m=1}\mathscr{D}^2_m\psi(t+\tfrac{1}{2})\right\}_{t=0}\frac{\arcsin\gamma}{\gamma}, \tag{5.4}$$

$$\mathscr{S}_2(\gamma) = \frac{(-1)^n}{2^{n+1}}\frac{n!}{\gamma}\frac{d^n}{d\gamma^n}\left[\nu_n\left(\frac{1-\gamma}{1+\gamma}\right)^{1/2}\right] + \sum_{j=0}^{n-1}\frac{n!}{2\sqrt{\pi}}\frac{(1-\gamma^2)^{-j-1/2}}{(n-j-1)!(\Gamma\frac{1}{2}-j)}\{\hat{\nu}_{nj}\psi(t+\tfrac{1}{2})\}_{t=0}, \tag{5.5}$$

$$\mathscr{S}_3(\gamma) = (\tfrac{1}{2})_n(-1)^n 2^{n+1}\pi\frac{d^n}{d\gamma^n} \times\left[\frac{1+(-1)^n}{2}\sigma_1 - \frac{1-(-1)^n}{2}\sigma_2\gamma\right](1-\gamma^2)^{-1/2} \times \arcsin\gamma, \tag{5.6}$$

where the new entities appearing above can be explicitly defined as follows:

$$\mathscr{D}_m \equiv m^2 + \frac{1}{4\pi^2}\frac{d^2}{dt^2}, \tag{5.7a}$$

$$\mathscr{D}_m = m^2 + (1-\gamma^2)\frac{d^2}{d\gamma^2} - \gamma\frac{d}{d\gamma}, \tag{5.7b}$$

$$\nu_n \equiv 2^{2n}\pi^2\left\{\frac{1+(-1)^n}{2}\sigma_1 + \frac{1-(-1)^n}{2}\sigma_2\right\}, \tag{5.8}$$

$$\hat{\nu}_{nj} \equiv 2^{2n}\pi^2\left\{\frac{1+(-1)^n}{2}\prod_{m=0}^{n/2}\mathscr{D}^2_{2m+1} + \frac{1-(-1)^n}{2}\mathscr{D}_0\prod_{m=1}^{(n+1)/2}\mathscr{D}^2_{2m}\right\}\prod_{k=1}^{n-j}\mathscr{D}_{k-n/2}, \tag{5.9}$$

$$\sigma_1 \equiv \prod_{m=0}^{n/2} D^2_{2m+1}, \quad \sigma_2 \equiv D_0\prod_{m=1}^{(n+1)/2} D^2_{2m}. \tag{5.10}$$

These therefore complete the analytical evaluation of Δ_2.

VI. APPLICATIONS TO CERTAIN SYSTEMS AND CONCLUSION

Employing the values of Δ_1 and Δ_2 given in the previous section one can evaluate the Padê approximants $[L/M]$, where $L+M = 0,1,2$. For this purpose we can give the explicit expressions of these entities in terms of Δ_1, Δ_2, and ω as follows:

$$[0/0] = 1,$$
$$[0/1] = (1-\Delta_1\omega)^{-1},$$
$$[1/0] = 1+\Delta_1\omega,$$
$$[0/2] = [1-\Delta_1\omega+(\Delta_1^2-\Delta_2)\omega^2]^{-1},$$
$$[1/1] = [\Delta_1+(\Delta_1^2-\Delta_2)\omega][\Delta_1-\Delta_2\omega]^{-1},$$
$$[2/0] = 1+\Delta_1\omega+\Delta_2\omega^2. \tag{6.1}$$

These equalities show that the first column of the Padé table has no pole. Therefore it does not produce any value for ω. Physically reasonable ω values obtained from [0/1], [1/1], and [0/2] for certain three particle systems are presented in Table I. As can be noticed easily the approximation for He is in good agreement with the exact results.[15] However, discrepancy between the calculated and exact values increases as the atomic number increases in helium isoelectronic series. This difficulty can be removed by using higher order Padé approximants or different types of basis function. The hydrogen anion case ought to be fundamentally different than heliumlike systems, since it seems to have negative values in its ω spectrum. The mathematical character of the hydrogen anion is under a detailed investigation and possibly will be published in the future. The selection of three-particle systems as examples is due to the fact that they are the most realistic systems which have symmetric eigenfunctions under coordinate exchange transformation among all N-particle systems. However, for integer α values, similar calculations can be handled and some approximate values can be obtained since integer α values make it possible to evaluate Δ_2 analytically. However, half-integer α values are also as much realistic as the integer ones. In this case at least numerical methods can be utilized to this end. Due to the fact that all evaluations use a symmetric basis function φ_0, all results will correspond to the symmetric eigenfunctions of the system under consideration. Although these results give

TABLE I. Energy values[a] obtained from several Padé approximants for heliumlike systems.

	E_{01}[b]	E_{02}[c]	E_{11}	E_{PK}[d]
H^-	0.3854	0.5349	0.5642	0.5278
He	2.5000	2.8659	2.8927	2.9037
Be^{++}	12.2627	13.2915	13.3347	13.6555
C^{+4}	29.4034	31.4083	31.4766	32.4062
O^{+8}	53.9221	57.2169	57.3176	59.1565

[a] To obtain the energy values in eV all columns must be multiplied by -27.196 eV.
[b] E_{LM} has been calculated from the pole of $[L/M]$ Padé approximant.
[c] The [0/2] term has two poles, but only one of them is physically meaningful.
[d] These results are due to Pekeris *et al.* (Ref. 15).

some information about the mathematical structure of the systems spectrum, the Pauli principle makes them physically meaningless almost for all atoms and molecules. Since we did not consider contributions due to the spin of particles there remains only a few systems such as heliumlike atoms which have symmetric eigenfunctions in the physical sense. On the other hand there does not seem to be any clue which shows the existence of a spectral series corresponding to symmetric behavior of the system in the experimental results. Therefore the symmetric basis function φ_0 will be useful only for a few systems, in the physical sense. However, these calculations can be realized for other types of basis functions without extra effort. The studies to this end are under a condensed work.

As can be easily noticed, spectral coefficients keep all information about the system under investigation. Therefore more accurate results await the values of higher order hyperspherical spectral coefficients. The work for this purpose has been almost completed. After finalizing some details it will be the subject of a coming publication.

[1]P-O. Löwdin, in *Perturbation Theory and Its Applications in Quantum Mechanics*, edited by C. H. Wilcox (Wiley, New York, 1966).
[2]D. Grau, Int. J. Quantum Chem. **XI**, 931 (1977).
[3]G. E. Baker, Jr., Phys. Rev. **124**, 768 (1961).
[4]G. A. Baker, Jr. and J. L. Gammel, *The Padé Approximant in Theoretical Physics* (Academic, New York, 1970).
[5]L. A. Copley and D. Masson, Phys. Rev. **164**, 2059 (1967).
[6]C. Pommerenke, J. Math. Anal. Appl. **41**, 775 (1973).
[7]S. Wilson, Mol. Phys. **39**, 525 (1980).
[8]G. A. Baker, Jr., *Essentials of Padé Approximants* (Academic, New York, 1975).
[9]M. Demiralp and E. Şuhubi, J. Math. Phys. **18**, 777 (1977).
[10]M. Demiralp, J. Chem. Phys. **72**, 3828 (1980).
[11]A. Erdelyi, *Higher Transcendental Functions*, Vol. II (McGraw Hill, New York, 1953).
[12]W. Magnus, F. Oberhettinger, and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, Berlin 1966).
[13]M. Demiralp and N. A. Baykara, "Analytic evaluation of certain zeroth order coulombic hyperangular interaction integrals," J. Math. Phys. **22**, 2427 (1981).
[14]Y. L. Luke, *The Special Functions and Their Approximations*, Vol. I (Academic, New York, 1969).
[15]C. L. Pekeris, Phys. Rev. **112**, 1649 (1958); **115**, 1216 (1959); **126**, 143 (1962); **126**, 1470 (1962).

Perturbation expansion of S matrix for background scattering

S. Bosanac [a),b)]
Quantum Theory Project, Williamson Hall, University of Florida, Gainesville, Florida 32611

(Received 21 May 1981; accepted for publication 6 October 1981)

The S matrix near a pole is parametrized into the contribution from the resonance and the background scattering. We develop a perturbation theory for the background scattering, based on the Jost function formalism. A closed expression is found up to the second order in the coupling strength between the channels. A brief comparison with the other formalism is also made and the advantages of the present theory are shown.

PACS numbers: 03.80. + r, 11.20.Dj, 24.10.Dp

1. INTRODUCTION

The major problem in the theory of resonances is to give a qualitative and quantitative description of how they are observed in the scattering cross sections. In the collisions where the scattering amplitude is given by the contribution of only a few partial waves, e.g., nuclear reactions, electron-atom scattering or atom-surface collisions, the problem is always reduced to a question: how the S matrix is parametrized in the vicinity of a resonance. The simplest answer is given using the complex energy formalism in which a resonance is represented by a pole of the S matrix.[1,2] In general, the pole is a complex number,[3] and since we will regard the S matrix as a function of the wavenumber rather than the energy, we can write for a general element of the S matrix in the vicinity of a pole k_0

$$S_{m,n} \sim \beta_{m,n}/(k - k_0) + b_{m,n}, \tag{1.1}$$

where k is the real wavenumber corresponding to the collision energy. Broadly speaking (1.1) is a three parameter formula: k_0 gives the position of the resonance [$\mathrm{Re}(k_0)$] and its width [$\mathrm{Im}(k_0)$], while the residue measures the height of the resonance and the background term $b_{m,n}$ describes its shape.

The parametrization (1.1) is the well-known Breit–Wigner formula,[4,5] except that instead of the wavenumber they used the energy as the variable. However, this is not the major obstacle since by multiplying both the numerator and the denominator of (1.1) by $(k + k_0)$, we obtain the usual Breit–Wigner form. The form (1.1) gives quite a good description of the S matrix near a pole and would be of interest to relate the three parameters to the coupling matrix between the channels. Several schemes were proposed[6–9]; however, the one due to Feshbach[7] has been used in many applications. The advantage of the Feshbach formalism is that it can also be used away from a resonance, and therefore, it can serve as a general method for the computation of the S matrix in the problems where the channels can be separated into the closed and open ones. The weakness of the method has been discussed,[10] especially when it is used as a basis for the perturbation theory of the poles, residues, and, as we will show in the present work, the background term in (1.1). The two major difficulties are (a) the contribution of the closed channels are separated from the contribution of the open channels and (b) the use of the complete set of the eigenfunctions of the uncoupled channels. As the result, in the first case, it is not clear how to treat the resonances which are the true resonances in the uncoupled open channels, while in the second case, one has to include in the theory the channels which are not directly present at a particular scattering energy. More about the second point will be briefly discussed in the Sec. 4 of the present work.

An alternative perturbation approach for calculating the poles and the residues was developed and applied to the Regge poles.[11] It is based on the fact that the poles of the S matrix are also the roots of the equation

$$\mathrm{Det}(J) = 0, \tag{1.2}$$

where J is the Jost function. In the case of the Regge poles it was demonstrated how the theory is applied even in the instances where the Feshbach theory was criticized. In particular it was shown how the residues are calculated when the resonance originated as a true resonance in the uncoupled open channels. In the present work such resonances will not be treated since we will be only interested in the compound state resonances, i.e., the resonances which are the bound states in the uncoupled channels. In Sec. 2 of the present work we will briefly summarize the main points in the theory as applied to the energy poles of the S matrix.[10]

The same idea was applied to the perturbation problem in a single channel case.[12] It was found that the first-order perturbation correction agreed with the usual Rayleigh–Schrödinger theory; however, the second-order is different since it does not involve the use of the complete set of eigenfunctions of the unperturbed Hamiltonian. The theory was generalized to the multichannel case[10] where also the degenerate problem was studied and it was shown explicitly where the advantages of the present ideas are.

In this work we would like to study how the theory gives the background term in (1.1) under the following assumptions: (a) The coupling between the channels is weak, (b) the poles represent the compound state resonances, and (c) the poles are not degenerate in the unperturbed Hamiltonian. These restrictions are not essential, except the weak coupling assumption, since generalization to these cases is straightforward.

The derivation of the background term is not a unique procedure. Let us briefly discuss this point for an arbitrary holomorphic function $f(z)$ with the well separated first-order poles. Such a function is a representative of the S-matrix

[a)]On leave of absence from R. Bošković Institute, 41001 Zagreb, Croatia, Yugoslavia.
[b)]This work was supported in part from the grant NSF F6F006-Y.

 0022-2488/83/010107-08$02.50

elements. Let us assume that some of the poles have a small imaginary part, much smaller than the separation between the neighboring poles. Therefore we can write in the neighborhood of a pole

$$f(z) = \beta/(z - z_0) + g(z), \tag{1.3}$$

where z_0 is a pole of $f(z)$. The function $g(z)$ is now analytic in the circle, the radius of which is determined by the distance to the closest pole of z_0. Hence, $g(z)$ can be expanded in the circle in a Taylor series around any point. In the representation (1.1) we retain only the leading term in such an expansion, which we designate by b. It is obvious that its value is not given uniquely and depends on the point around which $g(z)$ is expanded. We must therefore impose certain conditions which will give its value uniquely. For example, we can choose that the difference

$$\Delta = f(z) - \beta/(z - z_0) - b \tag{1.4}$$

is minimal on the real axis of z. In particular we can assume that

$$\int_{\mathrm{Re}\, z_0 - \delta}^{\mathrm{Re}\, z_0 + \delta} |\Delta|^2 \, dx = \min, \tag{1.5}$$

where the variable of variation is b. The last condition is equivalent to the equation

$$\int_{\mathrm{Re}\, z_0 - \delta}^{\mathrm{Re}\, z_0 + \delta} \left(f(x) - \frac{\beta}{x - z_0} \right) dx = b\delta, \tag{1.6}$$

where δ is the interval on the real axis around Re z_0 in which (1.4) is the best representation of the S matrix. In practice solving (1.6) is not straightforward. Therefore, we will assume that the point of expansion of $g(z)$ coincides with z_0, in which case the formalism greatly simplifies. However, we should have in mind that this may not be the best choice for b, as we have shown in the preceding discussion.

2. PERTURBATION THEORY FOR THE POLES AND RESIDUES

In this section we will briefly review the perturbation theory of the poles and residues of the S matrix based on the Jost function formalism.[10] Let the set of equations describing inelastic processes involving n channels in the matrix notation be

$$\psi'' = (V - K^2)\psi \tag{2.1}$$

where $K_i^2 = k^2 - E_i$ and V is the $n \times n$ potential matrix. Let us assume that the first O channels correspond to the open channels and the subsequent C channels to the closed channels, i.e.,

$$\begin{aligned} &K_i^2 > 0, \quad i = 1,2,3,\ldots,O, \\ &K_i^2 \leqslant 0, \quad i = O+1,\ldots,n. \end{aligned} \tag{2.2}$$

Let us also assume that the off-diagonal elements of V are small compared to the diagonal ones, hence they can be treated as a perturbation

$$V = V_0 + \epsilon V', \tag{2.3}$$

where V' is zero on the diagonal. In such a case the regular solution of (2.1) is given in the form of the integral equation

$$\psi = \psi_0 + \frac{\epsilon}{2i} K^{-1} \int_0^r G(r, r') V'(r') \psi(r') dr', \tag{2.4}$$

where

$$G(r, r') = f_0^+(r') f_0^-(r) - f_0^+(r) f_0^-(r'), \tag{2.5}$$

where ψ_0 and $f_0^{\pm}$ are the regular and irregular solutions of (2.1), respectively, when $\epsilon = 0$. The regular solution is defined with the boundary condition

$$\psi_0 \to 0, \quad r \to 0, \tag{2.6}$$

and the irregular

$$f_0^{\pm}(r) \sim \exp(\mp iKr), \quad r \to \infty. \tag{2.7}$$

The Jost function is then[13]

$$J = J_0 - \frac{\epsilon}{2i} K^{-1} \int_0^{\infty} dr f_0^-(r) \, V'(r) \, \psi(r), \tag{2.8}$$

and the roots of the equation

$$F = \mathrm{Det}(J) = 0 \tag{2.9}$$

in the variable k give the poles of the S matrix, which are interpreted as the bound states and resonances of the system. Approximate solutions of (2.9) are obtained if we set $\epsilon = 0$, in which case

$$F = \mathrm{Det}(J_0) = j_1 j_2 \cdots j_n = 0, \tag{2.10}$$

where j_k are the diagonal elements of J_0. Therefore, the set of the poles of the S matrix correspond in the zeroth order of ϵ to the set of poles in the uncoupled channels. In our treatment we will assume that the set of poles, obtained from (2.10), are not degenerate, i.e., the poles from different channels are not equal.

Let us designate by κ one of the roots of (2.9). Since F is also a function of ϵ, we can expand κ in the power series

$$\kappa(\epsilon) = k_0 + \epsilon k_1 + \tfrac{1}{2}\epsilon^2 k_2 + \cdots, \tag{2.11}$$

where k_0 is a solution of (2.10). For simplicity we will assume that k_0 represents a bound state in the pth closed channel. This restriction is not essential since the perturbation theory of resonances can be equally applied to the resonances in the uncoupled channels. However, a bound state in the uncoupled channels produces the Feshbach type resonance and they are by far more important in the study of inelastic collision processes than the shape resonances, how the other type is usually referred to.

As was shown, the coefficients in (2.11) are[10]

$$k_1 = 0 \tag{2.12}$$

and

$$\begin{aligned} k_2 = &- \frac{1}{2K_p j_p' j_p^+} \sum_{l=1 \neq p}^{n} \frac{1}{K_l j_l} \\ &\times \left(\int_0^{\infty} \psi_p V_{pl} \psi_l \, dr \int_r^{\infty} f^-(r') V_{lp}(r') \, \psi_p(r') \, dr' \right. \\ &\left. + \int_0^{\infty} dr \, \psi_p V_{pl} f_l^- \int_0^r dr' \, \psi_l(r') V_{lp}(r') \psi_p(r') \right), \end{aligned} \tag{2.13}$$

where the wave functions correspond to the unperturbed solutions of (2.1). The unperturbed Jost functions in (2.13) are defined from the asymptotic form of the regular solution ψ_0 in the l th channel,

$$\psi_l \sim j_l^+ \exp(iK_l r) + j_l \exp(-iK_l r), \quad r \leftarrow \infty, \tag{2.14}$$

and j_p' is defined as

$$j_p' = \frac{dj_p}{dk}. \tag{2.15}$$

Similarly the residues of the S matrix can be calculated using the perturbation theory. It was shown[10] that for the S-matrix element $S_{m,n}$, corresponding to the open channels, the appropriate residue has the parametrization

$$\lim_{k \to \kappa}(k-\kappa)S_{m,n} = (\beta_m \beta_n)^{1/2} \tag{2.16}$$

and that β_m and β_n have the expansion

$$\beta_m = \beta_m^{(0)} + \epsilon \beta_m^{(1)} + \tfrac{1}{2}\epsilon^2 \beta_m^{(2)} + \cdots. \tag{2.17}$$

It turns out that

$$\beta_m^{(0)} = \beta_m^{(1)} = 0 \tag{2.18}$$

and

$$\beta_m^{(2)} = -\frac{1}{2K_m j_m^2} \frac{1}{j_p' j_p^+ K_p} \left(\int_0^\infty dr\, \psi_m V_{mp} \psi_p \right)^2, \tag{2.19}$$

where again ψ_m and ψ_p refer to the unperturbed regular solutions of (2.1).

3. PERTURBATION EXPANSION OF BACKGROUND TERM

Near a resonance the S matrix has a parametrization in the form, as first given by Breit and Wigner,

$$S_{m,n} \sim \frac{(\beta_m \beta_n)^{1/2}}{k-\kappa} + b_{m,n}, \tag{3.1}$$

where $b_{m,n}$ is the background term. We have shown in the previous section how to obtain κ and β in the form of a perturbation expansion in ϵ, defined in (2.3). We will now show that $b_{m,n}$ can also be obtained in an analogous manner. To do this, let us recall a useful representation of the S matrix in terms of the functions F, defined in (2.9). For the elastic channels we have [14]

$$S_{m,m} = F(-K_m)/F, \tag{3.2}$$

where $-K_m$ means that the channel wavenumber K_m in F is replaced by its negative value. Near a pole we can write

$$\begin{aligned} S_{m,m} &\sim \frac{F_0(-K_m) + (k-\kappa)F_1(-K_m)}{(k-\kappa)F_1 + \frac{1}{2}(k-\kappa)^2 F_2} \\ &\sim \frac{\beta_m}{k-\kappa} + \frac{F_1(-K_m)}{F_1} - \frac{1}{2}\frac{F_2}{F_1}\beta_m, \end{aligned} \tag{3.3}$$

where the index of F designates the derivative with respect to k at the pole κ. Therefore the background term $b_{m,m}$ is given by

$$b_{m,n} = \frac{F_1(-K_m)}{F_1} - \frac{1}{2}\frac{F_2}{F_1}\beta_m. \tag{3.4}$$

To obtain the elastic background term as a power series in ϵ, let us first calculate F_1 in the form

$$\frac{\partial F}{\partial k} \equiv F_1 = F_1^{(0)} + \epsilon F_1^{(1)} + \tfrac{1}{2}\epsilon^2 F_1^{(2)} + O(\epsilon^3), \tag{3.5}$$

where the terms of higher order than ϵ^2 are neglected. Since F is a determinant of the matrix J we have

$$\begin{aligned} F_1 = \sum_{m=1}^{n} \mathrm{Det}^{(m)}(J) = \sum_{m=1}^{n} (&\mathrm{Det}_0^{(m)}(J) \\ &+ \mathrm{Det}_1^{(m)}(J) + \epsilon^2/2\, \mathrm{Det}_2^{(m)}(J)), \end{aligned} \tag{3.6}$$

where $\mathrm{Det}^{(m)}$ designates the derivative of the mth column of J. If $m \neq p$, where p is the index of the pth closed channel for which

$$j_p = 0, \tag{3.7}$$

we easily find

$$\mathrm{Det}_0^{(m)}(J) = \mathrm{Det}_1^{(m)}(J) = 0, \tag{3.8}$$

where we have taken into account that

$$j_p \sim (\kappa - k_0) j_p' \sim \tfrac{1}{2}\epsilon^2 k_2 j_p' + O(\epsilon^3), \tag{3.9}$$

where k_2 is given by (2.13). It follows that

$$\begin{aligned} &\mathrm{Det}_2^{(m)}(J) \\ &\quad = 2P(j)\frac{j_m'}{j_m}\left(\frac{J_{pm}^{(1)} J_{mp}^{(1)}}{j_m} - \frac{J_{pm}'^{(1)} J_{mp}^{(1)}}{j_m'}\right), \quad m \neq p, \end{aligned} \tag{3.10}$$

where

$$P(j) = \prod_{i \neq p} j_i \tag{3.11}$$

and

$$J_{pm}^{(1)} = -\frac{1}{2iK_p}\int_0^\infty dr\, f_p^- V_{pm}\, \psi_m. \tag{3.12}$$

When $p = m$, the derivative of the pth column will give for the diagonal (p,p) element of J

$$J_{pp}' \sim j_p' + \tfrac{1}{2}\epsilon^2(k_2 j_p'' + J_{pp}'^{(2)}) + O(\epsilon^3); \tag{3.13}$$

hence,

$$\mathrm{Det}_0^{(p)}(J) = j_p' P(j), \quad \mathrm{Det}_1^{(p)}(J) = 0, \tag{3.14}$$

and

$$\begin{aligned} \mathrm{Det}_2^{(p)}(J) = P(j)\Bigg(&k_2 j_p'' + J_{pp}'^{(2)} \\ &+ \sum_{l=1\neq p}^{n}\left(\frac{j_p'}{j_l} J_{ll}^{(2)} - \frac{2}{j_l} J_{lp}'^{(1)} J_{pl}^{(1)}\right) \\ &- 2j_p' \sum_{m>l\neq p} \frac{1}{j_l j_m} J_{lm}^{(1)} J_{ml}^{(1)}\Bigg), \end{aligned} \tag{3.15}$$

where

$$J_{mm}^{(2)} = \frac{1}{2K_m}\int_0^\infty dr\, f_m^- \sum_{i=1}^{n} \frac{V_{mi}}{k_i} \int_0^r G_i(r,r')\, V_{im}\psi_m\, dr' \tag{3.16}$$

which was obtained from (2.8) as the second iteration. We should recall that $F(-K_m)$ is obtained from (3.6) by replacing K_m with $-K_m$ in Eqs. (3.10), (3.14), and (3.15). In such a case we have for the ratio $F_1(-K_m)/F_1$,

$$\begin{aligned} \frac{F_1(-K_m)}{F_1} = &\frac{j_m^+}{j_m} - \epsilon^2 \frac{j_m^+}{j_m} \\ &\times\left(\frac{1}{j_p'}\frac{\partial}{\partial k}\left(\frac{j_p^+}{4K_m K_p j_m j_m^+}\right.\right. \end{aligned}$$

$$\times\left(\int_0^\infty dr\, f_p^- V_{pm}\psi_m\right)^2\Bigg)$$

$$+\frac{1}{4K_mK_p j_m j_m^+ j_p^+}$$

$$\times\Bigg(\int_0^\infty dr\,\psi_m V_{mp}\psi_p \int_0^r f_p^+ V_{pm}\psi_m\, dr'$$

$$+\int_0^\infty \psi_m V_{mp} f_p^+\, dr \int_r^\infty \psi_p V_{pm}\psi_m\, dr'\Bigg)$$

$$+\frac{1}{4K_m j_m j_m^+}\sum_{l=1\neq p}^{n}\frac{1}{K_l j_l}$$

$$\times\Bigg(\int_0^\infty dr\,\psi_m V_{ml} f_l^- \int_0^r \psi_l V_{lm}\psi_m\, dr'$$

$$+\int_0^\infty \psi_m V_{ml}\psi_l\, dr \int_r^\infty f_l^- V_{lm}\psi_m\, dr'\Bigg)\Bigg). \tag{3.17}$$

To obtain $b_{m,m}$ we still require the ratio F_2/F_1. Since β_m is of the order $O(\epsilon^2)$ we have to calculate F_2/F_1 only to the zeroth order in ϵ. Hence

$$\frac{\partial^2 F}{\partial k^2}\equiv F_2 = P(j)\left(j_p'' + 2\sum_{m=1\neq p}^{n}\frac{j_m' j_p'}{j_m}\right) \tag{3.18}$$

and $b_{m,m}$ is

$$b_{m,m} = \frac{F_1(-K_m)}{F_1} - \beta_m/2\left(\frac{j_p''}{j_p'} + 2\sum_{l=1\neq p}^{n}\frac{j_l'}{j_l}\right) + O(\epsilon^3). \tag{3.19}$$

Let us now turn our attention to the inelastic background term $b_{m,n}$. From the representation of the inelastic S-matrix elements[14]

$$S_{m,n}^2 = \frac{F(-K_m)F(-K_n)}{F^2} - \frac{F(-K_m,-K_n)}{F}. \tag{3.20}$$

Near the pole we can write approximately

$$S_{m,n}^2 = \frac{\beta_m\beta_n}{(k-\kappa)^2} + \frac{1}{k-\kappa}\Bigg(\beta_n b_{m,m} + \beta_m b_{n,n} - \frac{F_0(-K_m,-K_n)}{F_1}\Bigg), \tag{3.21}$$

where we have used (3.4) and the definition of the residue. We have shown that the elastic background terms can be written as

$$b_{m,m} = b_{m,m}^{(0)} + \tfrac{1}{2}\epsilon^2 b_{m,m}^{(2)} \tag{3.22}$$

in which case

$$\beta_n b_{m,m}^{(0)} + \beta_m b_{n,n}^{(0)} - \frac{F_0(-K_m,-K_n)}{F_1} = O(\epsilon^3) \tag{3.23}$$

if the leading term of $F_0(-K_m,-K_n)$, which is of the order ϵ^2, is calculated by a procedure similar to calculating F_1. Noting that $\beta = O(\epsilon^2)$ and taking into account (3.22) we notice that the only contribution of the order ϵ^3 comes from $F_0(-K_m,-K_n)$, hence (3.21) is

$$S_{m,n}^2 = \frac{\beta_m\beta_n}{(k-\kappa)^2} - \frac{1}{k-\kappa}\left(\frac{F_0^{(3)}(-K_m,-K_n)}{F_1^{(0)}} + O(\epsilon^4)\right), \tag{3.24}$$

where the index of F means that this function is calculated to this order in ϵ.

Let us briefly show how $F_0(-K_m,-K_n)$ is calculated to the order ϵ^3. Since the diagonal pth element of J is zero in the limit $\epsilon = 0$ we can replace $J_{p,p}$ by

$$J_{p,p} \sim \tfrac{1}{2}\epsilon^2(k_2 j_p' + J_{pp}^{(2)}) + \tfrac{1}{6}\epsilon^3(k_3 j_p' + J_{pp}^{(3)}), \tag{3.25}$$

and $F_0(-K_m,-K_n)$ is

$$F_0(-K_m,-K_n) = \epsilon^3 \underline{F}(-K_m,-K_n). \tag{3.26}$$

In the Eq. (3.25) we have used (3.9). The determinant $\underline{F}$ is the same as F_0 except that now the pth row and column are of the order ϵ^0. Therefore $\underline{F}$ should be calculated to the order ϵ, hence $F_0^{(3)}$ in (3.24) is

$$F_0^{(3)}(-K_m,-K_n) = \frac{d}{d\epsilon}\underline{F}(-K_m,-K_n), \tag{3.27}$$

where the derivative is calculated for $\epsilon = 0$. The calculation of (3.27) is straightforward, although lengthy and we will only give the final result.

$$F_0^{(3)}(-K_m,-K_n)$$

$$= \epsilon^3\frac{P(j)j_m^+ j_n^+}{8ij_m j_n}\left(\frac{2\langle\psi_m|V_{mp}|\psi_p\rangle\langle\psi_n|V_{np}|\psi_p\rangle\langle\psi_m|V_{mn}|\psi_n\rangle}{K_pK_mK_n j_p^+ j_m j_m^+ j_n j_n^+}\right) - \frac{\langle\psi_m|V_{mp}|\psi_p\rangle}{K_pK_m j_m j_m^+ j_p^+}\sum_{l\neq p,m}\frac{1}{K_l j_l}(\langle\psi_p|l|\psi_m\rangle$$

$$+\langle\psi_m|l|\psi_p\rangle) - \frac{\langle\psi_n|V_{np}|\psi_p\rangle}{K_pK_n j_n j_n^+ j_p^+}\sum_{l\neq p,n}\frac{1}{K_l j_l}(\langle\psi_p|l|\psi_n\rangle + \langle\psi_n|l|\psi_p\rangle), \tag{3.28}$$

where we have used the notation

$$\langle \psi_n | V_{np} | \psi_p \rangle = \int_0^\infty dr\, \psi_n V_{np} \psi_p \tag{3.29}$$

and

$$\langle \psi_p | l | \psi_m \rangle = \int_0^\infty dr\, \psi_p V_{pl} \int_0^\infty dr' f_l^-(r_>)\, \psi_l(r_<)\, V_{lm} \psi_m. \tag{3.30}$$

In the derivation of (3.28) we have also used the expression for k_3 in (3.25) and in the Appendix we show how it is calculated.

We can now calculate the inelastic background term. From (3.24) we obtain

$$S_{m,n} \sim \frac{(\beta_m \beta_n)^{1/2}}{k-\kappa}\left[1-(k-\kappa)\frac{F_0^{(3)}(-K_m,-K_n)}{F_1^{(0)}\beta_m \beta_n}\right]^{1/2}; \tag{3.31}$$

hence,

$$b_{m,n} = -\frac{1}{2}\frac{F_0^{(3)}(-K_m,-K_n)}{F_1^{(0)}(\beta_m \beta_n)^{1/2}}. \tag{3.32}$$

When we use

$$\langle \psi_m | V_{mp} | \psi_p \rangle = ij_m (2\beta_m K_m K_p j_p^+ j_p')^{1/2}, \tag{3.33}$$

the background term is finally

$$\begin{aligned} b_{m,n} = \frac{1}{2}\Bigg[& \frac{\langle \psi_m | V_{mn} | \psi_n \rangle}{2i(K_m K_n)^{1/2} j_m j_n} \\ & + \frac{2^{1/2} j_n^+}{8(\beta_n K_p K_m j_p^+ j_p')^{1/2} j_m j_n} \\ & \times \sum_{l \neq p,m} \frac{\langle \psi_p | l | \psi_m \rangle + \langle \psi_m | l | \psi_p \rangle}{K_l j_l} \\ & + \frac{2^{1/2} j_m^+}{8(\beta_m K_p K_n j_p^+ j_p')^{1/2} j_m j_n} \\ & \times \sum_{l \neq p,n} \frac{\langle \psi_p | l | \psi_n \rangle + \langle \psi_n | l | \psi_p \rangle}{K_l j_l} \Bigg]. \end{aligned} \tag{3.34}$$

The first term in (3.34) we recognize as the ordinary distorted wave approximation for the scattering from the channel m in the channel n. The last two terms come from the coupling between the open channels m and n and the channel where the unperturbed bound state is. It should be pointed out that although the matrix elements (3.30) are second order in V their overall contribution to (3.14) is first order since these terms are divided by $\sqrt{\beta}$ which is first order in V.

4. TWO-STATE FORMULA

Let us apply the results to the simplest case: the two-state problem. At the same time we will compare the obtained results with the Feshbach formalism. We assume that the channel 1 is open and that the channel 2 is closed. Therefore there is only elastic collision in channel 1, with the possibility of the internal excitations in channel 2. The second-order correction to the resonance level, if the pth bound state of the channel 2 is excited, is now

$$\begin{aligned} k_2 = & -\frac{1}{2K_2 K_1 j_1 j_2' j_2^+} \\ & \times\Bigg(\int_0^\infty \psi_2 V_{21} \psi_1\, dr \int_r^\infty f_1^- V_{12} \psi_2\, dr' \\ & + \int_0^\infty dr\, \psi_2 V_{21} f_1^- \int_0^r dr'\, \psi_1 V_{12} \psi_2\Bigg), \end{aligned} \tag{4.1}$$

and the residue is

$$\beta_1 = -\frac{1}{2K_1 K_2 j_1^2 j_2' j_2^+}\left(\int_0^\infty dr\, \psi_1 V_{12} \psi_2\right)^2. \tag{4.2}$$

The results are in close analogy with the expression obtained from the Feshbach theory.[15] However, this is to be expected since the difference between the two approaches becomes evident when the complete set of the functions is introduced, and for the derivation of (4.1) and (4.2) such a set was not necessary. For example, when the present theory is applied to the ordinary perturbation problem in a one channel case, then the first-order correction to any bound state (or a resonance) is given in the form similar to a known expression in the Rayleigh–Schrödinger perturbation theory.[12] However, the second-order correction in the RS theory is given as the sum over the complete set of eigenfunctions of the unperturbed Hamiltonian, but within the present theory this sum is replaced by an expression involving only the state which is being perturbed.

In the derivation of (4.1) and (4.2) we did not require the use of the complete set of the unperturbed eigenfunctions, hence, the results of the two procedures are similar. (The only difference is that in the Feshbach formalism one makes perturbation expansion of energy while here we expand the wavenumber.) The difference between the two approaches becomes evident if the degenerate case is treated[10] or if higher-order corrections to the resonance level are calculated.

Let us now turn our attention to the background term. For a two-level system the formula (3.19) becomes

$$\begin{aligned} b_{1,1} = & \frac{j_1^+}{j_1} - \frac{j_1^+}{j_1} \\ & \times\Bigg[\frac{\partial}{\partial k}\left(\frac{j_2^+}{4K_1 K_2 j_2' j_1 j_1^+}\left(\int_0^\infty dr f_2^- V_{21} \psi_1\right)^2\right) \\ & + \frac{1}{4K_1 K_2 j_2^+ j_1 j_1^+}\Bigg(\int_0^\infty dr\, \psi_1 V_{12} \psi_2 \int_0^r f_2^+ V_{21} \psi_1\, dr' \\ & + \int_0^\infty \psi_1 V_{12} f_2^+\, dr \int_r^\infty \psi_2 V_{21} \psi_1\, dr'\Bigg) + \beta_1 \frac{j_1'}{j_1^+}\Bigg]. \end{aligned} \tag{4.3}$$

The first term we recognize as the unperturbed S matrix in the channel 1. The second term is a correction to the elastic S matrix, coming from the interaction with channel 2. This term can now be compared with the analog in the Feshbach formalism. We find that the Feshbach form of $b_{1,1}$, if the relevant equations are solved in the first-order distorted wave approximation[15] looks like

$$b_{1,1} - \frac{j_1^+}{j_1} \sim \sum_{j \neq p} \int \frac{(\langle \psi_1 | V_{12} | \psi_j \rangle)^2}{k^2 - k_j^2}, \tag{4.4}$$

where the sum/integral extends over the complete set of the functions of the channel 2. Hence the formula (4.4) involves

the use of the complete set of functions for the background term, which in Eq. (4.3) is not necessary. The sum in (4.4) is essentially replaced by a term which involves taking the derivatives of the unperturbed solutions with respect to k. Since it is not obvious how these derivatives can be calculated, we will briefly discuss the properties of the solutions of the one channel problem.

5. DISCUSSION

In the expressions for the perturbation coefficients of the poles, residues, and the background term, it is assumed that we know the complete solution of the unperturbed Hamiltonian. By complete, we understand that we know the regular and irregular solutions together with the Jost functions and their derivatives with respect to k. Therefore it would be appropriate to review some of the properties of the solutions of the radial Schrödinger equation

$$\psi'' = (V - k^2)\,\psi, \tag{5.1}$$

which is a representative of the uncoupled set of equations (2.1). The channel energy is here represented with k^2. However, we should first recall that the derivatives with respect to k of the Jost functions and the matrix elements in Eq. (3.17) are not the derivatives with respect to the channel wavenumbers in (2.1). Therefore, we should transform d/dk in these cases by

$$\frac{d}{dk} = \frac{dK_n}{dk}\frac{d}{dK_n} = \frac{k}{K_n}\frac{d}{dK_n} \tag{5.2}$$

in a particular channel n. In what follows we will assume derivatives with respect to the channel wavenumber k in (5.1) which must not be confused with k in (5.2).

Let us restrict our discussion of (5.1) to a particular set of potentials which occur in atomic collisions. However, the theory is of general validity, and for the potentials other than those discussed here, one should appropriately modify the relevant steps.

A typical potential in atomic collisions has a hard core[16] of the type

$$\lim_{r\to R} V(r) = V(R), \tag{5.3}$$

while for $r<R$, $V(r)$ is infinite. In such a case the regular solution of (5.1) is defined with the initial values

$$\psi(R) = 0, \quad \psi'(R) = 1. \tag{5.4}$$

On the other hand, the two irregular solutions $f^{\pm}$, defined as

$$\lim_{r\to\infty} f^{\pm}(r) = \exp(\mp ikr) \tag{5.5}$$

are finite in the limit $r\to R$. From the definition of the Jost functions,[17]

$$j^{\pm} = (1/2ik)(\psi f'^{\pm} - \psi' f^{\pm}), \tag{5.6}$$

we obtain

$$f^{\pm}(R) = -2ikj^{\pm}. \tag{5.7}$$

When k^2 is positive, i.e., k^2 correspond to an open channel, we can easily find the relevant quantities entering the expression for the background term (3.17). For example, the derivative of the wavefunction with respect to k satisfies the differential equation

$$\dot\psi'' = -2k\psi + (V - k^2)\,\dot\psi, \tag{5.8}$$

where $\dot\psi \equiv d\psi/dk$. The regular solution of (5.8) is

$$\dot\psi = i\int_R^r G(r, r')\,\psi(r')\,dr', \tag{5.9}$$

where $G(r, r')$ is given by (2.5). The derivatives of the Jost functions with respect to k can be calculated from (5.6) and they are

$$\dot j^{\pm} = -(k^{-1} \pm ir)j^{\pm} - (\exp(\mp ikr)/2ik) \times(\pm i\psi \pm ik\dot\psi + \dot\psi'), \tag{5.10}$$

where we have taken the limit $r\to\infty$. Therefore, in principle when $k^2>0$ there is no basic difficulty in obtaining all the relevant quantities entering the background term.

It is not at all evident that this is the case for the bound states, and this point needs little more discussion. Let us therefore assume that $k^2<0$, and its value corresponds to one of the bound states of (5.1). In such a case we can find a useful expression which relates the Jost functions to the normalization constant of the regular wavefunction. Multiplying (5.8) with ψ and (5.1) with $\dot\psi$, and subtracting these two equations, we obtain

$$\frac{d}{dr}(\dot\psi\psi' - \dot\psi'\psi) = 2k\psi^2. \tag{5.11}$$

Integrating the equation and using the relationship

$$\psi = j^+f^- + j^-f^+, \tag{5.12}$$

we obtain the well-known relationship[17]

$$\dot ij^-j^+ = \int_R^\infty \psi^2\,dr. \tag{5.13}$$

The last formula means that if we know ψ and $\dot j^-$ then j^+ can also be obtained. This fact will be useful a little later. However, let us discuss j^-. From (5.6) we obtain in the limit $r\to\infty$,

$$j^- = (\exp(ikr)/2ik)(ik\psi - \psi'), \tag{5.14}$$

where it is assumed that k is positive imaginary. In general, when k is not a bound-state wavenumber, the regular wavefunction will exponentially increase for large r as $\psi \sim \exp(-ikr)$, hence, no significant figures in the bracket of (5.14) will cancel, which means numerical stability. This also means that j^- can be calculated to any arbitrary accuracy, without too much numerical difficulty. Since this is the case, then also $\dot j^-$ and $\ddot j^-$ can be calculated from (5.10) without too much difficulty.

Similarly, we can show that f^- can be calculated from (5.1) by the backward integration and the procedure is numerically stable. Therefore, from now on we assume that j^-, f^-, and ψ are known functions for the bound states.

That the same procedure does not apply for j^+ can easily be verified by calculating (5.6) for $r\to\infty$. In such a case we have

$$j^+ = (\exp(-ikr)/2ik)(-ik\psi - \psi'), \tag{5.15}$$

and since $\psi\sim\exp(-ikr)$, when k is not at the bound state, the significant figures in (5.15) will cancel. This means that j^+ cannot be calculated from (5.15). However, we can use

(5.13) to obtain j^+ and since j^- and ψ are known, which was shown earlier, the procedure is numerically stable.

We still need j^+, which cannot be obtained from (5.13) since this expression holds only for the bound states and is not valid in its small neighborhood. Therefore, we should look for a different way to calculate j^+. Let us differentiate (5.12) with respect to k, and use (5.9) for $\dot{\psi}$, hence,

$$\dot{j}^+ f^- + j^+ \dot{f}^- = i \int_R^r G(r, r')\, \psi(r')\, dr' - \dot{j}^- f^+, \quad (5.16)$$

and if we use (2.5),

$$\dot{j}^+ f^- + j^+ \dot{f}^- = if^- \int_R^r f^+ \psi\, dr' - if^+ \int_R^r f^- \psi\, dr' - \dot{j}^- f^+. \quad (5.17)$$

In the limit $r \to \infty$ we can show, using (5.13), that the second and the third term cancel even though f^+ is an exponentially increasing function. Hence we are left with

$$\dot{j}^+ + irj^+ = i \int_R^r f^+(r')\, \psi(r')\, dr', \quad r \to \infty \quad (5.18)$$

or

$$\dot{j}^+ = ij^+ \lim_{r\to\infty} \left(\int_R^r f^+ f^-\, dr' - r \right), \quad (5.19)$$

which relates $\dot{j}^+$ to j^+ and the irregular solutions f^+ and f^-. It can be easily shown that the procedure (5.19) is numerically stable.

However, for the calculation of $\dot{j}^+$ we need $f^+(r)$, but the analysis shows that it is not possible to obtain $f^+(r)$ from the straightforward integration of (5.1). The reason is simple: f^+ is an exponentially increasing function and is not uniquely defined by the boundary condition (5.5). We can nevertheless obtain f^+ up to an undetermined constant by starting from the Wronskian

$$f^+ f'^- - f'^+ f^- = 2ik, \quad (5.20)$$

and if this expression is treated as the first-order nonhomogeneous equation for f^+, the solution is

$$f^+ = -f^- \left(C - 2ik \int_x^r \frac{dr'}{(f^-)^2} \right), \quad (5.21)$$

where C is a constant determined at the lower bound. Since k corresponds to a bound state and $f^- \sim \psi$, the integrand is singular at $r = R$ and any node of ψ.

Using (5.4) and (5.1) it can be shown that near $r = R$ the regular wavefunction has expansion

$$\psi(r) \sim \Delta + O(\Delta^3), \quad \Delta = r - R. \quad (5.22)$$

Similarly, near any node of ψ we have

$$\psi(r) \sim \psi'(r_n)\Delta_n + O(\Delta_n^3), \quad \Delta_n = r - R_n, \quad (5.23)$$

where R_n is the nth node of ψ. Let us now write for C

$$C = 2ik\,(j^+)^2 \left(\frac{1}{x - R} + \sum_{n=1}^{N} \frac{1}{x - R_n} \frac{1}{(\psi'(R_n))^2} \right), \quad (5.24)$$

where N is the number of the nodes of ψ. The solution f^+ now reads

$$f^+ = -2ikj^+ \psi \left[\frac{1}{r - R} + \sum_{n=1}^{N} \frac{1}{r - R_n} \frac{1}{(\psi'_n)^2} + \int_R^r dr' \times \left(\frac{1}{(r' - R)^2} + \sum_{n=1}^{N} \frac{1}{(r' - R_n)^2} \frac{1}{(\psi'_n)^2} - \psi^{-2} \right) \right], \quad (5.25)$$

where we have replaced x by R. It can easily be shown that f^+ satisfies (5.1). We can also show that in the limit $r \to R$ the value of f^+ is

$$f^+ = -2ikj^+ \psi'(R) = -2ikj^+, \quad (5.26)$$

which is equal to (5.7). Therefore (5.25) indeed represents the solution f^+ for the bound states. However, we can also show that any function

$$F^+ = f^+ + Cf^-, \quad (5.27)$$

where f^+ is defined in (5.25) and C is an arbitrary constant, also satisfies the just mentioned conditions. Hence, F^+ is also the irregular solution of (5.1), satisfying the boundary condition (5.5). However, the background term (3.17) is invariant to the transformation (5.27), which can be proved by noting that j^+ transforms as

$$j^+ \to j^+ - Cj^+ j^-, \quad (5.28)$$

in which case all the elements in (3.17) containing C will cancel.

In fact the whole perturbation theory is invariant to the transformation (5.27), regardless of whether k belongs to the bound state or not. This comes out from the fact that $G(r, r')$, defined by (2.5), is invariant to the transformation (5.27); therefore, the Jost function (2.8) is also invariant. Since the Jost function (2.8) is the basis of the perturbation theory, then also the perturbation theory is invariant to the transformation (5.27). As the conclusion we can say that (5.25) can be used in the calculation of the background term, although it does not represent the unique irregular solution of (5.1).

ACKNOWLEDGMENT

The author is thankful to Professor D. Micha for many helpful discussions related to this work.

APPENDIX

Here we will calculate the third-order correction to the resonance wavenumber (2.11), which is needed in the derivation of the inelastic background term $b_{m,n}$ in (3.34). The coefficient k_3 can be obtained in two ways: either by directly calculating $d^3\kappa/d\epsilon^3$ for $\epsilon = 0$ from the implicit equation (2.9) or by calculating $d\kappa/d\epsilon$ for a finite ϵ and then looking for the coefficients in the expansion

$$\frac{d\kappa}{d\epsilon} = -\frac{\partial F}{\partial \epsilon} \Big/ \frac{\partial F}{\partial k} = k_1 + \epsilon k_2 + \epsilon^2 / 2k_3. \quad (A1)$$

The last procedure is sometimes more convenient and will be used here. Since $k_1 = 0$ it follows that $\partial F/\partial \epsilon$ is exactly zero for $\epsilon = 0$ hence $\partial F/\partial \epsilon$ starts with the order ϵ. On the other hand, $\partial F/\partial k$ is of the order ϵ^0 but the next higher is ϵ^2, as was shown in Sec. 3. Therefore looking for k_3 in (A1) is equivalent to finding $\partial F/\partial \epsilon$ to the order ϵ^2.

The derivative of determinant F is equal to the sum of determinants in which we take the derivative of each column separately. Therefore, if we take the derivative of the column m which does not correspond to the column in which the

diagonal element is zero in the unperturbed case (in our case this is the pth column), then

$$\frac{\partial \underline{F}^{(m)}}{\partial \epsilon} = \epsilon \underline{F}^{(m)}, \tag{A2}$$

where $\underline{F}^{(m)}$ is a determinant in which the mth and the pth column are of the order ϵ^0, except the elements (m,m) and (p,p). These elements are

$$(p,p) = \tfrac{1}{2}\epsilon(k_2 j_p' + J_{pp}^{(2)}) \tag{A3}$$

and

$$(m,m) = \epsilon J_{mm}^{(2)}. \tag{A4}$$

We now look for the contribution in $\underline{F}^{(m)}$ which is of the order ϵ, and this can be obtained by calculating $\partial \underline{F}^{(m)}/\partial\epsilon$ for $\epsilon = 0$. When this is done we find

$$\frac{\partial \underline{F}^{(m)}}{\partial \epsilon} = P(j)\left(-\frac{J_{pm}^{(1)} J_{mp}^{(2)}}{2 j_m} - \frac{J_{pm}^{(2)} J_{mp}^{(1)}}{j_m} + \sum_{l \neq p,m} \frac{J_{ml}^{(1)} J_{lp}^{(1)} J_{pm}^{(1)} + J_{mp}^{(1)} J_{pl}^{(1)} J_{lm}^{(1)}}{j_m j_l} \right). \tag{A5}$$

Similarly we can calculate (A5) when $m = p$. In such a case we find

$$\frac{\partial F^{(p)}}{\partial \epsilon} = P(j)\left(\frac{J_{pp}^{(3)}}{2} - \sum_{m \neq p} \frac{J_{pm}^{(1)} J_{mp}^{(2)}}{j_m} - \frac{1}{2} \sum_{m \neq p} \frac{J_{mp}^{(1)} J_{pm}^{(2)}}{j_m} + \sum_{i \neq p} \sum_{m \neq p,i} \frac{J_{ip}^{(1)} J_{pm}^{(1)} J_{mi}^{(1)}}{j_i j_m} \right). \tag{A6}$$

Therefore we now have

$$\frac{\partial F}{\partial \epsilon} = \epsilon^2 \sum_m \frac{\partial F^{(m)}}{\partial \epsilon}. \tag{A7}$$

Since $\partial F/\partial k$ is

$$\frac{\partial F}{\partial k} = j_p' P(j), \tag{A8}$$

the third-order coefficient k_3 is from (A1),

$$k_3 = -\frac{2}{j_p'}\left(\frac{J_{pp}^{(3)}}{2} - \frac{3}{2} \sum_{m \neq p} \frac{J_{pm}^{(2)} J_{mp}^{(1)} + J_{pm}^{(1)} J_{mp}^{(2)}}{j_m} + 3 \sum_{m \neq p} \sum_{i \neq p,m} \frac{J_{mi}^{(1)} J_{ip}^{(1)} J_{pm}^{(1)}}{j_m j_i} \right). \tag{A9}$$

The expression for k_3 can be much simplified if we use the explicit forms of the matrix elements in (A9). For example $J_{pm}^{(1)}$ is given by (3.12). The higher-order elements are obtained by iterating Eq. (2.4) and putting the series in (2.8). After some algebra, when the following form of the Green's function (2.5) is used,

$$G(r,r') = (\psi(r') f^-(r) - \psi(r) f^-(r'))/j, \tag{A10}$$

we obtain the final form of k_3,

$$k_3 = -\frac{3}{4iK_p j_p'} \sum_{m,k} \frac{1}{K_m K_k j_m j_k} \int_0^\infty dr \times f_p^- V_{pm} \int_0^\infty dr' f_m^-(r_>)\psi_m(r_<) V_{mk}(r') \times \int_0^\infty dr'' f_k^-(r_>')\psi_k(r_<') V_{kp}(r'') \psi_p(r''). \tag{A11}$$

[1] A. J. F. Siegert, Phys. Rev. **56**, 750 (1939).
[2] A. Donnachie, Rep. Prog. Phys. **36**, 695 (1973).
[3] For more details about the general properties of the resonance poles see R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966).
[4] G. Breit and E. P. Wigner, Phys. Rev. **49**, 519 (1936).
[5] H. A. Bethe and G. Placzek, Phys. Rev. **51**, 450 (1937).
[6] P L. Kapur and R. Peierls, Proc. R. Soc. London Ser. A **166**, 277 (1938).
[7] H. Feshbach, Ann. Phys. (NY) **5**, 357 (1958).
[8] L. Eisenbud and E. P. Wigner, Phys. Rev. **72**, 29 (1947).
[9] L. Fonda and R. G. Newton, Ann. Phys. (NY) **10**, 490 (1960).
[10] S. Bosanac, "Perturbation theory of inelastic resonances," J. Math. Phys. **23**, 2131 (1982).
[11] S. Bosanac, J. Math. Phys. **21**, 1881 (1980).
[12] S. Bosanac, Fizika (Zagreb) **12**, 77 (1980).
[13] R. G. Newton, Ann. Phys. (NY) **4**, 29 (1958).
[14] R. G. Newton, J. Math. Phys. **2**, 188 (1961).
[15] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Claredon, Oxford, 1971), Chap. XIII.4.
[16] S. Bosanac, Croat. Chem. Acta **49**, 471 (1977).
[17] V. de Alfaro and T. Regge, *Potential Scattering* (North-Holland, Amsterdam, 1965).

Inverse scattering with coinciding-pole reflection coefficients

Kay R. Pechenick and Jeffrey M. Cohen
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 22 July 1981; accepted for publication 2 September 1981)

In inverse scattering theory, algorithms for solving the Gel'fand–Levitan equation normally break down when poles in the reflection coefficient coincide. Here we present a method for treating an arbitrary number of coinciding poles. We give the first explicit solutions for 3, 4, 5, 6, 8, and 10 poles.

PACS numbers: 03.80. + r, 94.20. − y

1. INTRODUCTION

The Abel integral equation has been the basis for most work on ionospheric structure determination during the last half century. Using this approximate method, ionospheric electron densities have been computed from scattering data. With the same data it is possible to employ an exact full-wave method, based on the Gel'fand–Levitan equation,[1] to obtain a much improved determination of the ionospheric electron density. Since the data used is identical for the approximate and full-wave theories, there is no need to modify experimental equipment; the difference in treatment is essentially computational.

In principle, the full-wave inverse scattering method is exact. However, in practice, approximate analytic or numerical methods are normally employed to solve the Gel'fand–Levitan equation. To circumvent the possibility of round-off errors, numerical instabilities, etc., in solving the Gel'fand–Levitan equation numerically, we have solved the equation exactly, using a generalization of Kay's[2] procedure for rational function reflection coefficients. Previous attempts along these lines have given usable results when the number of poles in the reflection coefficient is not too large (3 poles[3]; 1 pole[4]; 3 poles[5]; 1, 2, and 3 poles[6]),

In previous communications[7,8] we presented a generalized procedure for finding exact solutions to the Gel'fand–Levitan equation in inverse scattering theory. Our procedure is applicable to the case in which the reflection coefficient $r(k)$ is a rational function of the wave number k. Using our procedure, we can calculate the scattering kernel $K(x,t)$ from $r(k)$, and we obtain the potential $V(x)$, which is related to the scattering kernel by the equation

$$V(x) = 2\frac{d}{dx}K(x,x). \tag{1}$$

One step in the procedure involves the solution of n simultaneous linear equations (with complex coefficients), where n is the number of poles of $r(k)$. (There is one set of n simultaneous equations for each value of the distance x.) The procedure breaks down whenever two or more poles coincide, because then the corresponding rows of the determinant of the coefficients are equal, so the determinant is zero, for all values of x. Here we present a modified procedure which overcomes this difficulty.

2. GEL'FAND–LEVITAN EQUATION

In this section we solve the Gel'fand–Levitan equation for the case

$$r(k) = \text{const}/(k - k_1)^n, \tag{2}$$

in which there are n coinciding poles. Because of the requirement[3] that $r(0) = -1$, the constant in Eq. (2) must be $-(-k_1)^n$. Thus we have

$$r(k) = -(-k_1)^n/(k - k_1)^n. \tag{3}$$

The additional requirement[3,9] that $r^*(k) = r(-k)$ (for all real k) forces k_1 to be purely imaginary. With these two conditions satisfied, it is automatically true that $|r(k)| \leqslant 1$ for all real k, another necessary property[2,3] of the reflection coefficient. Finally, k_1 must be in the lower half-plane, because we assume that $r(k)$ is analytic[2,9] for all k in the upper half-plane.

As in Refs. 7 and 8, we begin with the Gel'fand–Levitan equation

$$R(x+t) + K(x,t) + \int_{-\infty}^{x} K(x,z)R(z+t)dz = 0, \tag{4}$$

which can be rewritten as

$$R_1(x+t) + K_1(x,t) + \int_{-t}^{x} K_1(x,z)R_1(z+t)\,dz = 0, \tag{5}$$

where

$$R(x) = R_1(x)\theta(x) \tag{6}$$

and

$$K(x,t) = K_1(x,t)\theta(x+t). \tag{7}$$

We again let

$$R(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikx}r(k)\,dk. \tag{8}$$

Substituting (3) into (8), and using (6), we find that

$$R_1(x) = Bx^{n-1}e^{-ik_1x}, \tag{9}$$

where

$$B = -(ik_1)^n/(n-1)! \tag{10}$$

As in Refs. 7 and 8, we assume

$$K_1(x,t) = \sum_{\alpha} f_\alpha(x)e^{a_\alpha t}, \tag{11}$$

where the summation is over integer values of α from 1 to n and from -1 to $-n$; that is, the summation is over $2n$ values of α.

We now substitute (9) and (11) into (5). The result may be written in the form

$$E + F + J = 0, \tag{12}$$

in which

 0022-2488/83/010115-05$02.50

$$E = B(x+t)^{n-1}e^{-ik_1(x+t)}, \tag{13}$$

$$F = \sum_\alpha f_\alpha(x)e^{a_\alpha t}, \tag{14}$$

and

$$J = \int_{-t}^{x} \sum_\alpha f_\alpha(x)e^{\alpha_\alpha z}B(z+t)^{n-1}e^{-ik_1(z+t)}dz. \tag{15}$$

To evaluate the integral in (15), we make the substitution $y = z + t$ and use the fact that

$$\int y^{n-1}e^{by}\,dy = \sum_{r=1}^{n}(-1)^{r+1}b^{-r}\frac{(n-1)!}{(n-r)!}y^{n-r}e^{by}. \tag{16}$$

We find that

$$J = C + D, \tag{17}$$

where

$$C = B\sum_\alpha\left\{f_\alpha(x)e^{-\alpha_\alpha t}(n-1)!\times\left[\sum_{r=1}^{n}\frac{(-1)^{r+1}}{(\alpha_\alpha - ik_1)^r}\frac{(x+t)^{n-r}}{(n-r)!}e^{(\alpha_\alpha - ik_1)(x+t)}\right]\right\} \tag{18}$$

and

$$D = -B\sum_\alpha f_\alpha(x)e^{-a_\alpha t}(-1)^{n+1}(n-1)!(a_\alpha - ik_1)^{-n}. \tag{19}$$

From (12) and (17), we have

$$C + D + E + F = 0. \tag{20}$$

We shall show that there is a solution to (20) for which $C + E = 0$ and $D + F = 0$ simultaneously.

If we set $a_{-\alpha} = -a_\alpha$ in the equation $D + F = 0$, the equation becomes

$$\sum_\alpha\left[f_\alpha(x)e^{a_\alpha t} + (ik_1)^n f_{-\alpha}(x)e^{a_\alpha t}\frac{(-1)^{n+1}}{(-a_\alpha - ik_1)^n}\right] = 0. \tag{21}$$

Equation (21) will certainly be satisfied if we let

$$f_\alpha(x) + \frac{(ik_1)^n(-1)^{n+1}}{(-a_\alpha - ik_1)^n}f_{-\alpha}(x) = 0 \tag{22}$$

for all α. Thus,

$$f_{-\alpha}(x) = (1 + a_\alpha/ik_1)^n f_\alpha(x) \tag{23}$$

and

$$f_\alpha(x) = (1 - a_\alpha/ik_1)^n f_{-\alpha}(x) \tag{24}$$

for all α. Equations (23) and (24) are consistent [for nonzero $f_\alpha(x)$] only if

$$k_1^{2n}/(a_\alpha^2 + k_1^2)^n = 1. \tag{25}$$

Solving for a_α^2 in (25), we obtain

$$a_\alpha^2 = k_1^2(\beta^{-m} - 1), \tag{26}$$

where

$$\beta = \exp(2\pi i/n) \tag{27}$$

and m takes on the integer values from 1 to n. Thus, we may let a_α and $a_{-\alpha}$ be the two square roots of $k_1^2(\beta^{-\alpha} - 1)$, for positive α.

Turning now to the equation $C + E = 0$, we rewrite it

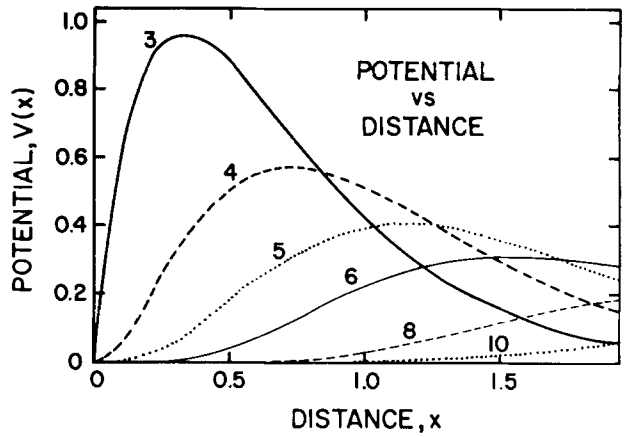


FIG. 1. Potential $V(x)$ vs distance x for 3, 4, 5, 6, 8, and 10 coinciding poles. All poles are at $-i$. In this and all other graphs, the potential is zero for negative x.

in the form

$$\sum_{s=0}^{n-1}\binom{n-1}{s}x^{n-1-s}t^s + \sum_\alpha\left(f_\alpha(x)e^{\alpha_\alpha x}\times\left\{\sum_{r=1}^{n}\frac{(-1)^{r+1}(n-1)!}{(\alpha_\alpha - ik_1)^r(n-r)!}\left[\sum_{s=0}^{n-r}\binom{n-r}{s}x^{n-r-s}t^s\right]\right\}\right) = 0. \tag{28}$$

The left side of Eq. (28) is a power series in t in which the coefficients are functions of x. Equation (28) is satisfied only if each coefficient is zero. For each power of t, this condition may be written as

$$\frac{x^{n-s-1}}{(n-s-1)!} + \sum_\alpha\left\{f_\alpha(x)e^{\alpha_\alpha x}\frac{(-1)^{n-s-1}}{(\alpha_\alpha - ik_1)^{n-s}}\times\left[\sum_{q=0}^{n-s-1}\frac{[(ik_1 - \alpha_\alpha)x]^q}{q!}\right]\right\} = 0, \tag{29}$$

which must be satisfied for $s = 0, 1, 2, \ldots, n-1$.

Using (23), we rewrite (29) in the form

$$\sum_{\alpha=1}^{n}M_{j\alpha}(x)f_\alpha(x) = \frac{x^{n-j}}{(n-j)!}, \tag{30}$$

$j = 1, 2, 3, \ldots, n$, where

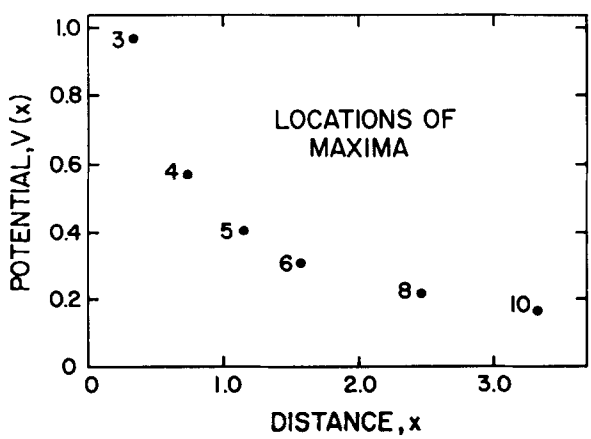


FIG. 2. Locations of maxima in potential vs distance graphs. Numbers represent the number of coinciding poles. All poles are at $-i$.

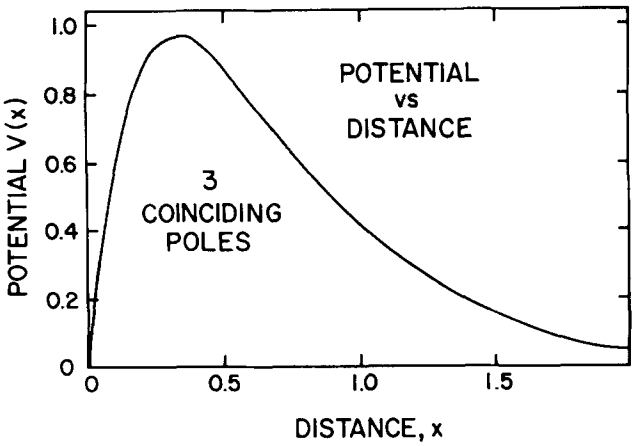


FIG. 3. Potential $V(x)$ vs distance x for 3 coinciding poles at $-i$.

$$M_{j\alpha}(x) = \frac{e^{a_\alpha x}}{(ik_1 - a_\alpha)^{n-j+1}} \sum_{q=0}^{n-j} \frac{[(ik_1 - a_\alpha)x]^q}{q!} + \frac{(ik_1 + a_\alpha)^{j-1}}{(ik_1)^n} e^{-a_\alpha x} \sum_{q=0}^{n-j} \frac{[(ik_1 + a_\alpha)x]^q}{q!}. \quad (31)$$

Equation (30) is a set of n simultaneous linear equations in n unknowns—the functions $f_\alpha(x)$. (That is, there is a set of n equations for each value of x.) The matrix $M_{j\alpha}$ is different from the matrix which would have been obtained using our original procedure.[7,8] The determinant of $M_{j\alpha}$ is nonzero.

After we have calculated the $f_\alpha(x)$ for a set of values of x, we calculate $K_1(x,x)$, which is given by

$$K_1(x,x) = \sum_{\alpha=1}^{n} f_\alpha(x)\left[e^{a_\alpha x} + (1 - ia_\alpha/k_1)^n e^{-a_\alpha x}\right]. \quad (32)$$

Equation (32) was derived by substituting (23) into (11).

The potential $V(x)$ is found from $K(x,x)$ using Equation (1). We have plotted graphs of V vs x for different values of n, all with $k_1 = -i$. There is no loss of generality in restricting k_1 to be $-i$ because multiplying k_1 by a positive real number only changes the scale. If ξ is a positive real number, then from Eq. (26),

$$a_\alpha(\xi k_1) = \xi a_\alpha(k_1). \quad (33)$$

That is, when we multiply k_1 by a constant, the a_α are multiplied by the same constant. (Here and in what follows we have inserted additional arguments into a_α, $M_{j\alpha}$, and other quantities to indicate the dependence on k_1.) By substituting Eq. (33) into (31) we deduce that

$$M_{j\alpha}(x/\xi,\xi k_1) = \xi^{-n+j-1} M_{j\alpha}(x,k_1). \quad (34)$$

Then, using Eqs. (30) and (34), we obtain

$$f_\alpha(x/\xi,\xi k_1) = \xi f_\alpha(x,k_1). \quad (35)$$

It follows from Eqs. (32) and (35) that

$$K_1(x/\xi,x/\xi,\xi k_1) = \xi k_1(x,x,k_1). \quad (36)$$

Because the potential V involves a derivative of K,

$$V(x/\xi,x/\xi,\xi k_1) = \xi^2 V(x,x,k_1). \quad (37)$$

In other words, if the poles move farther away from the origin by a factor of ξ, then the peak in the potential becomes higher by a factor of ξ^2 and moves to a value of x which is smaller by a factor of ξ.

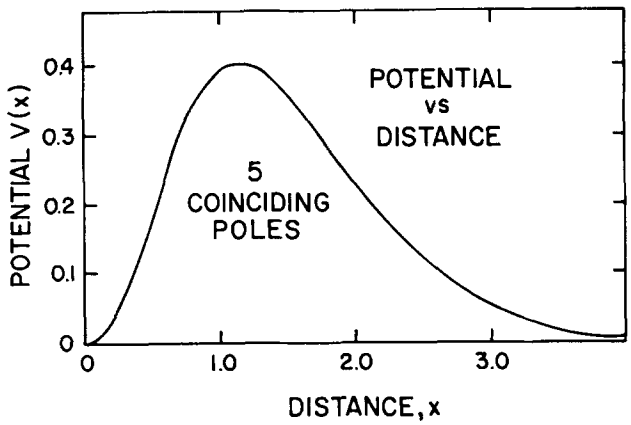


FIG. 5. Potential $V(x)$ vs distance x for 5 coinciding poles at $-i$.

3. RESULTS

In Fig. 1 we have plotted the potential versus distance for $n = 3, 4, 5, 6, 8$, and 10, all with $k_1 = -i$. In each case the potential is zero for all negative x, as can be seen from Eqs. (1) and (7). The potential rises to a maximum at a positive value of x and then decreases. As n increases, the peak in the potential moves downward and to larger x values. This can be seen more clearly in Fig. 2, which shows the position of the maximum for $n = 3, 4, 5, 6, 8$, and 10.

In Fig. 1, all curves are drawn to the same scale and thus can be directly compared. However, the scale chosen, al-

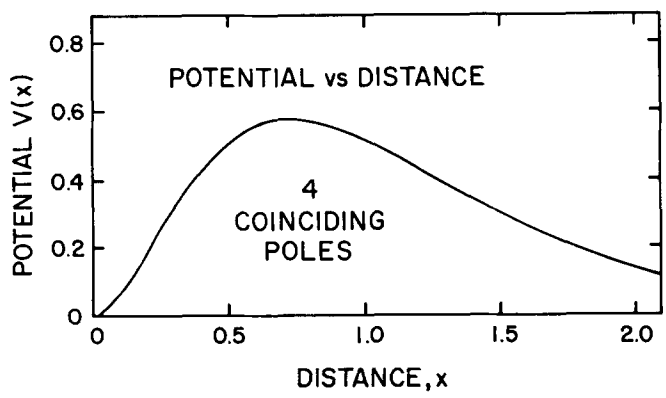


FIG. 4. Potential $V(x)$ vs distance x for 4 coinciding poles at $-i$.

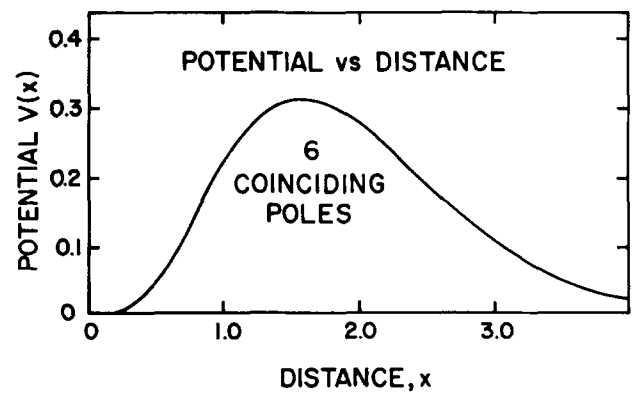


FIG. 6. Potential $V(x)$ vs distance x for 6 coinciding poles at $-i$.

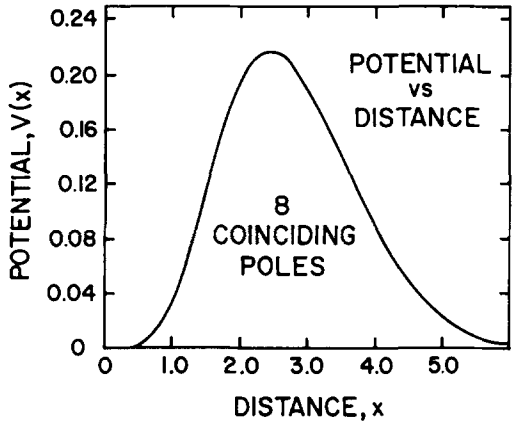


FIG. 7. Potential $V(x)$ vs distance x for 8 coinciding poles at $-i$.

though appropriate for the smaller values of n, is less advantageous for the larger values of n. Therefore, we have plotted the individual curves to more appropriate scales in Figures 3–8. Note that with increasing n, the curves rise more and more slowly as x increases from zero.

In graphs of potential V vs distance x for n coinciding poles at $k_1 = -i$, the value of x at which V is a maximum depends on n. (We let x_0 denote this value of x.) By altering the position of the poles, we can change x_0 to any positive value. For 3 poles at $k_1 = -i$, $x_0 = 0.33$, but if $k_1 = -0.33i$, $x_0 = 1$. For the 6-pole case, $x_0 = 1.57$ if $k_1 = -i$, while $x_0 = 1$ if $k_1 = -1.57i$. For 10 coinciding poles, $x_0 = 3.325$ if $k_1 = -i$, while $x_0 = 1$ if $k_1 = -3.325i$. Figure 9 shows 3-, 6-, and 10-pole cases, all with $x_0 = 1$.

In Fig. 10 we have plotted a 3-pole case with $x_0 = 1$ and a 10-pole case with $x_0 = 3$. (For the 10-pole case, $k_1 = -1.1083i$.)

We also considered the effect on the potential of starting with the case of 3 coinciding poles, all at $-i$, and then moving the poles slightly apart. Specifically, we considered

$$r(k) = \frac{k_1 k_2 k_3}{(k-k_1)(k-k_2)(k-k_3)}.$$

Case (a): $k_1 = k_2 = k_3 = -i$.

Case (b): $k_1 = -0.999i$, $k_2 = -i$, $k_3 = -1.001i$.

Case (c): $k_1 = -0.99i$, $k_2 = -i$, $k_3 = -1.01i$.

For case (a), we used the new procedure described in this paper. For cases (b) and (c), we used the procedure described

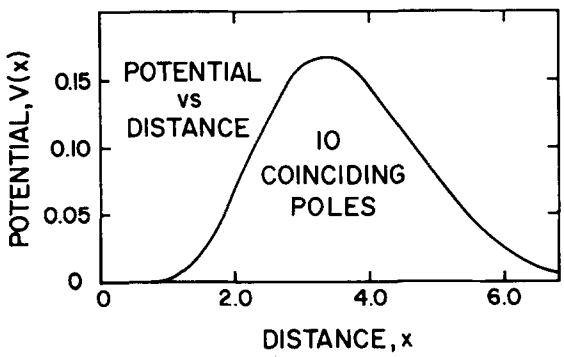


FIG. 8. Potential $V(x)$ vs distance x for 10 coinciding poles at $-i$.

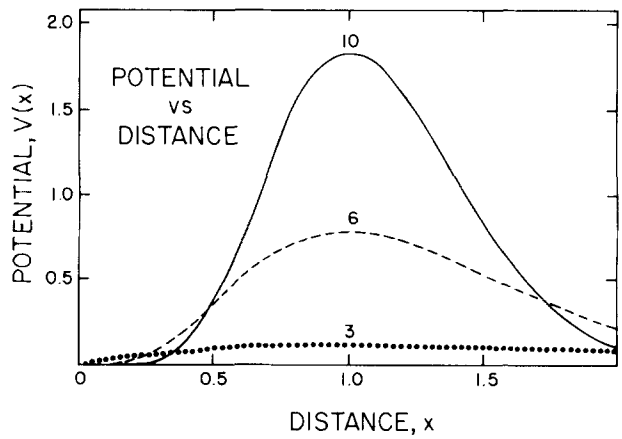


FIG. 9. Potential $V(x)$ vs distance x for three coinciding pole reflection coefficients: 3 poles at $-0.33i$, 6 poles at $-1.57i$, and 10 poles at $-3.325i$. Numbers represent the number of coinciding poles.

in our previous communications.[7,8] The results were that for the same values of x, the differences between the potentials in cases (a) and (b) were at most in the 5th significant digit. The differences between cases (a) and (c) were at most in the 4th significant digit. [If cases (b) and (c) were plotted to the same scale as case (a) in Fig. 3, the differences between (a), (b), and (c) would not be observable.] Thus in this case the procedure is stable.

4. DISCUSSION

In this communication we have presented a new inverse-scattering procedure for treating an arbitrarily large number of coinciding poles in the reflection coefficient. (In previous communications we treated arbitrarily large numbers of noncoinciding poles.) Since our procedure for solving the Gel'fand–Levitan equation is exact, we have circumvented the difficulties which often arise in numerical solutions, such as numerical instabilities and the use of excessive amounts of computer time and memory. We expect that our procedure will give rise to highly accurate on-line inverse scattering computational capability which should make possible improved ionosondes for ionospheric structure determination.

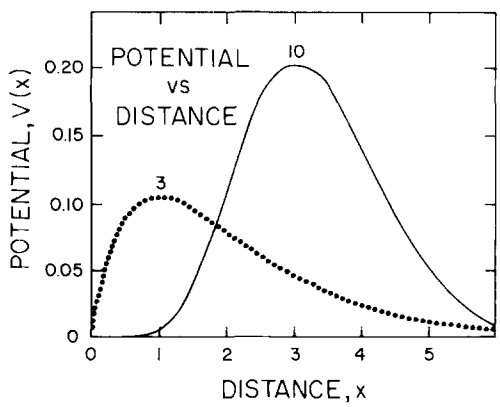


FIG. 10. Potential $V(x)$ vs distance x for two coinciding-pole reflection coefficients: 3 poles at $-0.33i$ and 10 poles at $-1.1083i$. Numbers represent the number of coinciding poles.

ACKNOWLEDGMENTS

One of us (J.M.C.) is indebted to Dr. Angelo J. Skalafuris for his hospitality at the Mathematics Research Center, where this work was begun. In addition, we would like to thank Prof. Peter Lax for his hospitality at the Courant Institute of Mathematical Sciences, and for emphasizing the importance of this work.

This work was supported in part by a grant from the Air Force Office of Scientific Research and in part by the National Science Foundation.

[1]M. Gel'fand and B. M. Levitan, Izv. Akad. Nauk SSSR, Ser. Math. **15**, 309 (1951), Am. Math. Soc. Transl. **1**, 253 (1955).
[2]Irvin Kay, Commun. Pure Appl. Math. **13**, 371 (1960).
[3]Saeyoung Ahn and Arthur K. Jordan, IEEE Trans. Antennas Propag. **24**, 879 (1976).
[4]H. E. Moses, Stud. Appl. Math. **60**, 177 (1979).
[5]H. E. Moses (unpublished).
[6]K. R. Pechenick and J. M. Cohen, University of Pennsylvania Report (1980).
[7]K. R. Pechenick and J. M. Cohen, Phys. Lett. A **82**, 156 (1981).
[8]K. R. Pechenick and J. M. Cohen, J. Math. Phys. **22**, 1513 (1981).
[9]I. Kay and H. E. Moses, Nuovo Cimento **3**, 276 (1956).

120,124

A constructive approach to bundles of geometric objects on a differentiable manifold[a)]

M. Ferraris and M. Francaviglia
Istituto di Fisica Matematica "J.-L. Lagrange," Università di Torino, Via C. Alberto 10, 10123, Torino, Italy

C. Reina
Istituto di Fisica, Università di Milano, Via Celoria 16, 20133, Milano, Italy

(Received 17 August 1981; accepted for publication 23 December 1981)

A constructive approach to bundles of geometric objects of finite rank on a differentiable manifold is proposed, whereby the standard techniques of fiber bundle theory are extensively used. Both the point of view of transition functions (here directly constructed from the jets of local diffeomorphisms of the basis manifold) and that of principal fiber bundles are developed in detail. These, together with the absence of any reference to the current functorial approach, provide a natural clue from the point of view of physical applications. Several examples are discussed. In the last section the functorial approach is also presented in a constructive way, and the Lie derivative of a field of geometric objects is defined.

PACS numbers: 04.20. − q, 02.40. + m

1. INTRODUCTION

Classical tensor calculus, developed during the latter part of the last century by Ricci and Levi-Civita, was soon found to be the most appropriate formalism for studying local physical laws in an invariant way. After its application to special and general relativity,[1] tensor calculus became a common tool in mathematical physics and the main formal link between geometry and physics itself.

As early as 1918 it was, however, discovered that certain local structures which are relevant both to physics and geometry do not have tensorial character, the most well-known example being, of course, given by connections [Levi-Civita (1917),[1a] Weyl (1918),[2] and Cartan (1923)[3]]. Early attempts to give definitions of "geometric objects" general enough to also include such nontensorial entities date back to the thirties [Schouten and Haantjes (1936)[4]], but a fully satisfactory and intrinsic definition was found only after the work of Nijenhuis during the fifties [Nijenhuis (1952),[5] (1960),[6] Haantjes and Laman (1953 a,b),[7] and Kuiper and Yano (1955)[8]]. More recently, the matter was reconsidered by Salvioli (1972),[9] who gave a natural and beautiful description grounded on a "functorial approach".

Roughly speaking, an object defined on a differentiable manifold is a geometric object if we know its transformation law for any change of local coordinates. Tensors are obviously geometric objects, but of a very restricted type; their transformation laws are in fact "homogeneous" and involve only the Jacobian matrix of the coordinate transformation. To allow more general objects like, for example, connections, higher derivatives of the coordinate transformation must be taken into account.

In recent years, owing to their greater generality, geometric objects other than tensors began to enter physical applications, because in many cases using objects more general than tensors is essential [see, e.g., Anderson (1967),[10] Krupka (1979a,b),[11] Kijowski and Tulczyjew (1978),[12] Prastaro (1980),[13] (1981),[14] Modugno (1981),[15] Pommaret (1978),[16] Ferraris, and Francaviglia, and Reina (1981)[17]]. In fact, in spite of the widely known and systematic use of tensorial methods in mathematical physics, restricting ones attention to tensors may often turn out to be misleading.

Motivated by physical applications we have reconsidered the mathematical foundations of the theory of geometric objects, providing for them a new direct approach, which adapts the nice construction proposed by Haantjes and Laman (1953a,b) to the more flexible language of differential geometry of fiber bundles. Our approach is less general than that of Salvioli because it refers explicitly to geometric objects having finite rank. However, it has the advantage of being constructive and able to handle in a simple, intrinsic, and detailed way most of the bundles of geometric objects which are relevant to mathematical physics. It, in fact, provides explicit constructions for the "lifting functors" of Salvioli's method and allows much easier calculations.

2. FIBER BUNDLES ON MANIFOLDS

1. Fields of geometric objects naturally arise as sections of suitable bundles. In the following we shall restrict ourselves to the bundles of geometric objects having finite rank, because they have the property of being fiber bundles.

Therefore, let us begin by recalling the concepts of fiber bundle theory we shall need later. We adopt the following definition.

Definition 2.1: Let $\mathbb{M}$, $\mathbb{F}$ be C^∞-manifolds and G a Lie group. A fiber bundle over $\mathbb{M}$ (with structure group G and standard fiber $\mathbb{F}$) is a quintuple $(\mathbb{B}, \mathbb{M}, \pi; \mathrm{G}, \mathbb{F})$, where π: $\mathbb{B}\to\mathbb{M}$ is a surjective map from a differentiable manifold $\mathbb{B}$ onto $\mathbb{M}$, if the following conditions are satisfied. (i) G acts effectively and differentiably on $\mathbb{F}$; (ii) there exist an open covering $\{U_a\}$ of $\mathbb{M}$ and homeomorphisms (called local tri-

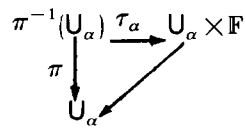
[a)]Work sponsored by C. N. R.–G. N. F. M.

 0022-2488/83/010120-05$02.50

vializations)

$$\tau_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{F}$$

such that the diagram



is commutative; (iii) there exist maps $m_{\alpha\beta}: U_{\alpha\beta} = U_\alpha \cap U_\beta \to G$ (called transition functions) such that

$$\tau_\alpha \cdot \tau_\beta^{-1}(p, f) = (p, m_{\alpha\beta}(p) f) \quad \forall p \in U_{\alpha\beta},\ f \in \mathbb{F}.$$

Remark 2.2: The transition functions above satisfy the compatibility relations $m_{\beta\alpha}(p) = [m_{\alpha\beta}(p)]^{-1}$ and $m_{\alpha\beta}(p) m_{\beta\gamma}(p) m_{\gamma\alpha}(p) = \mathbf{1} \in G$, $p \in U_\alpha \cap U_\beta \cap U_\gamma$. Therefore, they form a 1-cocycle with values in the sheaf of germs of local differentiable functions from $\mathbb{M}$ to G [for more details see Hirzebruck (1978)[18]].

Remark 2.3: Note that given a covering $\{U_\alpha\}$ of $\mathbb{M}$ and a set of G-valued transition functions $m_{\alpha\beta}$ satisfying the properties of Remark 2.2 one can construct a fiber bundle $\mathbb{B}$ over $\mathbb{M}$ with structure group G and standard fiber $\mathbb{F}$. We first form the disjoint union $\tilde{\mathbb{B}}$ of all the sets $U_\alpha \times \mathbb{F}$. The bundle $\mathbb{B}$ is then obtained from $\tilde{\mathbb{B}}$ by identifying the points $(p, f) \in U_\alpha \times \mathbb{F}$ and $(p, m_{\alpha\beta}(p) f) \in U_\beta \times \mathbb{F}$ for any α, β and $p \in U_{\alpha\beta}$. One can show that such a reconstruction does not depend on the choice of the covering $\{U_\alpha\}$.

2. As examples of the preceding construction we may quote the following.

Example 2.4: A Lie group acts naturally on itself on the left (or on the right). Therefore, one can construct fiber bundles having the structure group G itself as standard fiber. These are called principal G–bundles and will be denoted by $(\mathbb{P}, \mathbb{M}, \pi; G)$. Principal G bundles may be characterized as follows. A quadruplet $(\mathbb{P}, \mathbb{M}, \pi; G)$ is a principal G bundle if and only if the following prescriptions are satisfied: (i) G is a Lie group, $\mathbb{P}$ and $\mathbb{M}$ are C^∞ manifolds, and $\pi: \mathbb{P} \to \mathbb{M}$ is a surjective map of maximal rank; (ii) there exists a right (or left) action $R: \mathbb{P} \times G \to \mathbb{P}$ of G on $\mathbb{P}$ which is free [i.e., if $p \in \mathbb{P}$, $g \in G$, and $R(p, g) = p$ then g is the identity of G], differentiable, and such that $\mathbb{M} = \mathbb{P}/G$ [i.e., $\forall p \in \mathbb{P}$, $g \in G$, $\pi[R(p, g)] = \pi(p)$].

Example 2.5: Let $(\mathbb{P}, \mathbb{M}, \pi; G)$ be a principal G-bundle, $\mathbb{F}$ be a manifold, and $\rho: G \to \mathscr{D}(\mathbb{F})$ be a representation of G into the group $\mathscr{D}(\mathbb{F})$ of diffeomorphisms of $\mathbb{F}$. According to Remark 2.3 one can construct a fiber bundle $(\mathbb{B}, \mathbb{M}, \pi'; \rho(G), \mathbb{F})$ by using the composition of ρ with the transition functions of $\mathbb{P}$. An alternative well-known procedure consists in taking the quotient of the manifold $\mathbb{P} \times \mathbb{F}$ with respect to the equivalence relation defined by the following group action, $\hat{\rho}: \mathbb{P} \times \mathbb{F} \times G \to \mathbb{P} \times \mathbb{F}$, induced by

$$\hat{\rho}: (p, f, g) \mapsto (pg, \rho(g)^{-1} f). \tag{1}$$

To this bundle we shall give the name of bundle of objects of type ρ associated with $\mathbb{P}$.

Example 2.6: In particular, whenever G admits a representation $\lambda: G \to GL(V)$ in the linear group of some vector space V, by the same procedure we can construct bundles having V as standard fiber and λ (G) as structure group. These are called vector bundles over $\mathbb{M}$ (associated with $\mathbb{P}$).

Example 2.7: Whenever G admits a representation $\alpha: G \to IGL(A)$ in the affine group of some affine space A we obtain affine bundles (associated with $\mathbb{P}$), having A as standard fiber. Note that any vector bundle can be improperly considered as an affine bundle by identifying GL (V) with its isomorphic copy contained in IGL(V), where V is considered as an affine space. This procedure can be inverted, in the sense that given an affine bundle $(\mathbb{B}, \mathbb{M}, \pi; G, A)$ we may define an associated vector bundle $\mathbb{B}'$ having as fiber the vector space V underlying the affine fibers A of $\mathbb{B}$.

3. BUNDLES OF GEOMETRIC OBJECTS

1. Among the fiber bundles over $\mathbb{M}$ with given fiber $\mathbb{F}$ and structure group G, we shall describe here an important subclass, whose transition functions $m_{\alpha\beta}(p)$ are constructed starting only from the differentiable structure of $\mathbb{M}$. This is the original viewpoint of Haantjes and Laman, which will here be briefly recalled and set up in slightly different language, in order to prepare us for the alternative description which will be given later.

Let $\{(U_\alpha, \varphi_\alpha)\}$ be an atlas of $\mathbb{M}$. For any pair of overlapping charts $((U_\alpha, \varphi_\alpha), (U_\beta, \varphi_\beta))$, there exists a (local) C^∞ diffeomorphism

$$\Phi_{\alpha\beta} = \varphi_\alpha \cdot \varphi_\beta^{-1}: \varphi_\beta(U_{\alpha\beta}) \to \varphi_\alpha(U_{\alpha\beta}) \tag{2}$$

between open subsets of $\mathbb{R}^n$. The local diffeomorphisms $\Phi_{\alpha\beta}$ are usually called coordinate transformations. Our next task will then be to construct transition functions out of these local diffeomorphisms of $\mathbb{R}^n$.

First of all, we note that for any point $p \in U_{\alpha\beta}$ and for any $\Phi_{\alpha\beta}$ one can construct a local diffeomorphism $\tilde{\Phi}_{\alpha\beta}(p)$ of $\mathbb{R}^n$ into itself such that $\tilde{\Phi}_{\alpha\beta}(p)(0) = 0$, by defining

$$\tilde{\Phi}_{\alpha\beta}(p): x \mapsto \Phi_{\alpha\beta}[x + \varphi_\beta(p)] - \varphi_\alpha(p) \tag{3}$$

for any $x \in \mathbb{R}^n$ such that $x + \varphi_\beta(p) \in \varphi_\beta(U_{\alpha\beta})$. The local diffeomorphisms $\tilde{\Phi}_{\alpha\beta}(p)$ satisfy the following conditions: $[\tilde{\Phi}_{\alpha\beta}(p)]^{-1} = \tilde{\Phi}_{\beta\alpha}(p)$ and $\tilde{\Phi}_{\alpha\beta}(p) \cdot \tilde{\Phi}_{\beta\sigma}(p) \cdot \tilde{\Phi}_{\sigma\alpha}(p) = \mathrm{id}_{\mathbb{M}}$ for any $p \in U_\alpha \cap U_\beta \cap U_\sigma$.

2. Now let $\mathscr{P}_0(\mathbb{R}^n)$ be the pseudogroup of all local diffeomorphisms Ψ of $\mathbb{R}^n$ into itself such that $\Psi(0) = 0$. We remind the reader that $[D\Psi(0)]^{-1}$ exists, where the linear map $D\Psi(0): \mathbb{R}^n \to \mathbb{R}^n$ denotes the derivative of Ψ at 0. For any $\Psi \in \mathscr{P}_0(\mathbb{R}^n)$ we define $t^k(\Psi)$ to be the Taylor expansion of Ψ at 0 up to and including the order $k \geqslant 0$. Two local diffeomorphisms $\Psi, \Psi' \in \mathscr{P}_0(\mathbb{R}^n)$ are said to agree to the order k (at 0) if $t^k(\Psi) = t^k(\Psi')$. This is obviously an equivalence relation. The equivalence class $j^k(\Psi)$ may be represented as follows:

$$j^k(\Psi) = (0, D\Psi(0), D^2\Psi(0), \ldots D^k\Psi(0)),$$

where the symmetric r-linear operators $D^r\Psi(0): (\mathbb{R}^n)^r \to \mathbb{R}^n$ denote the rth order derivatives of Ψ at 0.

Let $G^k(n; \mathbb{R}) = \{j^k(\Psi) \mid \Psi \in \mathscr{P}_0(\mathbb{R}^n)\}$ denote the quotient set of $\mathscr{P}_0(\mathbb{R}^n)$ under the above equivalence relation. It is easy to show that $G^k(n; \mathbb{R})$ is a (real) Lie group with respect to the natural composition law:

$$j^k(\Psi) \cdot j^k(\Psi') = j^k(\Psi \cdot \Psi')$$

In particular, when $k = 1$ we recover the general linear group $GL(n;\mathbb{R})$.

3. Since all the local diffeomorphisms $\tilde{\Phi}_{\alpha\beta}(p)$ defined in Sec. 2.1 belong to $\mathscr{P}_0(\mathbb{R}^n)$ we may consider their k th order jets $j^k(\tilde{\Phi}_{\alpha\beta}(p)) \in G^k(n;\mathbb{R})$. Then, for any $U_{\alpha\beta}$ and any integer $k \geqslant 0$ we may define functions $\Phi^k_{\alpha\beta}: U_{\alpha\beta} \to G^k(n;\mathbb{R})$ as follows:

$$\Phi^k_{\alpha\beta}: p \mapsto j^k[\tilde{\Phi}_{\alpha\beta}(p)]. \tag{4}$$

One may easily check that the functions $\Phi^k_{\alpha\beta}$ defined in this way are C^∞ functions from $U_{\alpha\beta}$ into $G^k(n;\mathbb{R})$. Moreover, they satisfy the following conditions:

(i) $[\Phi^k{}_{\alpha\beta}(p)]^{-1} = \Phi^k{}_{\beta\alpha}(p)$

(ii) $\Phi^k{}_{\alpha\beta}(p)\cdot\Phi^k{}_{\beta\sigma}(p)\cdot\Phi^k{}_{\sigma\alpha}(p) = \mathbb{1}$,

where $\mathbb{1} = j^k(\mathrm{id}_n)$ denotes the identity of the group $G^k(n;\mathbb{R})$.

Remark 3.1: By this last result we see that the functions $\Phi^k_{\alpha\beta}$ may be considered as transition functions for a fiber bundle having structure group $G^k(n;\mathbb{R})$, because they form a 1-cocycle with values in the group $G^k(n;\mathbb{R})$.

4. Now let $\mathbb{F}$ be a manifold on which a Lie group G acts effectively and differentiably and ρ: $G^k(n;\mathbb{R}) \to \mathsf{G}$ be a group homomorphism onto G. Then we have maps

$$m_{\alpha\beta}: U_{\alpha\beta} \xrightarrow{\Phi^k_{\alpha\beta}} G^k(n;\mathbb{R}) \xrightarrow{\rho} \mathsf{G}$$

given by $m_{\alpha\beta}(p) = \rho(\Phi^k_{\alpha\beta}(p))$, which "lift" the differentiable structure of $\mathbb{M}$ into G-valued transition functions. From these data we can construct a fiber bundle $\mathbb{B}$ over $\mathbb{M}$ with standard fiber $\mathbb{F}$ and structure group G, which will be denoted by $(\mathbb{B}, \mathbb{M}, \pi; \mathbb{F}, \mathsf{G}, \rho)$. We then give the following definition:

Definition 3.2: $(\mathbb{B}, \mathbb{M}, \pi; \mathbb{F}, \mathsf{G}, \rho)$, where $\rho: G^k(n;\mathbb{R}) \to \mathsf{G}$, is called a bundle of geometric objects of type ρ of finite rank $(\leqslant k)$.

The bundles of geometric objects defined in this way fit into the scheme of Salvioli. It can be proved, in fact, that they satisfy all the properties listed in Salvioli (Ref. 9, p. 259).

We remark that the definitions given by Salvioli extend to also cover geometric objects of infinite rank. The direct approach we outlined above may also be extended to this case by relying on suitable Fréchet manifolds, i.e., by allowing the use of infinite jets of mappings.

5. We remark that, in differential geometry and in its recent applications to physics, a central role is played by principal fiber bundles and that, moreover, all fiber bundles can be considered as associated with some suitable principal fiber bundle.

Our next task is then to show that the construction we outlined above is, in fact, in agreement with this spirit, in the sense that all the bundles of geometric objects covered by Definition 3.2 are associated with certain principal bundles. This will provide us an alternative approach to the class of bundles considered, which, as we shall see later, is more manageable for applications.

Let us then proceed as follows. Given a C^∞-manifold $\mathbb{M}$ and an integer k $(1 \leqslant k < \infty)$, for any C^∞-function $h \in C^\infty(\mathbb{R}^n, \mathbb{M})$, the k th order jet $j^k(h)$ of h at $0 \in \mathbb{R}^n$ is naturally defined by reverting to any local parametrization of $\mathbb{M}$. We denote by $\mathbb{L}^k(\mathbb{M})$ the set of all the jets $j^k(h)$ such that h^{-1} exists. Let us now consider the quadruple $(\mathbb{L}^k(\mathbb{M}), \mathbb{M}, \pi^k; G^k(n,\mathbb{R}))$, where π^k: $\mathbb{L}^k(\mathbb{M}) \to \mathbb{M}$ is the canonical projection defined by $\pi^k[j^k(h)] = h(0)$. From the construction above, we see that there exists a canonical right-action of $G^k(n;\mathbb{R})$ on $\mathbb{L}^k(\mathbb{M})$, which is given by

$$(j^k(h), j^k(\Psi)) \mapsto j^k(h\cdot\Psi).$$

It is easy to check that this defines a principal $G^k(n;\mathbb{R})$-bundle over $\mathbb{M}$ [see Example 2.5].

Now let $\rho: G^k(n;\mathbb{R}) \to \mathsf{G}$ be a group homomorphism and $\mathbb{F}$ a manifold on which the Lie group G acts effectively and differentiably. We see immediately that the bundles of geometric objects of Definition 2.3 are the bundles of type ρ associated with $\mathbb{L}^k(\mathbb{M})$ in the sense of Example 2.6. Our claim is thence proved.

The principal bundles $\mathbb{L}^k(\mathbb{M})$ will be called bundles of k th order frames on $\mathbb{M}$. This terminology is motivated by the fact that $\mathbb{L}^1(\mathbb{M})$ is isomorphic with the bundle of linear frames of $\mathbb{M}$.

4. EXAMPLES

1. According to our previous remarks, all the bundles of geometric objects of types ρ are associated with some of the principal bundles $\mathbb{L}^k(\mathbb{M})$, which therefore are, in a sense, the prototype of such bundles.

Note that $\mathbb{L}^k(\mathbb{M})$ is associated with $\mathbb{L}^{k'}(\mathbb{M})$ whenever $k' \geqslant k$, thanks to the existence of a canonical epimorphism from $G^{k'}(n;\mathbb{R})$ onto $G^k(n;\mathbb{R})$. As a consequence, if a bundle $\mathbb{B}$ of geometric objects of type ρ is associated with $\mathbb{L}^k(\mathbb{M})$ it is also associated with all principal bundles $\mathbb{L}^{k'}(\mathbb{M})$ with $k' \geqslant k$. The smallest integer k such that $\mathbb{B}$ is associated with $\mathbb{L}^k(\mathbb{M})$ is called the rank of $\mathbb{B}$.

Example 4.1: All the bundles of tensors over $\mathbb{M}$ may be obtained as vector bundles associated with the bundle of geometric objects $\mathbb{L}^1(\mathbb{M})$, by means of suitable linear representations of $G^1(n;\mathbb{R})$.

For example, the tangent bundle $\mathbb{TM}$ is obtained from the canonical isomorphism $i: G^1(n;\mathbb{R}) \to GL(n;\mathbb{R})$ while the contangent bundle $\mathbb{T}^*\mathbb{M}$ is obtained from the inverse transpose isomorphism $i^*: G^1(n;\mathbb{R}) \to GL(n;\mathbb{R})$ defined by

$$i^*: j^1(\Psi) \mapsto {}^t[i(j^1(\Psi^{-1}))]. \tag{5}$$

The tensor bundles $\mathbb{T}^p_q(\mathbb{M})$ are then obtained by tensorizing the above constructions; analogously for the bundle $\Lambda^p(\mathbb{M})$ of differential p-forms.

Example 4.2: Let det: $GL(n;\mathbb{R}) \to \mathbb{R}^*$ be the determinant homomorphism. We denote by Δ the composition $\Delta = (\det)\cdot i: G^1(n;\mathbb{R}) \to \mathbb{R}^*$. From the linear representation Δ we can construct a line bundle $\det(\mathbb{M})$, called the determinant bundle of $\mathbb{M}$, whose sections are the fields of n-vectors on $\mathbb{M}$. Analogously, we can construct the dual bundle $\det^*(\mathbb{M})$ by using the linear representation $\Delta^* = (\det)\cdot i^*$. Its sections are the fields of n-covectors on $\mathbb{M}$ and, therefore, there is a natural isomorphism between the bundles $\det^*(\mathbb{M})$ and $\Lambda^n(\mathbb{M})$.

Example 4.3: Let $U(1)$ be the unitary group. By relying on $\det(\mathbb{M})$ one can construct a principal $U(1)$-bundle of geometric objects $U(\mathbb{M})$. This can be done by considering the epimorphism $\alpha: G^1(n;\mathbb{R}) \to U(1)$ defined by

$$\alpha: j^1(\Psi) \mapsto \exp[i \ln|\Delta(j^1(\Psi))|], \tag{6}$$

or shortly $\alpha = \exp(i \ln|\Delta|)$. The conjugate bundle $\mathsf{U}^*(\mathbb{M})$ is obtained in a completely analogous way, by relying instead on the epimorphism $\alpha^* = \exp(-i \ln|\Delta|)$. To the bundle $\mathsf{U}(\mathbb{M})$, which enters some recent unified theory of gravitation and electromagnetism [Ferraris and Kijowski (1981)[19]], we shall give the name of unitary bundle of $\mathbb{M}$.

Example 4.4: Let us now consider the bundle $\mathbb{L}^2(\mathbb{M})$. We define a natural left action of $G^2(n;\mathbb{R})$ on the vector space $\mathbb{T}^1_2(\mathbb{R}^n) = \mathbb{R}^n \otimes (\mathbb{R}^n)^* \otimes (\mathbb{R}^n)^*$ by the following explicit relation:

$$(\mathscr{g}^i_j, \mathscr{g}^i_{jk})(\Gamma^a_{bc}) = (\mathscr{g}^i_a \Gamma^a_{bc} \overline{\mathscr{g}}^b_j \overline{\mathscr{g}}^c_k + \mathscr{g}^i_a \overline{\mathscr{g}}^a_{jk}), \tag{7}$$

where $(\mathscr{g}^i_j, \mathscr{g}^i_{jk})$ and Γ^a_{bc} are canonical coordinates in $G^2(n;\mathbb{R})$ and $\mathbb{T}^1_2(\mathbb{R}^n)$, respectively, and $(\overline{\mathscr{g}}^i_j, \overline{\mathscr{g}}^i_{jk})$, denotes the inverse of $(\mathscr{g}^i_j, \mathscr{g}^i_{jk})$. The fiber bundle $\mathsf{C}(\mathbb{M})$ associated with $\mathbb{L}^2(\mathbb{M})$ via the affine representation (7) above is an affine bundle of geometric objects, whose sections are easily recognized to be the linear connections over $\mathbb{M}$. For this reason the bundle $\mathsf{C}(\mathbb{M})$ will be called the connection bundle of $\mathbb{M}$. It is easy to check that the vector bundle canonically associated with $\mathsf{C}(\mathbb{M})$ is the tensor bundle $\mathbb{T}^1_2(\mathbb{M})$.

Example 4.5: We can now define a further bundle by "taking the trace" of $\mathsf{C}(\mathbb{M})$, namely by considering the following left action of $G^2(n;\mathbb{R})$ on $(\mathbb{R}^n)^*$:

$$(\mathscr{g}^i_j, \mathscr{g}^i_{jk})(A_a) = (A_a \overline{\mathscr{g}}^i_j + \mathscr{g}^k_a \overline{\mathscr{g}}^a_{jk}), \tag{8}$$

where A_a are coordinates in $(\mathbb{R})^*$. The mapping (8) is obtained by taking a suitable trace in (7). It is easily seen that (8) is truly an action of $G^2(n;\mathbb{R}) \times (\mathbb{R}^n)^*$ into $(\mathbb{R}^n)^*$ and that it defines an affine bundle $\mathsf{D}^*(\mathbb{M})$ over $\mathbb{M}$, which will be called the dilatation bundle of $\mathbb{M}$. This terminology is suggested by the fact that the sections of $\mathsf{D}^*(\mathbb{M})$ are linear connections on the vector bundle $\det^*(\mathbb{M}) \simeq \Lambda^n(\mathbb{M})$, whose structure group is the group of dilatations in $\mathbb{R}^n$. We can easily realize that the vector bundle associated with $\mathsf{D}^*(\mathbb{M})$ is the cotangent bundle $\mathsf{T}^*\mathbb{M}$. There exists, of course, a dual construction, which gives a bundle $\mathsf{D}(\mathbb{M})$ whose sections are connections on $\det(\mathbb{M})$.

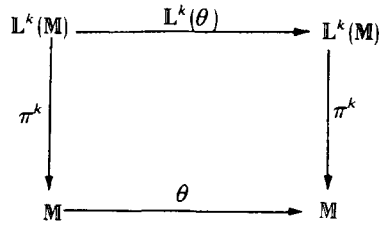
2. Other constructions involving the bundles of affine frames, projective frames, and spinor frames are currently under investigation and they will be the subject of further publication.

5. LIFT OF DIFFEOMORPHISMS AND LIE DERIVATIVES

In this last section we shall prove our main concern, i.e., we shall show that the construction presented above enables one to define in an intrinsic and canonical way the functorial lift of (local) diffeomorphisms of $\mathbb{M}$ to any bundle of geometric objects of type ρ and finite rank. This canonical lifting will provide more explicit formulas for the Lie derivative of a field of geometric objects.

Note added in proof: A more extended version, containing a detailed discussion of $\mathsf{U}(\mathbb{M})$ bundles and their role in providing a possible characterization of the electric charge, will appear in J. Math. Pures Appl. Phys.

1. Let $k \geqslant 1$ be an integer. Let θ: $\mathbb{M} \to \mathbb{M}$ be a local diffeomorphism of $\mathbb{M}$. There exists a canonical lift $\mathbb{L}^k(\theta): \mathbb{L}^k(\mathbb{M}) \to \mathbb{L}^k(\mathbb{M})$ such that the following diagram is commutative:



and $\mathbb{L}^k(\theta)$ is a local diffeomorphism which commutes with the natural right action of $G^k(n;\mathbb{R})$ on $\mathbb{L}^k(\mathbb{M})$. In fact, the local diffeomorphism $\mathbb{L}^k(\theta)$ is defined by the following relation:

$$\mathbb{L}^k(\theta): j^k(h) \to j^k(\theta \cdot h), \tag{9}$$

where $h: U(0) \subset \mathbb{R}^n \to \mathbb{M}$ is a local diffeomorphism.

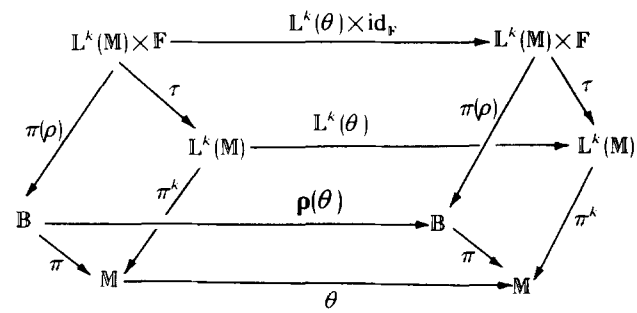
It is easy to prove that the lifting $\mathbb{L}^k: \theta \to \mathbb{L}^k(\theta)$ so defined satisfies the following properties:

$$\mathbb{L}^k(id_{\mathbf{M}}) = id_{\mathbb{L}^k(\mathbf{M})}, \tag{10}$$

$$\mathbb{L}^k(\theta_1 \cdot \theta_2) = \mathbb{L}^k(\theta_1) \cdot \mathbb{L}^k(\theta_2). \tag{11}$$

Therefore, $\mathbb{L}^k$ defines a (covariant) functor from the category of manifolds with local diffeomorphisms to the category of principle fiber bundles with principal fiber bundle morphisms.

2. The functorial construction above can be extended to any bundle of geometric objects of type ρ and finite rank k $(\mathbf{B}, \mathbb{M}, \pi; \mathbf{F}, \mathbf{G}, \rho)$ by the following procedure. First we remind the reader that, according to Sec. 3.5, the bundle $\mathbf{B}$ is associated with the principal bundle $\mathbb{L}^k(\mathbb{M})$ via the canonical projection $\pi(\rho)$: $\mathbb{L}^k(\mathbb{M}) \times \mathbf{F} \to \mathbf{B}$ defined by the group action ρ [in the sense of Example 2.5]. Let us denote by τ the projection of $\mathbb{L}^k(\mathbb{M}) \times \mathbf{F}$ onto the first factor $\mathbb{L}^k(\mathbb{M})$. Then there exists a local diffeomorphism $\boldsymbol{\rho}(\theta): \mathbf{B} \to \mathbf{B}$ such that the following (three-dimensional) diagram is commutative:



In fact, $\boldsymbol{\rho}(\theta)$ is defined by the following prescription:

$$\boldsymbol{\rho}(\theta): \pi(\rho)[j^k h, f] \to \pi(\rho)[j^k(\theta \cdot h), f], \tag{12}$$

for any $(j^k h, f) \in \mathbb{L}^k(\mathbb{M}) \times \mathbf{F}$. The relation (12) is well defined, because $\mathbb{L}^k$ commutes with the group action of $G^k(n;\mathbb{R})$.

It is easy to show that (12) defines a local isomorphism of bundles $\boldsymbol{\rho}(\theta): \mathbf{B} \to \mathbf{B}$ which, moreover, satisfies the required functorial properties:

$$\boldsymbol{\rho}(id_{\mathbf{M}}) = id_{\mathbf{B}}, \tag{13}$$

$$\boldsymbol{\rho}(\theta_1 \cdot \theta_2) = \boldsymbol{\rho}(\theta_1) \cdot \boldsymbol{\rho}(\theta_2). \tag{14}$$

Therefore, setting $\mathbf{B} = \boldsymbol{\rho}(\mathbb{M})$ we have a covariant functor $\boldsymbol{\rho}$ from the category of manifolds with local diffeomorphisms to the category of bundles of geometric objects of finite rank with local bundle-isomorphisms. It is obvious that in the particular case $\mathbf{F} = G^k(n;\mathbb{R})$ and $\boldsymbol{\rho} = id_{G^k(n;\mathbf{R})}$ the functor $\boldsymbol{\rho}$

reduces to the functor $\mathbb{L}^k$.

It is straightforward to prove that the covariant functor ρ defined above satisfies all the required properties in order to make $(\mathbb{B}, \mathbb{M}, \rho)$ a bundle of geometric objects in the sense of Salvioli.

3. In order to define the Lie derivative of a field of geometric objects along a vector field X on $\mathbb{M}$, we may now apply the standard procedure described in Salvioli (1972), making explicit the functor ρ.

Then let θ_t be the local 1-parameter group of diffeomorphisms generated by a vector field X on $\mathbb{M}$ and let $\beta:\mathbb{M}\to\mathbb{B}$ be a (local) section of a bundle of geometric objects $(\mathbb{B},\mathbb{M},\pi;\mathbb{F},\mathbf{G},\rho)$ of finite rank $k\geqslant 1$. The following relation,

$$\beta_t: x\in\mathbb{M}\to\rho(\theta_t)^{-1}[\beta\cdot\theta_t(x)]\in\pi^{-1}(x), \tag{15}$$

defines a one-parameter family of local sections of $\mathbb{B}$. Accordingly, we may define the Lie derivative of the (local) field of geometric objects β as follows:

$$L_X\beta:x\in\mathbb{M}\to\frac{d}{dt}[\beta_t(x)]\Big|_{t=0}. \tag{16}$$

It is easy to check that $L_X\beta$ defines a (local) field of vertical vectors over β, i.e., the following conditions hold;

(i) $\pi_{\mathbf{B}}\cdot(L_X\beta)=\beta$,

where $\pi_{\mathbf{B}}:\mathbb{TB}\to\mathbb{B}$ is the canonical projection;

(ii) $[\mathbb{T}\pi\cdot(L_X\beta)](x)=x, \quad \forall x\in\mathbb{M}$,

where $\mathbb{T}\pi:\mathbb{TB}\to\mathbb{TM}$ is the tangent map of the bundle projection π.

For further properties of Lie derivatives of geometric objects we refer the reader to Salvioli (1972) or Yano (1955).[20]

Note added in proof. A more extended version, containing a detailed discussion of $\mathrm{U}(\mathbb{M})$ bundles and their role in providing a possible characterization of the electric charge, will appear in Annales Inst. H. Poincaré.

[1]A. Einstein, "Die Grundlage der allgemeinen Relativitätstheorie," Ann. Phys. **49**, 769–822 (1916).

[1a]T. Levi-Civita, "Nozione di parallelismo in una varietá qualunque e conseguente specificazione geometrica della curvatura riemanniana," Rend. Circ. Mat. Palermo **42**, 73–205 (1917).

[2]H. Weyl, *Raum, Zeit, Materie* (Springer, Berlin, 1918).

[3]E. Cartan, "Sur les variétées a connexion affine et la théorie de la relativité géneralisée," Ann. Ec. Norm. Sup **40**, 325–412 (1923).

[4]J. A. Schouten and J. Haantjes, "On the theory of Geometric Objects," Proc. London Math. Soc. **42**, 356–376 (1936).

[5]A. Nijenhuis, "Theory of Geometric Objects," Doctoral Thesis, University of Amsterdam (1952) (unpublished).

[6]A. Nijenhuis, "Geometric Aspects of Formal Differential Operations on Tensorfields," in *Proceedings of the International Congress of Mathematics*, 1958 (Cambridge U. P., Cambridge, 1960).

[7]J. Haantjes and G. Laman (1953a,b), "On the Definition of Geometric Objects I, II", Indag. Math. **15**, 208–222 (1953).

[8]N. H. Kuiper and K. Yano, "On Geometric Objecs and Lie Groups of Transformations," Indag. Math. **17**, 411–420 (1955).

[9]S. Salvioli, "On the theory of Geometric Objects," J. Diff. Geom. **7**, 257–278 (1972).

[10]J. L. Anderson, *Principles of Relativity Physics* (Academic, New York, 1967).

[11]D. Krupka, (a), "Reducibility Theorems for Differentiable Liftings in Fiber Bundles," Arch. Math. **2**, Scripta Fac. Sci. Nat. UJEP Brunensis, XV, 93–106 (1979); (b), "Differential Invariants," Lecture Notes, University of Brno, August, 1979.

[12]J. Kijowski and W. M. Tulczyjew, *A Symplectic Framework for Field Theories*, Lect. Notes in Phys. **107** (Springer, Berlin, 1978).

[13]A. Prastaro, "On the General Structure of Continuum Physics: I. Derivative Spaces," Boll. U. M. I. **17** B (5), 704–726 (1980).

[14]A. Prastaro, "On the General Structure of Continuum Physics: II. Differential Operators," Boll. U.M.I. (to appear).

[15]M. Modugno, R. Ragionieri, and G. Stefani, Ann. Inst. Henri Poincarè **34**, 465–496 (1981).

[16]J. F. Pommaret, *Systems of Partial Differential Equations and Lie Pseudogroups* (Gordon and Breach, New York, 1978).

[17]M. Ferraris, M. Francaviglia, and C. Reina, "Variational Formulations of Geometric Theories of Gravitation," preprint (1981) (unpublished).

[18]F. Hirzebruch, *Topological Methods in Algebraic Geometry* (Springer, Berlin, 1978).

[19]M. Ferraris and J. Kijowski, "Unified geometric theory of electromagnetic and gravitational interactions," Gen. Relativ. Gravit. **14**, 37–47 (1982).

[20]K. Yano, *The Theory of Lie Derivatives and its Applications* (North-Holland, Amsterdam, 1955).

Exact solutions of strong gravity in generalized metrics

R. Hojman
Departamento de Física, Facultad de Ciencia, Universidad de Santiago de Chile, Santiago, Chile

A. Smailagic[a)]
International Centre for Theoretical Physics, Trieste, Italy

(Received 30 July 1981; accepted for publication 11 September 1981)

We consider classical solutions for the strong gravity theory of Salam and Strathdee in a class of metrics with positive, zero, and negative curvature. It turns out that such solutions exist and their relevance for quark confinement is explored. Only metrics with positive curvature (spherical symmetry) give a confining potential in a simple picture of the scalar hadron. This supports the idea of describing the hadron as a closed microuniverse of the strong metric.

PACS numbers: 04.20.Jb

INTRODUCTION

We shall discuss strong gravity theory[1] when both (g and f) metrics admit the same three-parameter continuous group of motion described by infinitesimal generators ξ^{σ}_{a} ($\sigma = 0,1,2,3,$; $a = 1,2,3,$) and the matrix $M = \|\xi^{\sigma}_{a}\|$ is of rank two so that the minimum invariant varieties are two-dimensional surfaces of constant curvature. The three cases corresponding to positive, zero, and negative curvature will be considered. The first possibility corresponds to a spherically symmetric metric and its general solution (for both f and g metrics) has been found.[2] The possibility of having quark confinement in the background f metric was analyzed[3] for a particular case (taking $g_{\mu\nu} = \delta_{\mu\nu}$). The analysis is applied to our case in Sec. III. Exact solutions for the three cases are found in Sec. II. In Sec. I the mentioned symmetries are briefly described and their associated line elements in their most general form are deduced from the Killing equations.

I. THE SYMMETRIES AND ASSOCIATED METRICS

It is well known[4] that the most general form of the spherically symmetric line element is given by

$$ds^2 = C(r,t)dt^2 - 2D(r,t)dt\,dr - A(r,t)dr^2 - B(r,t)(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{I.1}$$

with usual interpretation of t, r, θ, and ϕ as radial coordinates.

Consider the case of the two-dimensional minimum invariant varieties of zero curvature. A space-time is said to be plane symmetric if it admits the three-parameter group generated by the transformations

$$\begin{aligned} &\bar{y} = y + c\,, \\ &\bar{z} = z + b\,, \\ &\bar{y} = y\cos\theta - z\sin\theta\,, \\ &\bar{z} = y\sin\theta + z\cos\theta\,. \end{aligned} \tag{I.2}$$

The infinitesimal generators of the group are

$$\|\xi^{\sigma}_{a}\| = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -z \\ 0 & 1 & y \end{vmatrix}. \tag{I.3}$$

It follows directly from the Killing equation

$$\xi^{\mu}_{a} g_{\sigma\tau,u} + g_{\mu\sigma}\xi^{\mu}_{a,\tau} + g_{\mu\tau}\xi^{\mu}_{a,\sigma} = 0, \tag{I.4}$$

that the most general line element admitting this group is given by

$$ds^2 = C(w,x)dw^2 - 2D(w,x)dw\,dx - A(w,x)dx^2 - B(w,x)(dy^2 + dz^2)\,. \tag{I.5}$$

In the same way, if the infinitesimal generators of the group are given by

$$\|\xi^{\sigma}_{a}\| = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & z \\ 1 & -z & \frac{1}{2}(e^{-2y} - z^2) \end{bmatrix}, \tag{I.6}$$

that is to say, when the minimum invariant varieties of the three-parameter group are two-dimensional surfaces of negative curvature, the line element is

$$ds^2 = C(w,x)dw^2 - 2D(w,x)dwdx - A(w,x)dx^2 - B(w,x)(dy^2 + e^{2y}dz^2)\,. \tag{I.7}$$

The expressions (1.1), (1.5), and (1.7) can be summarized in the line element

$$ds^2 = C(w,x)dw^2 - 2D(w,x)dwdx - A(w,x)dx^2 - B(w,x)\{dy^2 + F(y)dx^2\}\,, \tag{I.8}$$

with $F(y) = \sin^2 y$, 1, and e^{2y}, respectively.

II. FIELD EQUATIONS

The field equations for the f and g metrics are[2]

$$\begin{aligned} &R^{g}_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R^{g} = k_g T^{g}_{\mu\nu}\,, \\ &R^{f}_{\mu\nu} - \tfrac{1}{2}f_{\mu\nu}R^{f} = k_f T^{f}_{\mu\nu}\,, \end{aligned}$$

with

$$\begin{aligned} &R^{g}_{\mu\nu} = g^{\alpha\beta}R^{g}_{\alpha\mu\beta\nu}\,, \quad R^{g} = g^{\mu\nu}R^{g}_{\mu\nu}\,, \\ &R^{f}_{\mu\nu} = f^{\alpha\beta}R^{f}_{\alpha\mu\beta\nu}\,, \quad R^{f} = f^{\mu\nu}R^{f}_{\mu\nu}\,, \end{aligned} \tag{II.1}$$

[a)] Partly supported by the National Science Foundation (United States of America) and Samoupravna Interesna Zajednica Za Nauku (Yugoslavia).

 0022-2488/83/010125-04$02.50

and the tensor $T_{\mu\nu}$ given by

$$T^g_{\mu\nu} = \frac{M^2}{8\pi k_f}\left(\frac{f}{g}\right)^v (f-g)^{\alpha\beta}\left[(f-g)^{\rho\tau}\{u g_{\mu\nu}(g_{\alpha\rho}g_{\beta\tau}-g_{\alpha\beta}g_{\rho\tau}) + 2g_{\alpha\mu}(g_{\beta\rho}g_{\nu\tau}-g_{\beta\nu}g_{\rho\tau})\} + 2(g_{\mu\alpha}g_{\nu\beta}-g_{\mu\nu}g_{\alpha\beta})\right], \tag{II.2}$$

$$T^f_{\mu\nu} = \frac{M^2}{8\pi k_f}\left(\frac{g}{f}\right)^u (f-g)^{\alpha\beta}\left[(f-g)^{\rho\tau} v f_{\mu\nu}(g_{\alpha\rho}g_{\beta\tau}-g_{\alpha\beta}g_{\rho\tau}) - 2(g_{\mu\alpha}g_{\nu\beta}-g_{\mu\nu}g_{\alpha\beta})\right], \quad u+v=\tfrac{1}{2},$$

and they follow from the f-g Lagrangian

$$\mathcal{L} = -\frac{1}{k_g}(-g)^{1/2}R^g - \frac{1}{k_f}(-f)^{1/2}R^f - \frac{M^2}{4k_f}(-g)^u(-f)^v \times (f-g)^{\alpha\beta}(f-g)^{\sigma\tau}(g_{\alpha\sigma}g_{\beta\tau}-g_{\alpha\beta}g_{\sigma\tau}). \tag{II.3}$$

We shall search for exact solutions of these equations in the case that the metric coefficients depend only on one variable (kind of solitonic solution). Choosing this coordinate to be x, the f and g metrics are given by

$$ds_g^2 = g_{\mu\nu}dx^\mu dx^\nu = \gamma(x)dw^2 - 2\delta(x)dwdx - \alpha(x)dx^2 - \beta(x)\{dy^2 + F(y)dz^2\}, \tag{II.4}$$

$$ds_f^2 = f_{\mu\nu}dx^\mu dx^\nu = C(x)dw^2 - 2D(x)dwdx - A(x)dx^2 - B(x)\{dy^2 + F(y)dz^2\}.$$

We can simplify the above expressions by appropriate coordinate transformations, though we should keep in mind that such transformations must be performed simultaneously on both metrics in order to preserve the invariance of the theory.

If one defines $\tilde{w} = w + \psi$ with $d\psi/dx = -\delta/\gamma$ we can write the f and g metrics as

$$ds_g^2 = \gamma dw^2 - \alpha dx^2 - x^2(dy^2 + Fdz^2),$$
$$ds_f^2 = Cdw^2 - 2Ddwdx - Adx^2 - B(dy^2 + Fdz^2). \tag{II.5}$$

The nonzero components of the curvature are

$$R^f_{00} = \frac{C}{2\Delta}\left\{C'' + \frac{R'C'}{B} - \frac{C'\Delta'}{2\Delta}\right\}; \quad \Delta \equiv AC - D^2,$$
$$R^f_{01} = -\frac{D}{2\Delta}\left\{C'' + \frac{B'C'}{B} - \frac{C'\Delta'}{2\Delta}\right\},$$
$$R^f_{11} = -\frac{B''}{B} + \frac{B'^2}{2B^2} + \frac{B'\Delta'}{2B\Delta} - \frac{A}{2\Delta}\left\{C'' + \frac{B'C'}{B} - \frac{C'\Delta'}{2\Delta}\right\}, \tag{II.6}$$
$$R^f_{22} = R_{33}/F = \epsilon(F) - \frac{C}{2\Delta}\left\{B'' + \frac{B'C'}{C} - \frac{B'\Delta'}{2\Delta}\right\},$$

with

$$\epsilon(F) \equiv \left(\frac{\dot{F}}{2F}\right)^2 - \frac{\ddot{F}}{2F} = \begin{cases} 1 \text{ for } F = \sin^2 y \\ 0 \text{ for } F = 1 \\ -1 \text{ for } F = e^{2y}, \end{cases} \tag{II.6'}$$

prime and dot meaning x and y differentiation, respectively. Also we have

$$R^f = \frac{1}{\Delta}\left(C'' + \frac{B'C'}{B} - \frac{C'\Delta'}{2\Delta}\right) + \frac{C}{\Delta B}\left(B'' - \frac{B'^2}{2B} - \frac{B'\Delta'}{2\Delta}\right) + \frac{2}{B}\left\{\frac{C}{2\Delta}\left(B'' + \frac{B'C'}{C} - \frac{B'\Delta'}{2\Delta}\right) - \epsilon(F)\right\}. \tag{II.7}$$

The components of the g metric are obtained by the simple replacements $A\to\alpha$, $B\to\beta$, $C\to\gamma$, $D\to 0$ in (II.6).

Evaluation of the components of the tensors $T^g_{\mu\nu}$ and $T^f_{\mu\nu}$ gives

$$T^g_{00} = \frac{M^2}{4\pi k_f}\left(\frac{4\Delta}{9\alpha\gamma}\right)^v \gamma\left\{u\theta - \frac{\gamma}{\Delta}\left[\alpha + A\left(\frac{2\beta}{B} - 3\right)\right]\right\},$$
$$T^g_{01} = \frac{M^2}{4\pi k_f}\left(\frac{4\Delta}{9\alpha\gamma}\right)^v \frac{\alpha\gamma D}{\Delta}\left(3 - \frac{2\beta}{B}\right),$$
$$T^g_{11} = -\frac{M^2}{4\pi k_f}\left(\frac{4\Delta}{9\alpha\gamma}\right)^v \alpha\left\{u\theta - \frac{\alpha}{\Delta}\left[\gamma + C\left(\frac{2\beta}{B} - 3\right)\right]\right\},$$
$$T^g_{22} = -\frac{M^2}{4\pi k_f}\left(\frac{4\Delta}{9\alpha\gamma}\right)^v \beta\left\{u\theta - \frac{\beta}{B}\left(\frac{A\gamma + C\alpha}{\Delta} + \frac{\beta}{B} - 3\right)\right\}, \tag{II.8}$$

$$T^f_{00} = \frac{M^2}{4\pi k_f}\left(\frac{9\alpha\gamma}{4\Delta}\right)^u \left\{vC\theta + \gamma\left(\frac{C\alpha}{\Delta} + \frac{2\beta}{B} - 3\right)\right\},$$
$$T^f_{01} = -\frac{M^2}{4\pi k_f}\left(\frac{9\alpha\gamma}{4\Delta}\right)^u \left\{vD\theta + \frac{\alpha\gamma D}{\Delta}\right\},$$
$$T^f_{11} = -\frac{M^2}{4\pi k_f}\left(\frac{9\alpha\gamma}{4\Delta}\right)^u \left\{vA\theta + \alpha\left(\frac{A\gamma}{\Delta} + \frac{2\beta}{\Delta} - 3\right)\right\},$$
$$T^f_{22} = -\frac{M^2}{4\pi k_f}\left(\frac{9\alpha\gamma}{4\Delta}\right)^u \left\{vB\theta + \beta\left(\frac{A\gamma + C\alpha}{\Delta} + \frac{\beta}{B} - 3\right)\right\},$$

where

$$\theta \equiv \left| -\frac{\alpha\gamma}{\Delta} + \frac{A\gamma + C\alpha}{\Delta}\left(3 - \frac{2\beta}{B}\right) + \frac{\beta}{B}\left(6 - \frac{\beta}{B}\right) - 6\right|. \tag{II.8'}$$

It can be immediately seen from the relation $R^g_{01} - \frac{1}{2}g_{01}R^g = k_g T^g_{01}$ that $D = 0$ or $B = \frac{2}{3}x^2$. Following the previous work[2] we shall take the second possibility, $B = \frac{2}{3}x^2$.

Further simplification is achieved by noticing that the relations $\alpha T^g_{00} + \gamma T^g_{11} = 0$ and $AT^f_{00} + CT^f_{11} + 0$ hold. These facts, together with Eqs. (II.1) and (II.6) and the analogous expression for the components of the g metric Ricci tensor, imply $\alpha\cdot\gamma = \text{const}$ and $\Delta = \text{const}$ (θ is also constant). We make the choice $\alpha\cdot\gamma = 1$ for simplicity. After some manipulation one gets the equations

$$-\frac{3\epsilon}{2x^2} + \frac{1}{\Delta x^2}(C + xC') = -\frac{M^2}{4\pi}\left(\frac{9}{4\Delta}\right)^u \left\{\frac{(1-v)}{\Delta} + \frac{3v}{4}\right\},$$
$$\frac{1}{3\Delta}\left(C'' + \frac{2C'}{x}\right) = -\frac{M^2}{4\pi}\left(\frac{9}{4\Delta}\right)^u \times \left\{\frac{2}{3}v\theta + \frac{A\gamma + C\gamma^{-1}}{\Delta} - \frac{3}{2}\right\}, \tag{II.9}$$
$$\frac{\epsilon}{x^2} - \frac{1}{x^2}(\gamma + x\gamma') = -\frac{M^2}{4\pi}\frac{k_g}{k_f}\left(\frac{4\Delta}{9}\right)^v \times \left\{-\frac{(1+u)}{\Delta} + \frac{3u}{4}\right\}.$$

The solution of the first equation is

$$C(x)=\frac{3\Delta}{2}\left\{\epsilon-\frac{2\mu_f}{x}-\frac{2\lambda}{9}x^2\right\}, \tag{II.10}$$

where λ is a constant given by

$$\lambda=\frac{M^2}{4\pi}\left(\frac{9}{4\Delta}\right)^u\left\{\frac{(1-v)}{\Delta}+\frac{3v}{4}\right\}, \tag{II.11}$$

and μ_f is an integration constant. Analogously, $\gamma(x)$ is found to be

$$\gamma(x)=\frac{2\mu_g}{x}-\frac{\Lambda}{3}x^3, \tag{II.12}$$

with the constant Λ given by

$$\Lambda=\frac{M^2}{4\pi}\frac{k_g}{k_f}\left(\frac{4\Delta}{9}\right)^v\left\{\frac{3u}{4}-\frac{(1+u)}{\Delta}\right\}, \tag{II.13}$$

and μ_g is another intergration constant. From the second equations in (II.9) and (II.10) we get

$$A\gamma+C\gamma^{-1}=\tfrac{3}{2}\Delta+\tfrac{2}{3}, \tag{II.14}$$

and, together with (II.12), we obtain

$$A(x)=\frac{1}{\epsilon-2\mu_g/x-\Lambda x^2/3}\times\left\{\frac{2}{3}+\frac{3\Delta}{2}\left[1-\frac{\epsilon-2\mu_f/x-2\lambda x^2/9}{\epsilon-2\mu_g/x-\Lambda x^2/3}\right]\right\}. \tag{II.15}$$

So the relations (II.10), (II.12), and (II.15) represent a complete set of exact solutions of the field equations (II.1).

III. DISCUSSION

It has been proposed[3] that hadrons can be interpreted as closed microuniverses generated by the strong f gravity metric. The geodesics associated with the f metric may provide a clue to understanding confinement in hadron physics. For that purpose we shall be concentrating on the possibility of having confining potentials in the case of our solution (II.10).

To simplify calculations, let us put the f metric in diagonal form by performing a coordinate transformation given by

$$d\tau=\left(\frac{3}{2}\right)^{1/2}\left(dw-\frac{D}{C}dx\right). \tag{III.1}$$

The f metric turns out to be

$$f_{\mu\nu}=\mathrm{diag}(\tfrac{2}{3}C,-\Delta/C,-\tfrac{2}{3}x^2,-\tfrac{2}{3}x^2F). \tag{III.2}$$

In order to consider a possible confinement let us solve a Klein–Gordon equation in the background f metric. Since our analysis will be only qualitative, we are considering a scalar hadron although the realistic Dirac equation can be exactly solved[5] in the f metric (in the case of a spherically symmetric metric).

The Klein–Gordon equation

$$\frac{1}{(-f)^{1/2}}\partial_\mu\left((-f)^{1/2}f^{\mu\nu}\partial_\nu\Phi\right)+m^2\Phi=0, \tag{III.3}$$

turns out to be

$$\frac{1}{C}\partial_\tau^2\Phi-\frac{2}{3\Delta x^2}\partial_x(Cx^2\partial_x\Phi)-\frac{1}{x^2F^{1/2}}\partial_y(F^{1/2}\partial_y\Phi)-\frac{1}{x^2F}\partial_z^2\Phi+\frac{2}{3}m^2\Phi=0. \tag{III.4}$$

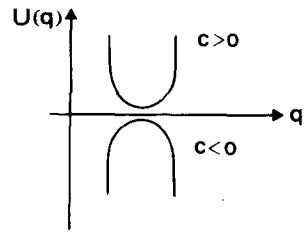


FIG. 1. Qualitative behavior of $U(q)$ when $\epsilon\neq0$ and sgn $\epsilon=-$ sgn Λ.

We can separate variables by writing

$$\Phi(x)=e^{-i\omega\tau}\frac{R(x)}{x}\eta(y)\zeta(z), \tag{III.5}$$

and the resulting equations are

$$\zeta(z)=e^{\pm ivz},$$

$$\frac{1}{\eta}(F^{1/2}\dot\eta)^{\cdot}+\frac{\nu^2}{F}=(F)^{1/2}l(l+1),$$

$$R''+\frac{C'}{C}R'+\frac{3\Delta}{2}\times\left[\frac{\omega^2}{C^2}-\frac{2m^2}{3C}-\frac{2}{3\Delta}\frac{C'}{Cx}+\frac{l(l+1)}{Cx^2}\right]R=0. \tag{III.6}$$

By a suitable change of variable $q=q(x)$ the last equation in (III.6) can be put in the Schrödinger form

$$R''(q)+[\omega^2-U(q)]R(q)=0, \tag{III.7}$$

with the potential given by

$$U(q)=\frac{2\omega^2}{3}C(q)+\left(\frac{2}{3\Delta}\right)^{1/2}\frac{C'(q)}{x(q)}-\frac{l(l+1)}{x^2(q)}C(q). \tag{III.7'}$$

This Schrödinger-type equation for the radial part can be studied qualitatively (at least) to look for the existence of bound states leading to (total or partial) quark confinement.

In the solution (II.10) we take $\mu_f=0$ and $\Lambda=2\lambda/9$. Since we have different choices for the constants ϵ and Λ the following possibilities arise:

(a) $\epsilon\neq0$ and ϵ and Λ have different signs.

The potential is given by

$$U(q)=\pm\frac{3\Delta}{2}|\Lambda|\left\{\frac{l(l+1)}{\sin^2 q}-\frac{2+2m^2/3|\Lambda|}{\cos^2 q}\right\},\quad q\in(0,\pi/2). \tag{III.8}$$

See Fig. 1.

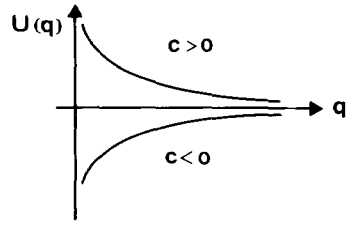


FIG. 2. Qualitative behavior of $U(q)$ when $\epsilon\neq0$ and sgn $\epsilon=$ sgn Λ.

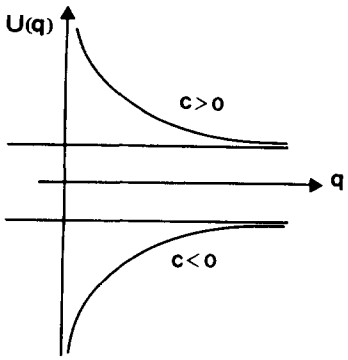


FIG. 3. Qualitative behavior of $U(q)$ when $\epsilon = 0$.

(b) $\epsilon \neq 0$, and ϵ and Λ have the same sign.

The potential is

$$U(q) = \pm \frac{3\Delta}{2}\left\{\frac{l(l+1)}{\sinh^2 q} + \frac{2 - 2m^2/|\Lambda|}{\cosh^2 q}\right\}, \quad q\in(0,\infty). \tag{III.9}$$

See Fig. 2.

(c) $\epsilon = 0$.

$$U(q) = \text{const} \pm \frac{2m^2}{3|\Lambda|}\frac{1}{q^2}, \quad q\in(0,\infty). \tag{III.10}$$

See Fig. 3.

As is obvious from this qualitative analysis, the only potential that increases infinitely and thus can give rise to discrete eigenvalues of the radial solution is (III.8) with positive sign (i.e., $\epsilon = 1$ and $\Lambda < 0$). It turns out that this is exactly the potential obtained in Ref. 3, where the explicit solutions for the eigenvalues can be found.

So although the exact classical solutions exist in the wider class of metrics (I.8), only the metric with positive curvature can give a confining potential in this simple picture and the idea of regarding hadrons as closed microuniverses of strong gravity is strongly supported. The metrics with negative and zero curvature (open and flat universes) can produce no confining potential. Besides, it follows that the choice of parameters relevant to the confining potential (III.8) automatically gives $C > 0$ and there can be no radiation in the sense of Hawking.[6] In Ref. 3 radiation is avoided by a special choice of parameters that turns out to be the only possibility.

This analysis lacks the presence of color that is naturally incorporated in the theory of strong interaction (QCD). However, the study of the f-g theory with color, either in the simple SU(2) form[7] or in the more general form incorporating Weyl symmetry,[8] showed the change of f metric to be of order $1/r^2$. This term is irrelevant for long-distance behavior where confinement occurs, and spin 1 gauge bosons are relevant to distinguish $\bar{q}q$ from qq states.

ACKNOWLEDGMENTS

We would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and United Nations Educational Scientific and Cultural Organization for hospitality at the International Centre for Theoretical Physics, Trieste, Italy, where part of this work was done. We also would like to thank Mrs. Lupita Estrada for carrying out the typing of the definitive version of this article and to Mr. José Rangel for his drawings.

Finally, one of us (R.H.) is indebted to Professor Marcos Rosenbaum for hospitality at the Centro de Estudios Nucleares, Universidad Nacional Autonoma de México, where part of this work was performed.

[1] C. J. Isham, A. Salam, and J. Strathdee, Phys. Rev. D **3**, 867 (1971).
[2] C. J. Isham and D. Storey, Phys. Rev. D **18**, 1047 (1978).
[3] A. Salam and J. Strathdee, Phys. Lett. **67B**, 429 (1977).
[4] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).
[5] E. Van Beveren, C. Dullemond, and T. A. Rijken, preprint THEF-NYM-79.11, Institute for Theoretical Physics, University of Nijmegen, The Netherlands, 1979.
[6] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).
[7] W. A. Sayed and A. Smailagic, Nuovo Cimento A **54**, 435 (1980).
[8] A. Salam and J. Strathdee, Phys. Rev. D **18**, 1323 (1978).

Gödel-type universe with a perfect fluid and a scalar field

S. K. Chakraborty and N. Bandyopadhyay
Presidency College, Physics Department, Calcutta, India

(Received 21 July 1981; accepted for publication 11 September 1981)

The paper contains, along with a brief review of solutions of general relativistic field equations when the metric is of a particular cylindrically symmetric stationary form, a new solution of the same general category when the energy-momentum tensor is due to a perfect fluid plus a scalar field. It turns out that under these constraints, the space-time is completely homogeneous and contains closed timelike lines. There is, however, a nonuniqueness in the interpretation as one can introduce a Maxwellian electromagnetic field of arbitrary strength along with the perfect fluid and the scalar field.

PACS numbers: 04.20.Jb, 04.20.Cv

I. INTRODUCTION

In a recent paper Raychaudhuri and Guha Thakurta[1] have shown that the stationary, cylindrically-symmetric line element

$$ds^2 = dt^2 - dr^2 - dz^2 - 2m(r)d\psi dt - l(r)d\psi^2 \qquad (I.1)$$

will represent a homogeneous space-time (i.e., admits four linearly independent Killing vectors) only when $m(r)$ and $l(r)$ satisfy the following conditions:

$$D \equiv (l + m^2)^{1/2} = A_1 e^{ar} + A_2 e^{-ar},$$
$$\frac{1}{D}\frac{dm}{dr} = C, \qquad (I.2)$$

or

$$D = Ar,$$
$$\frac{1}{D}\frac{dm}{dr} = C',$$

or

$$D = \text{const}, \qquad (I.3)$$

A_1, A_2, A, C, and C' being arbitrary constants. Specific choices of these constants yield the solutions of Gödel,[2] Ozsvath,[3] Som and Raychaudhuri,[4] Reboucas,[5] Novello,[6] and Gegenberg and Das.[7] The Gödel solution,

$$ds^2 = dt^2 - dr^2 - dz^2 + 2\sqrt{2}\sinh^2 r\, d\psi\, dt + (\sinh^4 r - \sinh^2 r)d\psi^2, \qquad (I.4)$$

represents the only homogeneous space-time with a perfect fluid content[8] (all other perfect fluid solutions, e.g., the Hoenselaers and Vishveshwara solution,[9,10] are reducible to the Gödel solution), the fluid obeying the equation of state: density = pressure (uniform). In the original version of his solution Gödel solved Einstein's field equations retaining the cosmological Λ term, and thereby arrived at a uniform dust distribution (vanishing pressure).

In the paper by Raychaudhuri and Guha Thakurta it has been shown that if the condition $C = \sqrt{2}\,a$ (which, incidentally, must be satisfied by all perfect fluid solutions) is relaxed, homogeneous space-times of the form (I.1) allows the introduction of an electromagnetic field along with uniform perfect fluid distribution, but the equation of state of the fluid is now changed to an inequality: density > pressure. The electromagnetic field may have a uniform distribution of sources or may satisfy source-free Maxwell's equations.

Ozsvath, retaining the cosmological Λ term in Einstein's field equations, and considering a material velocity vector (a Killing vector) different from the velocity vector in a co-moving system, has interpreted the homogeneous space-time of the form (I.1) as due to a uniform dust distribution along with an electromagnetic field satisfying source-free Maxwell's equations.

The Som and Raychaudhuri metric

$$ds^2 = dt^2 - dr^2 - dz^2 - (r^2 - a^2r^4)d\psi^2 + 2ard\psi dt \qquad (I.5)$$

is a homogeneous space-time of the form (I.1) with m and l satisfying (I.3). Here the universe contains a uniform distribution of charged dust with charge density = twice dust density (in general relativistic unit) and an associated electromagnetic field.

The Reboucas metric

$$ds^2 = dt^2 - dr^2 - dz^2 + \frac{4\Omega}{a}\cosh ar\, d\psi\, dt + \left[\frac{\Omega^2 + \alpha^2}{\Omega^2 - \alpha^2}\cosh^2 ar + 1\right]d\psi^2, \qquad (I.6)$$
$$\text{with } a^2 = 2(\Omega^2 - \alpha^2),$$

represents for $\alpha \neq 0$, a nonvanishing electromagnetic field along with a perfect fluid distribution. With $\alpha = 0$ the electromagnetic field vanishes and the solution is transformable to the Gödel solution.

Novello considered solutions of Einstein's field equations with a cosmological Λ term and an energy-momentum tensor corresponding to a vortex dominated non-Stokesian fluid which is characterized by a linear relationship between the anisotropic pressure (Π^i_j) and the vortex tensor Ω^i_j[$\equiv \omega^i \omega_j - (\omega^2/3)\delta^i_j$, ω^i being the vorticity vector of the fluid]. He arrived at the homogeneous space-time of the form (I.1) with

$$m(r) = \frac{2}{(\gamma - 2)^{1/2}}\cos \alpha r,$$
$$l(r) = -\left(\frac{\gamma^2 + 2}{\gamma^2 - 2}\right)\cos^2 \alpha r + 1,$$

where α is a constant and the constant γ relates Π^i_j and Ω^i_j: $\Pi^i_j = -\gamma^2 \Omega^i_j$. A perfect fluid coupled with an electromag-

 0022-2488/83/010129-04$02.50

netic field was shown by him to be a realization of such a vortex-dominated non-Stokesian fluid.

More recently Gegenberg and Das have found a class of exact solutions to the combined Einstein–Maxwell–Klein–Gordon field equations, thereby demonstrating the possibility of still another interpretation of sources of homogeneous space-time of the form (I.1). Their solutions,

$$ds^2 = -(dr^2 + dz^2 + r^2 d\psi^2) + [a(r)d\psi + dt]^2,$$

with $a(r) = \pm (8\pi)^{1/2} m r^2$, represents a universe consisting of a complex, charged, massive, Klein–Gordon field $\phi = \exp[i(L\psi + Et)]$, with mass m, along with an associated constant magnetic field $B = \pm 2^{1/2} mz + B_0$ along the symmetry axis (B_0, L, E are arbitrary parameters). The authors assumed validity of a sort of Weyl–Majumdar–Papapetrau condition. The magnetic field vanishes and the metric becomes static if the K.G. field is taken to be massless.

In view of all these facts it seemed worthwhile to investigate the possibility of existence of space-times of the form (I.1) with a combination of a perfect fluid and a real, uncharged K.G. field as its source. As shown in this paper such a possibility exists only if the K.G. field is massless (unlike the case of charged fields considered by Gegenberg and Das). It further turns out that the resulting space-time is homogeneous with the possibility of incorporating an electromagnetic field with or without a homogeneous distribution of sources in addition to the perfect fluid and K.G. field.

II. THE FIELD EQUATIONS FOR THE FLUID CUM SCALAR FIELD

With the usual Lagrangian for the massive scalar field as

$$L = -\tfrac{1}{2}[\phi_{,\mu}\phi^{,\mu} - M^2\phi^2],$$

the Einstein equations are

$$\begin{aligned} G_{\alpha\beta} &\equiv R_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} R \\ &= 8\pi[(p+\rho)v_\alpha v_\beta - p g_{\alpha\beta} \\ &\quad + \tfrac{1}{2}(\phi_{,\mu}\phi^{,\mu} - M^2\phi^2) g_{\alpha\beta} \\ &\quad - \tfrac{1}{2}\phi_{,\alpha}\phi_{,\beta}]. \end{aligned} \tag{II.1}$$

If v^μ is an eigenvector of $G_{\mu\nu}$ (we shall show in the last section that if v^μ is not an eigenvector of $G_{\mu\nu}$ then no solution of the desired type exists), then we have either

$$v^\mu \phi_{,\mu} = 0$$

or

$$\phi_{,\mu} = \alpha v_\mu. \tag{II.2}$$

With $\phi_{,\mu} = \alpha v_\mu$, v_μ is hypersurface orthogonal and hence by a coordinate transformation the line element (I.1) can be reduced to static form (with $m = 0$). Now for the line element (I.1), $A^\mu \equiv \delta^\mu_0$ is an eigenvector of $G_{\mu\nu}$. If we demand that the fluid velocity vector coincides with A^μ, i.e., the coordinate system is co-moving, we get

$$\phi_{,0} = 0 \tag{II.3}$$

Written out explicitly for the metric (I.1), with $x^1 = r$, $x^2 = z$, $x^3 = \psi$. The equations (II.1) now give

$$0 = 8\pi\left[\frac{p-\rho}{2} + \frac{1}{2}M^2\phi^2 - \phi_{,2}\phi^{,2}\right]D, \tag{II.4}$$

$$D_{11} - \frac{m_1^2}{2D} = 8\pi\left[\frac{p-\rho}{2} + \frac{1}{2}M^2\phi^2 - \phi_{,1}\phi^{,1}\right]D, \tag{II.5}$$

$$\frac{1}{2}\left(\frac{mm_1}{D}\right)_1 = 8\pi\left[\frac{\rho+3p}{2} + M^2\phi^2\right]D, \tag{II.6}$$

$$\frac{1}{2}\left(\frac{ml_1 - lm_1}{D}\right)_1 = 8\pi[-m(p+\rho) - \phi_{,3}\phi^{,0}]D, \tag{II.7}$$

$$\frac{1}{2}\left(\frac{m_1}{D}\right)_1 = 0, \tag{II.8}$$

$$\frac{1}{2}\left(\frac{l_1 + mm_1}{D}\right)_1 = 8\pi\left[\frac{p-\rho}{2} + \frac{1}{2}M^2\phi^2 - \phi_{,3}\phi^{,3}\right]D, \tag{II.9}$$

and the equation for the scalar field is

$$\Box\phi = -M^2\phi. \tag{II.10}$$

Again the vanishing of G_{12}, G_{13}, and G_{23} implies that only one of $\phi_{,1}$, $\phi_{,2}$, and $\phi_{,3}$ is nonzero.

In case $\phi_{,3} \neq 0$, $\phi_{,1} = \phi_{,2} = 0$, Eq. (II.10) gives $D =$ constant, which in turn implies $\phi =$ constant from (II.4), (II.5), and (II.9) and hence the scalar field vanishes.

If again $\phi_{,1} \neq 0$, $\phi_{,2} = \phi_{,3} = 0$ Eqs. (II.4) and (II.5) give

$$D_{11} - \frac{m_1^2}{2D} = 8\pi\phi_{,1}^2 D. \tag{II.11}$$

But from (II.4), (II.6), and (II.9)

$$D_{11} = 8\pi\left[\frac{\rho+3p}{2} + \frac{1}{2}M^2\phi^2\right]D, \tag{II.12}$$

and from (II.6) and (II.8)

$$\frac{m_1^2}{2D} = 8\pi\left[\frac{\rho+3p}{2} + \frac{1}{2}M^2\phi^2\right]D. \tag{II.13}$$

From (II.11)–(II.13), $\phi_{,1}$ vanishes.

Lastly let us consider $\phi_{,2} \neq 0$ and $\phi_{,1} = \phi_{,3} = 0$. Since the metric (I.1) obviously admits the Killing vector $\xi^\mu \equiv \delta^\mu_2$ the vanishing of the Lie derivative of $T^{\mu\nu}$ with respect to ξ^μ yields

$$\phi_{,2,2} = 0 \quad \text{or} \quad \phi = \alpha z, \tag{II.14}$$

where α is an arbitrary constant and a trivial constant of integration is omitted. Equation (II.10) now gives $M = 0$. In this case, (II.4), (II.6), and (II.7) imply p and ρ are constants and

$$\frac{m_1}{D} = 4\sqrt{\pi(2p + \alpha^2)} = c, \tag{II.15}$$

where c is a constant. From (II.6) and (II.9)

$$\frac{D_{11}}{D} = 16\pi p = a^2, \tag{II.16}$$

where a is a constant. The general solution of (II.16) is

$$D = A_1 e^{ar} + A_2 e^{-ar}, \tag{II.17}$$

A_1 and A_2 being arbitrary constants. In the case $p = 0$, we have

$$D = a'r \tag{II.18}$$

or

$$D = a'', \tag{II.19}$$

where a' and a'' are arbitrary constants. Using (II.15) and (II.17)

$$m = (c/a)(A_1 e^{ar} - A_2 e^{-ar} + B), \tag{II.20}$$

where B is an arbitrary constant. The pressure p and the density ρ are given by

$$p = \frac{a^2}{16\pi}, \quad \rho = \frac{a^2}{16\pi} + 2\alpha^2 = \frac{c^2}{8\pi} - \frac{3a^2}{16\pi}, \tag{II.21}$$

the constants a, c, and α being related by

$$c^2/2 - a^2 = 8\pi\alpha^2. \tag{II.22}$$

The solutions for m corresponding to (II.18) and (II.19) are, respectively,

$$m = \frac{ca'}{2} r^2 + B', \tag{II.23}$$

and

$$m = ca'' r + B'', \tag{II.24}$$

where B' and B'' are constants. With density ρ given by

$$\rho = 2\alpha^2 = c^2/8\pi. \tag{II.25}$$

III. ALTERNATIVE INTERPRETATION FOR THE METRIC

It would be noted that Eqs. (II.15) and (II.17) are identical with the condition deduced by Raychaudhuri and Guha Thakurta for homogenity so that the solutions we are seeking are all homogeneous. Further (II.21) requires $c^2 > 2a^2$, which is again the condition that these authors found for satisfying the Einstein equations with a distribution of perfect fluid and electromagnetic field (the electromagnetic field satisfying the Maxwell equations with or without source). It thus appears that these metrics admit an alternative interpretation—and one is tempted to ask whether one can combine all three to give a different interpretation. It is shown below that this is indeed possible. The field equations will now be

$$a^2 - \frac{c^2}{2} = 8\pi\left[\frac{p-\rho}{2} - \tau_1^1\right], \tag{III.1}$$

$$0 = 8\pi\left[\frac{p-\rho}{2} + \tau_2^2 + \alpha^2\right], \tag{III.2}$$

$$\frac{c^2}{2} = 8\pi\left[\frac{\rho+3p}{2} + \tau_0^0\right], \tag{III.3}$$

$$a^2 - \frac{c^2}{2} = 8\pi\left[\frac{p-\rho}{2} + \tau_3^3\right], \tag{III.4}$$

$$0 = \tau_0^3, \tag{III.5}$$

where τ_β^α represents the electromagnetic stress-energy tensor. Equations (III.1)–(III.5) along with the Rainich conditions yield

$$p = \frac{a^2}{16\pi}, \quad \rho = \frac{1}{16\pi}(c^2 - a^2) + \alpha^2, \tag{III.6}$$

$$\tau_1^1 = -\tau_2^2 = \tau_3^3 = -\tau_0^0 = \frac{1}{16\pi}\left[a^2 - \frac{c^2}{2} + 8\pi\alpha^2\right], \tag{III.7}$$

$$\tau_3^0 = \frac{m}{8\pi}\left[a^2 - \frac{c^2}{2} + 8\pi\alpha^2\right]. \tag{III.8}$$

The reality condition for the electromagnetic field ($\tau_0^0 > 0$) constrains ρ and p as

$$\rho > p + 2\alpha^2. \tag{III.9}$$

A non-unique interpretation of the sources of the electromagnetic field is possible. One choice is

$$F^{13} = -F^{31} = \pm\frac{1}{2D}(c^2 - 2a^2 - 16\pi\alpha^2)^{1/2},$$
$$F^{01} = -F^{10} = mF^{31}, \tag{III.10}$$
$$\sigma(\equiv\text{charge density}) = \frac{c}{8\pi}(c^2 - 2a^2 - 16\pi\alpha^2)^{1/2},$$

which corresponds to a homogeneous distribution of sources. Another choice is

$$F^{13} = -F^{31} = \pm\frac{1}{2D}(c^2 - 2a^2 - 16\pi\alpha^2)^{1/2}\cos\theta,$$
$$F^{20} = -F^{02} = \pm\tfrac{1}{2}(c^2 - 2a^2 - 16\pi\alpha^2)^{1/2}\sin\theta, \tag{III.11}$$
$$F^{10} = -F^{01} = \pm\frac{m}{2D}(c^2 - 2a^2 - 16\pi\alpha^2)^{1/2}\cos\theta,$$

which satisfy source-free Maxwell equations for $\theta = -cz$.

IV. EXISTENCE OF CLOSED TIMELIKE LINES

The coordinate ψ in (I.1) can be treated as an angular coordinate if $g_{33} \to 0$ and $g_{33}/g_{11} \to r^2$ as $r \to 0$. One can then use the transformations $X = r\cos\psi$, $Y = r\sin\psi$ and obtain analyticity of the metric at $r = 0$.[11] The choice of constants for which ψ can be treated as an angular coordinate is seen to be

$$A_1 + A_2 = 0 \text{ and } B = A_2 - A_1. \tag{IV.1}$$

With this choice of constants, the t, z, r constant lines are closed timelike lines for

$$r > \frac{2}{a}\coth^{-1}\frac{2c}{a} \quad (A_1 > 0)$$

or (IV.2)

$$r < \frac{2}{a}\coth^{-1}\frac{2c}{a} \quad (A_1 < 0).$$

V. ABSENCE OF STATIONARY SOLUTIONS WHEN THE FLOW VECTOR IS NOT AN EIGENVECTOR OF $G_{\mu\nu}$

Since $A^\mu \equiv \delta_0^\mu$ is an eigenvector of $G_{\mu\nu}$ (with eigenvalue λ, say), contraction of $G_{\mu\nu}$ with A^μ yields

$$(p+\rho)v_0 v_1 - \phi_{,0}\phi_{,1} = 0, \tag{V.1}$$
$$(p+\rho)v_0 v_2 - \phi_{,0}\phi_{,2} = 0, \tag{V.2}$$
$$8\pi\left[(p+\rho)v_0^2 - p - \tfrac{1}{2}\phi_{,\alpha}\phi^{,\alpha} - \phi_{,0}^2\right] = \lambda, \tag{V.3}$$
$$8\pi\left[(p+\rho)v_0 v_3 + mp - \frac{m}{2}\phi_{,\alpha}\phi^{,\alpha} - \phi_{,0}\phi_{,3}\right] = -m\lambda. \tag{V.4}$$

Equations (V.1)–(V.4) give

$$\frac{v_1}{\phi_{,1}} = \frac{v_2}{\phi_{,2}} = \frac{mv_0 + v_3}{m\phi_{,0} + \phi_{,3}} = \frac{\phi_{,0}}{(p+\rho)v_0} = \Theta, \text{ say} \tag{V.5}$$

where Θ may not be constant. For $\Theta \neq 0$, (V.5) can be rewritten as

$$v_1 = \Theta\phi_{,1}, \quad v_0 = \phi_{,0}/\Theta(p+\rho), \tag{V.6}$$

$$v_2 = \Theta\phi_{,2}, \quad v_3 = \Theta\phi_{,3} + m\phi_{,0}\left[\Theta - \frac{1}{(p+\rho)\Theta}\right].$$

For $\Theta^2 = 1/(p+\rho)$, $v_\mu = \Theta\phi_{,\mu}$ and then the R^0_0 and R^3_0 equations imply $m =$ constant, and so by a transformation of the ψ coordinate (I.1) can be reduced to static form. The other case of interest with $\Theta \neq 0$ is

$$v_1 = v_2 = 0 = \phi_{,1} = \phi_{,2}, \tag{V.7}$$

$$v_0 = \frac{\phi_{,0}}{\Theta(p+\rho)}, \quad v_3 = \Theta\phi_{,3} + m\phi_{,0}\left[\Theta - \frac{1}{\Theta(p+\rho)}\right],$$

and the corresponding field equations are

$$p = \rho, \tag{V.8}$$

$$D_{11} - \frac{m_1^2}{2D} = 0, \tag{V.9}$$

$$\left(\frac{l_1 + mm_1}{2D}\right)_1 = 8\pi D(p'v^3v_3 - \phi_{,3}\phi^{,3}), \tag{V.10}$$

$$\left(\frac{mm_1}{2D}\right)_1 = 8\pi D(p'v_0v^0 - \phi_{,0}\phi^{,0}), \tag{V.11}$$

$$\left(\frac{m_1}{2D}\right)_1 = 8\pi D(p'v_0v^3 - \phi_{,0}\phi^{,3}), \tag{V.12}$$

$$\left(\frac{ml_1 - lm_1}{2D}\right)_1 = 8\pi D(p'v_3v^0 - \phi_{,3}\phi^{,0}), \tag{V.13}$$

and $\Box\phi = 0 = m^2\phi_{,0,0} + 2m\phi_{,0,3} - D^2\phi_{,0,0} + \phi_{,3,3}$, (V.14)

where $p' = p + \rho = 2p$.

The normality condition for the flow vector yields the further equation

$$\frac{\phi^2_{,0}}{\Theta p'^2} - \frac{\Theta^2}{D^2}(m\phi_{,0} + \phi_{,3})^2 = 1. \tag{V.15}$$

From (V.6) and (V.12)

$$\left(\frac{m_1}{D}\right)_1 = 0 \quad \text{or} \quad \frac{m_1}{D} = k, \tag{V.16}$$

where k is a constant. From (V.9) and (V.16)

$$D = A_1e^{k'r} + A_2e^{-k'r}, \tag{V.17}$$

so that

$$m = 2(A_1e^{k'r} - A_2e^{-k'r} + \text{const}), \tag{V.18}$$

where $2k'^2 = k^2$.

Again a rather involved manipulation using (V.9), (V.10), (V.11), and (V.15) yields

$$k'^2/8\pi + \phi^2_{,0} = p', \tag{V.19}$$

$$\phi^2_{,0} = \Theta^2p'^2, \tag{V.20}$$

and

$$p'(\phi_{,3} + m\phi_{,0})^2 = 0. \tag{V.21}$$

(V.21) implies either

$$p' = 0 \tag{V.22}$$

or

$$m\phi_{,0} + \phi_{,3} = 0. \tag{V.23}$$

Condition (V.22) implies, in view of (V.19) and (V.20), $m =$ const, and so that metric is transformable to the static form. Condition (V.23) implies, since $\phi = \phi(\psi,t)$, $m =$ const, with obvious conclusion. For the particular case $\Theta = 0$, Eqs. (V.5), (V.14), and (V.10) imply $\phi_{,0} = \phi_{,3} = 0$, so that the scalar field vanishes.

ACKNOWLEDGMENT

The authors are thankful to Professor A. K. Raychaudhuri for suggesting the problem and extending valuable comments.

[1] A. K. Raychaudhuri and S. N. Guha Thakurta, Phys. Rev. D **22**, 802 (1980).
[2] K. Gödel, Rev. Mod. Phys. **21**, 447 (1949).
[3] I. Ozsvath, *Perspective in Geometry and Relativity* (Indiana University, Bloomington, Indiana, 1966).
[4] M. M. Som and A. K. Raychaudhuri, Proc. R. Soc. London Ser. A **304**, 81 (1968).
[5] M. J. Reboucas, Centro Brasileiro De Pesquisas Fisicas Report No. A0023 (1978, unpublished).
[6] M. Novello, Centro Brasileiro De Pesquisas Fisicas Report No. A0006 (1979, unpublished).
[7] J. D. Gegenberg and A. Das, J. Math. Phys. **22**, 1736 (1981).
[8] I. Ozsvath, J. Math. Phys. **6**, 590 (1965).
[9] C. Hoenselaers and C. V. Vishveshwara, Gen. Relativ. Gravit. **10**, 43 (1979).
[10] S. K. Chakraborty, Gen. Relativ. Gravit. (in press).
[11] S. C. Maitra, J. Math. Phys. 7, 1025 (1966).

Existence of solutions of integral equations in the thermodynamics of one-dimensional fermions with repulsive delta function potential

C. K. Lai
14 Pond Crest Road, Danbury, Connecticut 06810

(Received 12 July 1977; accepted for publication 25 September 1981)

We prove by an iteration scheme that the solutions of coupled integral equations in the thermodynamics of the fermions in one dimension with delta function potential exist.

PACS numbers: 05.30.Fk

I. INTRODUCTION

Since the early 1950s, people have searched for one dimensional models of many body systems that are mathematically attractable. In 1950, Tomonaga first proposed a model for the electron gas that can be solved exactly.[1] The model treats the elementary excitations as bosons and further assumes that the momentum transfer equals the energy transfer. Obviously the last condition is an approximation though it is argued that the condition would be satisfied near the fermi surface. A more realistic model with conventional quadratic kinetic energy is the delta function model of the Hamiltonian

$$H = -\sum \frac{\partial^2}{\partial x_i^{\,2}} + 2c\sum_{i>j}\delta(x_i - x_j), \quad c>0. \tag{1}$$

The ground state energy of (1) for the bosons was first obtained by Lieb and Liniger[2] in 1965, assuming that the wave function is a finite sum of plane waves with coefficients to be determined from some transcendental equations. That is, the wave function is assumed of the form

$$\psi = \sum \alpha_P \exp\left(ip_{P_1}x_1 + ip_{P_2}x_2 + \cdots + ip_{P_N}x_N\right) \tag{2}$$

if

$$x_1 < x_2 < \cdots < x_N, \tag{3}$$

where $P_1, P_2, \ldots P_N$ is a permutation of $1,2,\ldots,N$. The coefficients α_P are to be determined by matching the wave function (2) with the wave functions in other regions than (3) and a set of transcendental equations thus result. This assumption is known as Bethe's hypothesis. It was originated by Bethe[3] in the early forties when he studied the Hamiltonian of the ferromagnets

$$H = -\frac{1}{2}\sum(\sigma_x\sigma_x' + \sigma_y\sigma_y' + \sigma_z\sigma_z'), \tag{4}$$

where the σs are the spin operators. Yet it is surprising that the hypothesis was successfully applied later to a number of quantum models in one dimension and classical models in two dimensions. These include the anisotropic ferromagnetic model, the quantum lattice gas model, the ice model, the ferroelectric model, and the delta function model mentioned earlier.

Now the fermion case of the delta function model (1) imposed a greater difficulty than the boson case because of the fermi statistic symmetries required for its wave functions. McGuire[4] and Flick and Lieb[5] first solved the one spin down and two spin down cases. Later Gaudin[6] and Yang[7] obtained the solution for the n spin down problem and more generally Yang classified all the solutions of any statistics using group theoretical approach. At about the same time, some other problems of the delta function model have also been solved. These include the S matrix for any finite number of particles,[8] the ground state energy of the fermion-boson mixture,[9] and the thermodynamics and excitation spectrum at finite temperature for the bosons.[10]

The thermodynamics in the fermion case again imposed further complication than the boson case. For now it is more difficult to determine all the excited states of the system. To achieve this, one has to locate all the solutions of the transcendental equations in Bethe's hypothesis. This was finally solved by the present author in a previous paper[11] showing that the roots of the transcendental equations were lying in strings in the complex plane. More specifically, the transcendental equations in Bethe's hypothesis for the case of N spin down are given by

$$e^{ipL} = \prod_{\Lambda'}\left(\frac{-p+\Lambda'-ic}{-p+\Lambda'+ic}\right), \tag{5}$$

$$\prod_{p'}\left(\frac{-p'+\Lambda-ic/2}{-p'+\Lambda+ic/2}\right) = -\prod_{\Lambda'}\left(\frac{-\Lambda'+\Lambda-ic}{-\Lambda'+\Lambda+ic}\right),$$

where the p's are the local momenta and the Λ's (M in number) are some auxiliary variables. In the ground state, the p's and Λ's are real numbers. In the excited states, the Λ's are complex numbers in the form of strings:

$$\Lambda = \xi + i\mu\eta + O(e^{-kL}), \quad \mu = -(m-1), \\ -(m-3),\ldots,(m-1), \tag{6}$$

where $\eta = c/2$, $k>0$ is a certain number, and ξ is real. The integer m defines the length of the string that contains m complex numbers. Thus the two numbers ξ and m define a string uniquely. Let $C(\xi,n)$ denote such a string. By the use of (6), (5) become equations for the p's and ξ's and in the limit L, N, $M\to\infty$ proportionally, they yield the following integral equations for the density functions of p's and $C(\xi,m)$:

$$\frac{1}{2}\pi = \rho + \rho_h - \frac{1}{2}\int_{-\infty}^{\infty} G_1(p-k)\rho\,dk \\ + \frac{1}{2}\int_{-\infty}^{\infty} G_0(p-k)\sigma_{1,h}\,dk, \tag{7}$$

$$\sigma_n + \sigma_{n,h} = \frac{1}{2}\int_{-\infty}^{\infty} G_0(p-k)(\sigma_{n+1,h} + \sigma_{n-1,h})\,dk, \quad n\geqslant 1$$

 0022-2488/83/010133-05$02.50

where $\sigma_{0,h} = \rho$ and

$$G_n(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega k} e^{-n\eta|\omega|}}{\cosh \eta\omega}\, d\omega. \tag{8}$$

The σ_m, $\sigma_{m,h}$ are defined by

$N\sigma_m\, d\xi$ = the number of ξ's for strings $c(\xi,m)$ in $[\xi,\xi + d\xi]$,

$N\sigma_{m,h}\, d\xi$ = the number of "holes" for the above ξ's in $[\xi,\xi + d\xi]$, (9a)

and the ρ and ρ_h are the density functions for the p's and its holes, respectively,

$\rho\, dk$ = the number of p's in the interval $[k,k + dk]$,

$\rho_h\, dk$ = the number of holes for the above p's in $[k,k + dk]$. (9b)

If one defines

$$\rho_h/\rho = \exp(\epsilon(p)/T), \quad \sigma_{m,h}/\sigma_m = \exp(\varphi_m(k)/T) \tag{10}$$

and minimizes the free energy $(E - TS)/L$ [subject to fixed N/L and $(N - 2M)/L$] to obtain the equilibrium distributions of ρ's and σ_m's, one arrives at the following equations for $\epsilon(k)$ and $\varphi_m(k)$:

$$A = p^2 - \epsilon - \frac{T}{2}\int_{-\infty}^{\infty} G_1(p-k)\ln(1 + e^{-\epsilon/T})dk - \frac{T}{2}\int_{-\infty}^{\infty} G_0(p-k)\ln(1 + e^{\varphi_1/T})dk, \tag{11a}$$

$$\varphi_1 = \frac{T}{2}\int_{-\infty}^{\infty} G_0(p-k)[\ln(1 + e^{\varphi_1/T}) - \ln(1 + e^{-\epsilon/T})]\, dk, \tag{11b}$$

$$\varphi_\nu = \frac{T}{2}\int_{-\infty}^{\infty} G_0(p-k)[\ln(1 + e^{\varphi_{\nu+1}/T}) + \ln(1 + e^{\varphi_{\nu-1}/T})]\, dk, \quad \nu \geqslant 2, \tag{11c}$$

$$\lim_{\nu\to\infty} \varphi_\nu = \nu\lambda T. \tag{11d}$$

The free energy F and the pressure P are then given by

$$\frac{F}{L} = A\frac{N}{L} - \frac{T}{2\pi}\int_{-\infty}^{\infty} \ln(1 + e^{-\epsilon/T})\, dk - \lambda\left(\frac{N-2M}{L}\right), \tag{12}$$

$$P = (T/2\pi)\int_{-\infty}^{\infty} \ln(1 + e^{-\epsilon/T})\, dk, \tag{13}$$

where A is the chemical potential. Here λ can be seen as the magnetic field. [The Hamiltonian (1) for a fermion system will have an additional term $\lambda(N - 2M)$, with $N - 2M$ being the total spin of the system.] Thus the solutions of (11a)–(11d) determine all the thermodynamic quantities.

It can easily be shown that in the limit $c\to 0$ (free fermions) and $c\to\infty$ (free fermions of only one species), (11a)–(11d) give the correct temperature distribution functions. It also yields the correct second virial coefficients for any value of c. But unlike the boson system the general existence of solutions in (11a)–(11d) has not been established. In the boson case, a single integral equation of similar nature was shown by Yang and Yang[10] to have solutions by iteration. The present author has long suspected that an iteration procedure can also be applied to (11a)–(11d) though it may not be obvious at first sight. In this paper, I prove that (11a)–(11d) can indeed be solved by iterations when the chemical potential A is negative and the magnetic field $\lambda = 0$. This not only establishes the existence of solutions for the case mentioned but also guarantees the convergence of the iteration method in numerical computation.

In the following, the existence theorems will be described in detail. The method of proof is quite extraordinary as it involves repeated use of inductions and seems to be the first time that it is applied to the coupled integral equations of type (11a)–(11d).

II. EXISTENCE THEOREMS

We use an iteration scheme to prove the existence of solutions for equations (11a)–(11d). The iteration is defined in operator forms as follows (T is scaled to be 1):

$$\epsilon^{(0)} = -A + p^2,$$
$$\varphi_\nu^{(0)} = 0, \quad \nu \geqslant 1,$$
$$\epsilon^{(1)} = -A + p^2 - \frac{G_1}{2}\ln(1 + e^{-\epsilon^{(0)}}) - \frac{G_0}{2}\ln[1 + \exp(\varphi_1^{(0)})],$$
$$\varphi_1^{(1)} = \frac{G_0}{2}\ln(1 + \exp\varphi_2^{(0)}) - \frac{G_0}{2}\ln[1 + \exp(-\epsilon^{(0)})],$$
$$\varphi_\nu^{(1)} = \frac{G_0}{2}\ln(1 + \exp\varphi_{\nu-1}^{(0)}) + \tfrac{1}{2}G_0\ln(1 + \exp\varphi_{\nu+1}^{(0)}), \quad \nu \geqslant 2 \tag{14}$$

and so on for $\epsilon^{(2)}$, $\varphi_\nu^{(2)}$, etc. Here we use G_0 and G_1 as the integral operators with kernals $G_0(p-k)$, $G_1(p-k)$ defined in (11a)–(11d). We also assume that the chemical potential A is negative and the magnetic field λ is zero in (11a)–(11d)

$$A < 0, \quad \lambda = 0.$$

That is, we are looking for solutions with boundary conditions

$$\lim_{\nu\to\infty} \varphi_\nu/\nu = 0.$$

Then for the iteration (14), the following theorems hold:

Theorem 1: In the iteration (14), $\epsilon^{(n)}$ forms a decreasing sequence

$$\epsilon^{(0)} > \epsilon^{(1)} > \epsilon^{(2)} > \cdots > \epsilon^{(n)} > \cdots, \tag{15}$$

and $\varphi_\nu(n)$ forms an increasing sequence:

$$\varphi_\nu^{(0)} \leqslant \varphi_\nu^{(1)} \leqslant \varphi_\nu^{(2)} \leqslant \cdots \leqslant \varphi_\nu^{(n)} \leqslant \cdots \tag{16}$$

(the sign $<$ in (16) holds strictly for $n \geqslant \nu$).

Theorem 2: $\epsilon^{(n)}$ is bounded below and $\varphi_\nu^{(n)}$ is bounded above.

Theorem 3:

$$\lim_{n\to\infty} \epsilon^{(n)}(p) = \epsilon(p)$$

and

$$\lim_{n\to\infty} \varphi_\nu^{(n)}(p) = \varphi_\nu(p) \tag{17}$$

exist and satisfy Eqs. (11a)–(11d).

Theorem 3 is a straightforward consequence of Theorems 1 and 2. Their proofs are given in Secs. III and IV.

III. PROOF OF THEOREM 1

The proof is based on several lemmas.

Lemma 1: The kernels $G_0(p-k)$ and $G_1(p-k)$ are positive functions

$$G_0(p-k)>0, \quad G_1(p-k)>0 \tag{18}$$

and

$$G_0 f \equiv \int_{-\infty}^{\infty} G_0(p-k) f\, dk = f,$$
$$G_1 f \equiv \int_{-\infty}^{\infty} G_1(p-k) f\, dk = f \tag{19}$$

if f is a constant function.

Proof: From (8), the kernel G_0 has the following closed form:

$$G_0(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-i\omega k}}{\cosh\eta\omega}\, d\omega = \frac{1}{2\eta\cosh(\pi K/2\eta)} > 0. \tag{20}$$

Also, the Fourier transform of G_1 is given by

$$\tilde{G}_1(\omega) = \frac{e^{-\eta|\omega|}}{\cosh\eta\omega} = \tilde{G}_0(\omega) e^{-\eta|\omega|}. \tag{21}$$

Thus by convolution,

$$G_1(k) = \int_{-\infty}^{\infty} G_0(k-k') \frac{\eta}{\eta^2 + k'^2}\, dk' > 0. \tag{22}$$

Now

$$\int_{-\infty}^{\infty} G_0(p-k)\, dk = \tilde{G}_0(0) = 1,$$
$$\int_{-\infty}^{\infty} G_1(p-k)\, dk = \tilde{G}_1(0) = 1. \tag{23}$$

Thus (19) is true if f is a constant function. The positiveness of G_0 and G_1 in (18) is very crucial and will be used in the proofs of all the remaining lemmas and theorems.

Lemma 2: The inequality

$$(1+e^{X'})/(1+e^{X}) \geqslant (1+e^{Y'})/(1+e^{Y}) \tag{24}$$

holds if

$$X \geqslant Y,$$
$$(X'-X) \geqslant (Y'-Y) \geqslant 0. \tag{25}$$

The proof is straightforward if one takes the logarithm of (24) and notes that the derivative of $\ln(1+e^{X})$ is an increasing function.

Lemma 3: Suppose x and x' are constants and $\omega,\omega',\varphi,\varphi'$ are functions satisfying the inequalities

$$\omega \leqslant x \leqslant \varphi, \tag{26}$$

$$\omega' - \omega \leqslant x' - x \leqslant \varphi' - \varphi. \tag{27}$$

Then the following inequality:

$$G_1 \ln[(1+\exp\omega')/(1+\exp\omega)]$$
$$< G_0 \ln[(1+\exp\varphi')/(1+\exp\varphi)] \tag{28}$$

holds for the integral operators G_0 and G_1 defined in (8).

Proof: By Lemma 1 and Lemma 2, one has

$$G_1 \ln\left[\frac{1+\exp\omega'}{1+\exp\omega}\right] \leqslant G_1 \ln\left[\frac{1+\exp x'}{1+\exp x}\right]$$
$$= G_0 \ln\left[\frac{1+\exp x'}{1+\exp x}\right]$$
$$\leqslant G_0 \ln\left[\frac{1+\exp\varphi'}{1+\exp\varphi}\right]. \tag{29}$$

This completes the proof. Note that in (29), the detailed comparison of the kernels G_0 and G_1 is not required once the inequalities of type (26) and (27) are established. This lemma will be used later when we deal with the iterations $\epsilon^{(n)}$ in (14), where both G_0 and G_1 are present.

Lemma 4: For fixed $n>0$ the following holds:

$$\varphi_1^{(n)} \leqslant \varphi_2^{(n)} \leqslant \cdots \leqslant \varphi_\nu^{(n)} \leqslant \cdots; \tag{30}$$

furthermore the sign $<$ holds strictly for $\nu \leqslant n$. The proof by induction is straightforward.

Lemma 5: For a fixed m, assume that

$$\epsilon^{(1)} > \epsilon^{(2)} > \cdots > \epsilon^{(m)}, \tag{31}$$
$$\varphi_\nu^{(1)} \leqslant \varphi_\nu^{(2)} \leqslant \cdots \leqslant \varphi_\nu^{(m)}, \quad \nu \geqslant 1. \tag{32}$$

Define

$$f_\nu^{(n)} = (1+\exp\varphi_\nu^{(n)})/(1+\exp\varphi_\nu^{(n-1)}), \quad \nu \geqslant 1,$$
$$f_0^{(n)} = [1+\exp(-\epsilon^{(n-1)})]/[1+\exp(-\epsilon^{(n)})]. \tag{33}$$

Then

$$f_{\nu+1}^{(n)} \geqslant f_\nu^{(n)} \tag{34}$$

holds for $n \leqslant m+1$.

Proof by induction: Suppose (34) holds for $n \leqslant j \leqslant m$. Then take

$$\varphi_{\nu+1}^{(j+1)} - \varphi_{\nu+1}^{(j)} = (G_0/2)\ln(f_{\nu+2}^{(j)} f_\nu^{(j-1)}), \tag{35}$$
$$\varphi_\nu^{(j+1)} - \varphi_\nu^{(j)} = (G_0/2)\ln(f_{\nu+1}^{(j)} f_{\nu-1}^{(j-1)}). \tag{36}$$

By the induction hypothesis, one has

$$\varphi_{\nu+1}^{(j+1)} - \varphi_{\nu+1}^{(j)} \geqslant \varphi_\nu^{(j+1)} - \varphi_\nu^{(j)}$$

and by Lemma 1 and Lemma 2, (34) holds for $n = j+1$. It is easy to show that (34) holds for $n=1$. This completes the induction. [(31) is required for the case $\nu = 1$.]

Lemma 6: For a fixed m, assume that

$$\varphi_\nu^{(1)} \leqslant \varphi_\nu^{(2)} \leqslant \cdots \leqslant \varphi_\nu^{(m)}, \quad \nu \geqslant 1. \tag{37}$$

Define the following integral equations:

$$y_1 = (G_0/2)\ln(1+\exp y_2),$$
$$y_2 = (G_0/2)\ln(1+\exp y_3)(1+\exp y_1),$$
$$\langle \vdots \rangle$$
$$y_\nu = (G_0/2)\ln(1+\exp y_{\nu+1}) \times (1+\exp y_{\nu-1}), \quad \nu \geqslant 2, \tag{38}$$

and so forth. Let us also define the iterations for the above equations by

$$y_1^{(0)} = y_2^{(0)} = \cdots = y_\nu^{(0)} = 0, \quad \nu \geqslant 1, \tag{39}$$

$$y_1^{(1)} = (G_0/2)\ln(1+\exp y_1^{(0)}) = \tfrac{1}{2}\ln 2,$$

$$y_\nu^{(1)} = (G_0/2)\ln(1+\exp y_{\nu-1}^{(0)}) \times (1+\exp y_{\nu+1}^{(0)}) = \ln 2, \quad \nu \geqslant 2, \tag{40}$$

and so forth, for $y_\nu^{(2)}, y_\nu^{(3)}, \cdots$. Thus except for y_1, y_ν satisfies the same integral equations as φ_ν in (11c) and the iterations $y_\nu^{(n)}$ are sequences of constant numbers. Then the following inequalities:

$$\varphi_\nu^{(n)} \leqslant y_\nu^{(n)} \leqslant \varphi_{\nu+1}^{(n)}, \tag{41}$$

$$\varphi_\nu^{(n)} - \varphi_\nu^{(n-1)} \leqslant y_\nu^{(n)} - y_\nu^{(n-1)} \leqslant \varphi_{\nu+1}^{(n)} - \varphi_{\nu+1}^{(n-1)} \tag{42}$$

hold for $n \leqslant m$.

Proof by induction: Comparing (39) with (14), it is easy to establish that (41) and (42) hold for $n = 0,1$. Suppose they hold for $n < m-1$. Then by Lemmas 1 and 2,

$$\begin{aligned}
\varphi_\nu^{(n+1)} - \varphi_\nu^{(n)} &= \frac{G_0}{2}\ln\left[\frac{1+\exp\varphi_{\nu+1}^{(n)}}{1+\exp\varphi_{\nu+1}^{(n-1)}}\right] \\
&\quad + \frac{G_0}{2}\ln\left[\frac{1+\exp\varphi_{\nu-1}^{(n)}}{1+\exp\varphi_{\nu-1}^{(n-1)}}\right] \\
&< (\text{the same form as above with the } \varphi\text{'s} \\
&\qquad \text{replaced by the } y\text{'s}) \\
&= y_\nu^{(n+1)} - y_\nu^{(n)} \\
&\leqslant \frac{G_0}{2}\ln\left[\frac{1+\varphi_{\nu+2}^{(n)}}{1+\exp\varphi_{\nu+2}^{(n-1)}}\right] \\
&\quad + \frac{G_0}{2}\ln\left[\frac{1+\exp\varphi_\nu^{(n)}}{1+\exp\varphi_\nu^{(n-1)}}\right] \\
&= \varphi_{\nu+1}^{(n+1)} - \varphi_{\nu+1}^{(n)}.
\end{aligned} \tag{43}$$

Similarly, one can show that (41) holds for $n \leqslant m$. This completes the proof.

Lemma 7: Suppose (31) and (32) are true so that up to $n = m$, $-\epsilon^{(n)}$ and $\varphi_\nu^{(n)}$ are increasing sequences. Define constants $x^{(n)}$ such that

$$x^{(0)} = 0$$

$$x^{(n)} = \tfrac{1}{2}\ln(1+\exp x^{(n-1)}) + \tfrac{1}{2}\ln(1+\exp y^{(n-1)}), \quad n \geqslant 1, \tag{44}$$

where $y^{(n)}$ are the sequences of constants defined in (40). Then

$$-\epsilon^{(n)} + \epsilon^{(n-1)} \leqslant x^{(n)} - x^{(n-1)} \leqslant \varphi_3^{(n)} - \varphi_3^{(n-1)}, \tag{45}$$

$$-\epsilon^{(n-1)} \leqslant x^{(n-1)} \leqslant \varphi_3^{(n-1)} \tag{46}$$

hold for $n \leqslant m$.

Proof by induction: Suppose (45) and (46) hold for $n \leqslant m-1$. Then by Lemma 3 and Lemma 6 and Eq. (34), one has

$$\begin{aligned}
-\epsilon^{(n+1)} + \epsilon^{(n)} &= \frac{G_1}{2}\ln[(1+\exp(-\epsilon^{(n)}))/ \\
&\quad \times(1+\exp(-\epsilon^{(n-1)}))] \\
&\quad + \frac{G_0}{2}\ln[(1+\exp\varphi_1^{(n)})/(1+\exp\varphi_1^{(n-1)})] \\
&\leqslant \frac{G_1}{2}\ln\left[\frac{1+\exp x^{(n)}}{1+\exp x^{(n-1)}}\right] \\
&\quad + \frac{G_0}{2}\ln\left[\frac{1+\exp y_1^{(n)}}{1+\exp y_1^{(n-1)}}\right] \\
&= (x^{(n+1)} - x^{(n)}) \\
&\leqslant \frac{G_0}{2}\ln f(\varphi_3^{(n)}) \\
&\quad + \frac{G_0}{2}\ln f(\varphi_2^{(n)}) \\
&\leqslant \frac{G_0}{2}\ln f(\varphi_4^{(n)}) \\
&\quad + \frac{G_0}{2}\ln f(\varphi_2^{(n)}) \\
&= \varphi_3^{(n+1)} - \varphi_3^{(n)}.
\end{aligned} \tag{47}$$

Similarly one can show that (46) holds for $n \leqslant m$. This completes the induction.

Now we can prove Theorem 1 by induction. It is easy to show that the inequalities (15) and (16) hold for $n = 0,1$. Suppose it holds for $n \leqslant m$; then it is obvious from (14) that

$$\epsilon^{(m)} < \epsilon^{(m+1)}, \tag{48}$$

$$\varphi_\nu^{(m)} \leqslant \varphi_\nu^{(m+1)}, \quad \nu \geqslant 2, \tag{49}$$

as G_0 and G_1 are positive. Thus it remains to show that

$$\varphi_1^{(m)} \leqslant \varphi_1^{(m+1)} \quad \text{for } \nu = 1.$$

Now by Lemma 3 and Lemma 7,

$$\begin{aligned}
-\epsilon_2^{(m)} + \epsilon_1^{(m-1)} &= \frac{G_1}{2}\ln\frac{1+\exp(-\epsilon^{(m-1)})}{1+\exp(-\epsilon^{(m-2)})} \\
&\quad + \frac{G_0}{2}\ln\left(\frac{1+\exp\varphi_1^{(m-1)}}{1+\exp\varphi_1^{(m-2)}}\right) \\
&\leqslant \frac{G_0}{2}\ln f(\varphi_3^{(m-1)}) \\
&\quad + \frac{G_0}{2}\ln f(\varphi_1^{(m-1)}) \\
&= \varphi_2^{(m)} - \varphi_2^{(m-1)}.
\end{aligned} \tag{50}$$

Similarly one can show that

$$-\epsilon^{(m-1)} \leqslant \varphi_2^{(m-1)}. \tag{51}$$

Then

$$\begin{aligned}
\varphi_1^{(m+1)} - \varphi_1^{(m)} &= \frac{G_0}{2}[\ln(1+\exp\varphi_2^{(m-1)})/ \\
&\quad \times(1+\exp\varphi_2^{(m-1)})] \\
&\quad - \frac{G_0}{2}[\ln(1+\exp(-\epsilon_2^{(m)}))/ \\
&\quad \times(1+\exp(-\epsilon^{(m-1)}))] > 0.
\end{aligned} \tag{52}$$

This completes the proof of the theorem.

IV. PROOF OF THEOREMS 2 AND 3

Proof of Theorem 2: First we will prove that $\varphi_\nu^{(n)}$ are bounded above. Let us define

$$\bar{\varphi}_\nu^{(n)} = \max \varphi_\nu^{(n)} \tag{53}$$

for the nth iteration. Consider the equations

$$x_1 = \tfrac{1}{2}\ln(1+\exp x_2),$$

$$x_\nu = \tfrac{1}{2}\ln(1+\exp x_{\nu-1})(1+\exp x_{\nu+1}), \quad \nu \geqslant 2, \tag{54}$$

whose solutions x_ν satisfy

$$x_\nu = \ln [\nu(\nu+2)]. \tag{55}$$

Now it is obvious that

$$\bar{\varphi}_\nu^{(0)} < x_\nu. \tag{56}$$

Suppose

$$\bar{\varphi}_\nu^{(n)} < x_\nu \quad \text{for } n \leqslant m. \tag{57}$$

Then

$$\begin{aligned}\bar{\varphi}_\nu^{(m+1)} &\leqslant \tfrac{1}{2} \ln [(1+\exp \bar{\varphi}_{\nu-1}^{(m)}) \\ &\quad \times (1+\exp \bar{\varphi}_{\nu+1}^{(m)})] \\ &< \tfrac{1}{2} \ln [(1+\exp x_{\nu-1}) \\ &\quad \times (1+\exp x_{\nu+1})] \\ &= x_\nu. \end{aligned} \tag{58}$$

Thus by induction, for all values of n

$$\varphi_\nu^{(n)} < x_\nu = \text{const.} \tag{59}$$

and $\varphi_\nu^{(n)}$ is bounded above. To prove $\epsilon^{(n)}$ is bounded below, one notes that from (14),

$$\epsilon^{(n)} > -A + p^2 - \tfrac{1}{2} G_1 \ln (1+\exp(-\epsilon^{(n-1)})) - C, \tag{60}$$

where

$$C = \ln (1+\exp x_1). \tag{61}$$

Consider the equation

$$\begin{aligned} f(p) &= -A + p^2 - \tfrac{1}{2} G_1 \ln [1+\exp(-f)] - C \\ &\equiv F(f). \end{aligned} \tag{62}$$

It has been shown that (62) can be solved by iteration and the solution satisfies[10]

$$f(0) \leqslant f(p) < -A + p^2 - C. \tag{63}$$

Thus $f(p) < \epsilon^{(0)}$. Since F is an increasing functional of f, (60) and (62) imply

$$\begin{aligned} &\epsilon^{(1)} > F(\epsilon^{(0)}) > F(f) = f(p) \\ &\epsilon^{(2)} > F(\epsilon^{(1)}) > F(f) = f(p), \end{aligned} \tag{64}$$

and so on. Therefore

$$\epsilon^{(n)} > f(p) > f(0) \tag{65}$$

and $\epsilon^{(n)}$ is bounded below. This completes the proof of Theorem 2.

Proof of Theorem 3: By Theorems 1 and 2, the limits

$$\lim \epsilon^{(n)}(p) = \epsilon(p), \quad \lim \varphi_\nu^{(n)}(p) = \varphi_\nu(p) \tag{66}$$

exist. It can easily be seen that the sequences approach the limit uniformly and that $\epsilon(p)$, $\varphi_\nu(p)$ satisfy (11a)–(11d). The boundary condition is also satisfied by (66). This completes the proof.

V. CONCLUSION

I have shown in a previous paper that the thermodynamics of the fermions of the delta-function model is determined by the integral equations (11a)–(11d). In this paper, I show that the solutions of (11a)–(11d) exist in the case of negative potential and zero magnetic field. The proof is based on the fact that the iterations $\epsilon^{(n)}$ and $\varphi_\nu^{(n)}$ of (14) are monotonic bounded sequences and thus approach the solutions of (11a)–(11d). For positive chemical potential and nonzero magnetic field, one can take the initial iteration $\varphi_\nu^{(0)} = 2\nu\lambda$ but the sequences $\epsilon^{(n)}$ and $\varphi_\nu^{(n)}$ are no longer monotonic. It is believed that they will still converge to the solutions of (11a)–(11d) though it cannot be proved by the procedure here.

ACKNOWLEDGMENT

The author would like to thank Madeline Chao for helpful discussions.

[1] S. Tomonaga, Prog. Theor. Phys. (Kyoto) **5**, 544 (1950).
[2] E. H. Lieb and W. Liniger, Phys. Rev. **130**, 1605 (1963).
[3] H. A. Bethe, Z. Physik **71**, 205 (1931).
[4] J. McGuire, J. Math. Phys. **6**, 432 (1965).
[5] E. H. Lieb and M. Flicker, Phys. Rev. **161**, 179 (1967).
[6] M. Gandin, Phys. Lett. **24A**, 55 (1967).
[7] C. N. Yang, Phys. Rev. Lett. **19**, 1312 (1967).
[8] J. B. McGuire, J. Math. Phys. **5**, 622 (1964); C. N. Yang, Phys. Rev. **168**, 1920 (1968).
[9] C. K. Lai and C. N. Yang, Phys. Rev. A **3**, 393 (1971).
[10] C. N. Yang and C. P. Yang, J. Math. Phys. **10**, 1115 (1969).
[11] C. K. Lai, Phys. Rev. A **8**, 2567 (1973); C. K. Lai, Phys. Rev. Lett. **26**, 1472 (1971).

The free boson gas in a weak external potential

J. V. Pulè
Department of Mathematical Physics, University College, Belfield, Dublin 4, Ireland[a)]
School of Theoretical Physics, Dublin Institute for Advanced Studies

(Received 22 July 1981; accepted for publication 25 September 1981)

The equation of state and the barometric formula for a free boson gas in a weak external potential are derived for very general potentials.

PACS numbers: 05.30.Jp

1. INTRODUCTION

Van den Berg[1] derives heuristically the equation of state and the barometric formula for a one-dimensional free boson gas in a weak external field of power form. These results are made rigorous by Lewis and Ven den Berg.[2] The aim of this paper is to give simple proofs of these results for a very wide class of potentials.

We distinguish between Bose–Einstein condensation, in which the total condensate is in the ground state so that the ground state is macroscopically occupied, and generalized Bose–Einstein condensation in which the condensate occupies the low-lying energy levels without any of these levels being necessarily macroscopically occupied. A detailed study of generalized Bose–Einstein condensation is now in preparation.[3] It is clear from the paper of Landau and Wilde[4] that the equation of state does not distinguish between the two types of condensation but depends only on the distribution of eigenvalues of the single particle Hamiltonian in the large volume limit. In Sec. 2 we obtain this limit distribution by Dirichlet–Neumann bracketing. The techniques used by Davies[5] can be adapted to give the same results but we think that the method used here is much more direct.

In Sec. 3 we consider the scaled distribution of the boson gas in the thermodynamic limit. We obtain the barometric distribution of the normal fluid and show that on the same scale the condensate is concentrated on the set of absolute minima of the potential.

To avoid unnecessary repetition we use the notation of Ref. 6. We are grateful to Professor J. T. Lewis and Dr. M. Van den Berg for many discussions on this problem and to Professors A. Verbeure and J. Messer for making available their manuscript on the treatment of this problem by correlation inequalities.

2. THE EQUATION OF STATE IN THE THERMODYNAMIC LIMIT

Let Λ^{1} be a bounded open region of $\mathbb{R}^{\nu}$ of unit volume whose boundary $\partial\Lambda^{1}$ is piecewise continuously differentiable.

For each $L>0$ let Λ^{L} be the region $\Lambda^{L}=\{x\in\mathbb{R}^{\nu}: L^{-1}x\in\Lambda^{1}\}$. Take H_L to be the self-adjoint operator on $\mathscr{H}_L = L^{2}(\Lambda^{L})$ determined by $-\frac{1}{2}\Delta + V(x/L)$ and the Dirichlet boundary condition. We assume that V is a nonnegative continuous function defined on $\overline{\Lambda^{1}}$. Let $E_1^L \leqslant E_2^L \leqslant E_3^L \leqslant \cdots$ be the eigenvalues of H_L and $\{\phi_n^L : n = 1,2,3,\ldots\}$ the corresponding eigenvectors.

Our objective in this section is to find the pressure in the thermodynamic limit

$$p_{\bar\rho} = \lim_{L\to\infty} -\frac{1}{\beta L^{\nu}} \sum_{k=1}^{\infty} \ln\left[1 - z(L)e^{-\beta E_k^L}\right], \tag{2.1}$$

where the fugacity $z(L)$ is determined by the constant density constraint

$$\bar\rho = \frac{1}{L^{\nu}} \sum_{k=1}^{\infty} \frac{z(L)}{e^{\beta E_k^L} - z(L)}. \tag{2.2}$$

Equivalently, if we put $\eta_k^L = E_k^L - E_1^L$ we require

$$p_{\bar\rho} = \lim_{L\to\infty} -\frac{1}{\beta L^{\nu}} \sum_{k=1}^{\infty} \ln\left[1 - \zeta(L)e^{-\beta\eta_k^L}\right], \tag{2.3}$$

where $\zeta(L)\in[0,1)$ is the unique solution of

$$\bar\rho = \frac{1}{L^{\nu}} \sum_{k=1}^{\infty} \frac{\zeta(L)}{e^{\beta\eta_k^L} - \zeta(L)}. \tag{2.4}$$

If we define the distribution function F^L on $[0,\infty)$ by

$$F^L(\eta) = \frac{1}{c_\nu L^{\nu}} \max\{k : \eta_k^L \leqslant \eta\},$$

where $c_\nu = \tau_\nu/\pi^{\nu}2^{\nu/2}$, τ_ν being the volume of the unit ball in $\mathbb{R}^{\nu}$, then

$$\frac{1}{L^{\nu}} \sum_{\eta_k^L\in\Lambda} f(\eta_k^L) = c_\nu \int_\Lambda f(\eta) F_L\,(d\eta). \tag{2.5}$$

If in addition f is continuously differentiable on (ϵ, ∞) and $\lim_{\eta\to\infty} f(\eta)F^L(\eta) = 0$, then

$$\frac{1}{L^{\nu}} \sum_{\eta_k>\epsilon} f(\eta_k^L) = -c_\nu \int_\epsilon^\infty f'(\eta)[F^L(\eta) - F^L(\epsilon)]\,d\eta. \tag{2.6}$$

From the work of Landau and Wilde[4] and (2.6) it follows that to find the equation of state it is sufficient to find $\lim_{L\to\infty} F_L(\eta)$. This is done in the following lemma.

Lemma 1: If $E_1^L \to 0$ as $L\to\infty$ the $\lim_{L\to\infty} F^L(\eta) = F(\eta)$, where $F(\eta) = \int_{V(x)<\eta} [\eta - V(x)]^{\nu/2}\, d^{\nu}x$.

Proof: Let $G^L(E) = \max\{k : E_k^L \leqslant E\}$ and $\{B_n^L : n\in I_-\}$ be the set of cubes in $\mathbb{R}^{\nu}$ of the form $[L^{-1/2}a_1, L^{-1/2}(a_1+1))\times\ldots\times[L^{-1/2}a_\nu, L^{-1/2}(a_\nu+1))$, with $a_1, \ldots, a_\nu\in\mathbb{Z}$, which intersect Λ^{1} and let $\{B_n^L : n\in I_+\}$ be the set of

[a)] Postal address.

 0022-2488/83/010138-05$02.50

those cubes that are contained in Λ^1. Also let

$$A_n^L = LB_n^L, \quad {}_+V_n^L = \sup_{x\in A_n^L} V(x/L) = \sup_{x\in B_n^L} V(x)$$

and

$${}_-V_n^L = \begin{cases} \inf_{x\in A_n^L} V(x/L) & \text{if } A_n^L \subseteq \overline{\Lambda}^L, \\ 0 & \text{if } A_n^L \not\subseteq \overline{\Lambda}^L, \end{cases}$$

$$= \begin{cases} \inf_{x\in B_n^L} V(x) & \text{if } B_n^L \subseteq \overline{\Lambda}^1, \\ 0 & \text{if } B_n^L \not\subseteq \overline{\Lambda}^1. \end{cases}$$

Denote by ${}_+V^L$ (respectively ${}_-V^L$) the piecewise constant function on $\cup_{n\in I_+} A_n^L$ (respectively $\cup_{n\in I_-} A_n^L$) with value ${}_+V_n^L$ (respectively ${}_-V_n^L$) in A_n^L and by ${}_+\Delta$ (respectively ${}_-\Delta$) the Laplacian with the Dirichlet (respectively, Neumann) boundary condition on the boundary of each A_n^L $n\in I_+$ (respectively I_-). Let ${}_+G^L(E)$ [respectively ${}_-G^L(E)$] be the distribution of the eigenvalues of $-\frac{1}{2}({}_+\Delta) + {}_+V^L$[respectively, $-\frac{1}{2}({}_-\Delta) + {}_-V^L$] Then by Dirichlet–Neumann bracketing (Ref. 7, Theorem XIII. Eq. 15)

$${}_-G^L(E) \geqslant G^L(E) \geqslant {}_+G^L(E). \tag{2.7}$$

Now

$${}_+G^L(E) = \sum_{\substack{n\in I_+ \\ {}_+V_n^L \leqslant E}} \eta_+[(E - {}_+V_n^L)L], \tag{2.8}$$

where η_+ is the distribution of the eigenvalues of the Laplacian on the unit cube with the Dirichlet boundary condition. By Ref. 7, for η_+ there is a constant c_+ such that

$$\eta_+(E) \geqslant c_\nu E^{\nu/2} - c_+(1 - E^{(\nu-1)/2}/2^{(\nu-1)/2}).$$

Therefore

$${}_+G^L(E) \geqslant c_\nu L^\nu \sum_{\substack{n\in I_+ \\ {}_+V_n^L \leqslant E}} (E - {}_+V_n^L)^{\nu/2} - c_+ \sum_{\substack{n\in I_+ \\ {}_+V_n^L \leqslant E}} \left(1 + L^{(\nu-1)/2} \frac{(E - {}_+V_n^L)^{(\nu-1)/2}}{2^{(\nu-1)/2}}\right). \tag{2.9}$$

Similarly

$${}_-G^L(E) \leqslant c_\nu L^\nu \sum_{\substack{n\in I_- \\ {}_-V_n^L \leqslant E}} (E - {}_-V_n^L)^{\nu/2} + c_- \sum_{\substack{n\in I \\ V_n^L \leqslant E}} \left(1 + L^{(\nu-1)/2} \frac{(E - {}_-V_n^L)^{(\nu-1)/2}}{2^{(\nu-1)/2}}\right). \tag{2.10}$$

From (2.9) we see that

$$\lim_{L\to\infty} \inf \frac{1}{L^\nu} G^L(\eta + E_1^L) \geqslant c_\nu \int_{\eta > V(x)} [\eta - V(x)]^{\nu/2}\, dx,$$

and from (2.10)

$$\lim_{L\to\infty} \sup \frac{1}{L^\nu} G^L(\eta + E_1^L) \leqslant c_\nu \int_{\eta > V(x)} [\eta - V(x)]^{\nu/2}\, dx.$$

Since $F^L(\eta) = (1/c_\nu L^\nu) G^L(\eta + E_1^L)$ the lemma follows: ■

Let H_L^0 be the Dirichlet Laplacian on Λ^L. It is know that if G_0^L is the distribution of eigenvalues of H_L^0, $\lim_{\lambda\to\infty} G_0^1(\lambda)/\lambda^{\nu/2} = c_\nu$ and therefore $G_0^1(\lambda)/\lambda^{\nu/2}$ is bounded by say c. Now $H_L^0 \leqslant H_L$ and so

$$\frac{G^L(\eta)}{L^\nu} \leqslant \frac{G_0^L(\eta)}{L^\nu} = \frac{G_0^1(L^2\eta)}{(L^2\eta)^{\nu/2}} \eta^{\nu/2} < c\eta^{\nu/2}.$$

Thus

$$F_L(\eta) < \frac{c}{c_\nu} (\eta + E_1^L)^{\nu/2}. \tag{2.11}$$

Therefore if f' is continuous and bounded on (ϵ, ∞) and $f'(\eta)\eta^{\nu/2}$ is integrable at infinity, by the Lebesgue dominated convergence theorem we deduce from (2.6) that

$$\lim_{L\to\infty} \frac{1}{L^\nu} \sum_{\eta_k > \epsilon} f(\eta_k^L) = -c_\nu \int_\epsilon^\infty f'(\eta)[F(\eta) - F(\epsilon)]\, d\eta.$$

Using these results the next theorem follows immediately from Ref. 4. Let $\rho_c \leqslant \infty$ be defined by

$$\rho_c = \beta c_\nu \int_0^\infty \frac{e^{\beta\eta}}{(e^{\beta\eta} - 1)^2} F(\eta)\, d\eta = \frac{1}{(2\pi\beta)^{\nu/2}} \int_{\Lambda^1} d^\nu x\, g_{\nu/2}(e^{-\beta V(x)}), \tag{2.12}$$

where $g_r(s) = \sum_{n=1}^\infty (s^n/n^r)$.

Theorem 1: When the mean density $\bar\rho$ is less than ρ_c, the grand canonical pressure, $p_{\bar\rho}$, is determined parametrically as a function of $\bar\rho$ by the pair of equations

$$\bar\rho = \beta c_\nu \int_0^\infty \frac{\bar\xi e^{\beta\eta}}{(e^{\beta\eta} - \bar\xi)^2} F(\eta)\, d\eta = \frac{1}{(2\pi\beta)^{\nu/2}} \int_{\Lambda^1} d^\nu x\, g_{\nu/2}(\bar\xi e^{-\beta V(x)}),$$

$$\beta p_{\bar\rho} = \beta c_\nu \int_0^\infty \frac{\bar\xi}{(e^{\beta\eta} - \bar\xi)} F(\eta)\, d\eta = \frac{1}{(2\pi\beta)^{\nu/2}} \int_{\Lambda^1} d^\nu x\, g_{1+\nu/2}(\bar\xi e^{-\beta V(x)}).$$

If ρ_c is finite and $\bar\rho$ is greater than ρ_c then $p_{\bar\rho}$ is independent of $\bar\rho$ and is given by

$$\beta p_{\bar\rho} = \beta c_\nu \int_0^\infty \frac{1}{(e^{\beta\eta} - 1)} F(\eta)\, d\eta = \frac{1}{(2\pi\beta)^{\nu/2}} \int_{\Lambda^1} d^\nu x\, g_{1+\nu/2}(e^{-\beta V(x)}).$$

Proof: It is clear that if $\bar\rho < \rho_c$, $\zeta(L) \to \bar\xi$, where $\bar\xi$ is the unique solution in [0, 1) of

$$\bar\rho = \beta c_\nu \int_0^\infty \frac{\bar\xi e^{\beta\eta}}{(e^{\beta\eta} - \bar\xi)} F(\eta)\, d\eta,$$

while if $\rho_c < \infty$ and $\bar\rho \geqslant \rho_c$, $\zeta(L) \to 1$. Since

$$\frac{-1}{\beta L^\nu} \sum_{k=1}^\infty \ln(1 - \zeta e^{-\beta\eta_k^L}) = c_\nu \int_0^\infty \frac{\zeta}{(e^{\beta\eta} - \zeta)} F_L(\eta)\, d\eta$$

is uniformly convergent, for $\zeta\in[0,\,1-\delta]$, $\delta>0$, to

$$c_\nu\int_0^\infty \frac{\zeta}{(e^{\beta\eta}-\zeta)}F(\eta)d\eta, \quad \text{for } \bar\rho<\rho_c,$$

the result follows without difficulty.

For $\bar\rho\geqslant\rho_c$ we use the inequality of Theorem 4.2 of Ref. 4, viz. for $\epsilon>0$,

$$0\leqslant -\frac{1}{\beta L^\nu}\sum_{\eta_k<\epsilon}\ln[1-\zeta(L)e^{-\beta\eta_k^L}]$$

$$\leqslant \frac{2\bar\rho}{e\beta\bar\zeta(L)e^{-\beta\epsilon}}F_L(\epsilon),$$

which tends to zero as $\epsilon\downarrow 0$. But

$$-\frac{1}{\beta L^\nu}\sum_{\eta_k\geqslant\epsilon}\ln(1-\zeta e^{-\beta\eta_k^L})$$

converges uniformly for $\zeta\in[0,1]$, which proves the result in the second case. ■

Note that if $\rho_c<\infty$ we have a phase transition in the sense that $p_{\bar\rho}$ has a singularity at $\bar\rho=\rho_c$. As was noted in the Introduction, this result is independent of how $\zeta(L)$ converges to 1 in the case $\bar\rho\geqslant\rho_c$ and therefore is independent of $(1/L^\nu)\zeta(L)/[1-\zeta(L)]$, which is the occupation density of the ground state. The ground state need not be the only energy level in which there is macroscopic occupation; in Ref. 2, for example, there are an infinite number of levels macroscopically occupied. On the other hand, there may not be macroscopic occupation of the ground state or any other level. In the following theorem we give a condition for the ground state, and the ground state only, to be macroscopically occupied.

Theorem 2: Let $\rho_k(\bar\rho)$ be the occupation density for the k th energy level, i.e.,

$$\rho_k(\bar\rho)=\frac{1}{L^\nu}\frac{\zeta(L)}{e^{\beta\eta_k^L}-\zeta(L)}.$$

Suppose that $\nu\geqslant 3$ and $E_2^L/(E_2^L-E_1^L)<\kappa<\infty$. Then, for $\bar\rho\geqslant\rho_c$,

$$\lim_{L\to\infty}\rho_1(\bar\rho)=\bar\rho-\rho_c$$

and

$$\lim_{L\to\infty}\rho_k(\bar\rho)=0, \quad k\neq 1.$$

Proof: To prove this theorem it is clearly sufficient to show that $(1/L^\nu)\Sigma_{k=2}^\infty[\zeta/(e^{\beta\eta_k^L}-\zeta)]$ converges uniformly for $\zeta\in[0,1]$.

$$\frac{1}{L^\nu}\sum_{k=2}^\infty\frac{\zeta}{e^{\beta\eta_k^L}-\zeta}=\beta c_\nu\int_0^\infty\frac{\zeta e^{\beta\eta}}{(e^{\beta\eta}-\zeta)^2}\left(F^L(\eta)-\frac{1}{c_\nu}\right)d\eta$$

and

$$0\leqslant F^L(\eta)-\frac{1}{c^\nu}<F^L(\eta)<c(\eta+E_1^L)^{\nu/2} \text{ by (2.11)}$$

$$<c\eta^{\nu/2}\left(1+\frac{E_1^L}{E_2^L-E_1^L}\right)^{\nu/2}$$

$$<c\eta^{\nu/2}(1+k^{\nu/2}).$$

Since, for $\nu\geqslant 3$, $\eta^{\nu/2-2}$ is integrable at zero this gives the required uniform convergence.

3. THE SCALED DENSITY DISTRIBUTION

In this section we investigate the spatial distribution, scaled in a suitable way, of the boson gas in the thermodynamic limit. For $A\subseteq\Lambda^1$ let $\nu^L(A)$ denote the fraction of the total number of particles which is in LA, i.e.,

$$\nu^L(A)=\frac{1}{\bar\rho L^\nu}\int_{LA}\sum_{k=1}^\infty\frac{\zeta(L)}{[e^{\beta\eta_k^L}-\zeta(L)]}|\phi_k^L(x)|^2\,d^\nu x$$

$$=\frac{1}{\bar\rho}\sum_{k=1}^\infty\frac{\zeta(L)}{e^{\beta\eta_k}-\zeta(L)}\int_A|\phi_k^L(x)|^2\,d^\nu x. \tag{3.1}$$

If we define the distribution function F_A^L on $[0,\infty)$ by

$$F_A^L(\eta)=\frac{1}{c_\nu}\sum_{\eta_k^L\leq\eta}\int_A|\phi_k^L(Lx)|^2\,d^\nu x$$

then

$$\nu^L(A)=\frac{c_\nu}{\bar\rho}\int_0^\infty\frac{\zeta(L)}{e^{\beta\eta}-\zeta(L)}F_A^L(d\eta). \tag{3.2}$$

For technical reason we consider only A's that are open and whose boundaries have zero Lebesgue measure. The result of this lemma coincides with that of Lemma 1 if $A=\Lambda^1$ but the proof of Lemma 1 does not require Wiener integration techniques

Lemma 2: $\lim_{L\to\infty}F_A^L(\eta)=F_A(\eta)$,

where

$$F(\eta)=\int_{\substack{V(x)<\eta\\ x\in A}}[\eta-V(x)]^{\nu/2}\,d^\nu x.$$

Proof: The kernel of the integral operator e^{-tH^L} is $G^L(x,y;t)$, where

$$G^L(x,y;t)=p(x-y;t)\mathbb{E}\left\{\exp\left\{-\int_0^t V(y+x(\tau))\,d\tau\right\};\ y+x(\tau)\in\overline{\Lambda^L},\ 0\leqslant\tau\leqslant t\,|x(0)=0,x(t)=x-y\right\}, \tag{3.3}$$

where

$$p(x,t)=(2\pi t)^{-\nu/2}\exp\{-\|x\|^2/2t\} \tag{3.4}$$

and $\mathbb{E}\{\cdot\}$ denotes the average value for all paths $x(\cdot)$ of a Wiener process on $\mathbb{R}^\nu$[8]

$$c_\nu\int_0^\infty e^{-t\eta}F_A^L(d\eta)=\frac{e^{tE_1^L}}{L^\nu}\int_{LA}G^L(x,x;t)\,d^\nu x. \tag{3.5}$$

Following Ray[8] and using (3.3) and Jensen's inequality, we have

$$\frac{1}{L^\nu}\int_{LA} G^L(x,x;t)\,d^\nu x \leqslant \frac{1}{L^\nu}\int_{LA} d^\nu x \frac{1}{(2\pi t)^{\nu/2}}\frac{1}{t}\int_0^t d\tau\ \mathbb{E}\{\exp\{-tV(x/L+x(\tau)/L)\};$$

$$x + x(\tau)\in \bar{\Lambda}^L,\ 0\leqslant\tau\leqslant t\,|x(0)=x(t)=0\}$$

$$=\frac{1}{(2\pi t)^{\nu/2}}\frac{1}{t}\int_0^t dt\,\mathbb{E}\left\{\frac{1}{L^\nu}\int_{x\in LA\cap\bar{\Lambda}^L - x(\tau)} d^\nu x \exp\left[-tV\left(\frac{x}{L}+\frac{x(\tau)}{L}\right)\right]|x(0)=x(t)=0\right\}$$

$$=\frac{1}{(2\pi t)^{\nu/2}}\frac{1}{t}\int_0^t dt\,\mathbb{E}\left\{\int_{x\in\bar{A}^{-1}\cap A + x(\tau)/L} d^\nu x \exp[-tV(x)]\,|x(0)=x(t)=0\right\}.$$

The last expression tends to

$$\frac{1}{(2\pi t)^{\nu/2}}\frac{1}{t}\int_0^t dt\,\mathbb{E}\left\{\int_{x\in A} d^\nu x \exp[-tV(x)]\,|x(0)=x(t)=0\right\} = \frac{1}{(2\pi t)^{\nu/2}}\int_A d^\nu x e^{-tV(x)} = c_\nu\int_0^\infty e^{-t\eta}F_A(d\eta).$$

Using the notation of Lemma 1 let $\{B_n^L: n\in\mathrm{I}\}$ be the set of B_n^L contained in A. Then

$$\frac{1}{L^\nu}\int_{LA} G^L(x,x;t)\,d^\nu x = \frac{1}{L^\nu(2\pi t)^{\nu/2}}\int_{LA} d^\nu x\,\mathbb{E}\left\{\exp\left[-\int_0^t V\left(\frac{x(\tau)}{L}\right)d\tau\right], x(\tau)\in\Lambda^L\,|x(0)=x(t)=x\right\}$$

$$\geqslant \frac{1}{L^\nu(2\pi t)^{\nu/2}}\sum_{n\in I}\int_{A_n^L} d^\nu x\,\mathbb{E}\left\{\exp\left[-\int_0^t V\left(\frac{x(\tau)}{L}\right)d\tau\right], x(\tau)\in A_n^L\,|x(0)=x(t)=x\right\}$$

$$\geqslant \frac{1}{L^\nu(2\pi t)^{\nu/2}}\sum_{n\in I} e^{-t_+V_n^L}\int_{A_n^L} d^\nu x\,\mathbb{E}\{1, x(\tau)\in A_n^L\,|x(0)=x(t)=x\}.$$

There exists a constant C such that

$$\frac{1}{L^\nu(2\pi t)^{\nu/2}}\int_{A_n^L} d^\nu x\,\mathbb{E}\{1, x(\tau)\in A_n^L\,|x(0)=x(t)=x\} > \frac{1}{L^{\nu/2}}\frac{1}{(2\pi t)^{\nu/2}}\left(1-C\frac{t^{1/4}}{L^{1/8}}\right).$$

This can be extracted from Arima,[9] Mizohata and Arima,[10] Van den Berg,[2] and various other sources. Therefore

$$\lim_{L\to\infty}\frac{1}{L^\nu}\int_{LA}\frac{G}{L}(x,x;t)\,d^\nu x \geqslant \frac{1}{(2\pi t)^{\nu/2}}\int_A d^\nu x e^{-tV(x)} = c_\nu\int_0^\infty e^{-t\eta}F_A(d\eta).$$

Hence

$$\int_0^\infty e^{-t\eta}F_A^L(d\eta)\to\int_0^\infty e^{-t\eta}F_A(d\eta),$$

which means (see, e.g., Ref. 11) $F_A^L(\eta)\to F_A(\eta)$. ■

Lemma 3: Let $A_0=\{x\in\Lambda^1: V(x)=0\}$ and suppose $\bar{A}\cap A_0=\emptyset$; then there exists $\epsilon_0>0$ and $c>0$ such that for $E_k^L<\epsilon_0$, $x\in A$, $L^{\nu/2}|\phi_k^L(Lx)|<c(E_k^L)^{\nu/4}$.

Proof: Since $\bar{A}\cap A_0=\emptyset$ we can find a $\delta>0$, $\sigma>0$ such that $V(x+y)>\delta$ if $|y|<\sigma$ and $x\in A$. Let $\epsilon_0<\delta/2$ and suppose $E_k<\epsilon$. Using Eq. 8 of Ref. 8 we have

$$e^{-tE_k^L}|\phi_k^L(Lx)|$$

$$\leqslant\mathbb{E}\left\{\exp\left[-\int_0^t V\left(x+\frac{x(\tau)}{L}\right)\right]|\phi_k^L(Lx+x(t))|,\right.$$

$$\left. Lx+x(\tau)\in\Lambda^L\right\}$$

$$\leqslant\left(\frac{eE_k^L}{\pi\nu}\right)^{\nu/4}\mathbb{E}\left\{\exp\left[-\int_0^t V\left(x+\frac{x(\tau)}{L}\right)\right],\right.$$

$$\left. Lx+x(\tau)\in\Lambda^L\right\},$$

using the bound for $\|\phi_k^L\|_\infty$ in Ref. 12,

$$\leqslant\left(\frac{eE_k^L}{\pi\nu}\right)^{\nu/4}[e^{-\delta t}\,\mathrm{Prob}\{\max\|x(\tau)\|<L\sigma\}$$

$$+\mathrm{Prob}\{\max\|x(\tau)\|>L\sigma\}]$$

$$\leqslant\left(\frac{eE_k^L}{\pi\nu}\right)^{\nu/4}\left[e^{-\delta t}+3\int_{\|y\|>L\sigma/4}p(y,t)\,d^\nu y\right],$$

using Lemma 2 of Ref. 8. Therefore

$$|\phi_k^L(Lx)|\leqslant\left(\frac{eE_k^L}{\pi\nu}\right)^{\nu/4}(e^{-(\delta-t_0)t}+3(2^{\nu/2})e^{-(L^2\sigma^2/64t^2-\epsilon_0)t})$$

By choosing $t^2=L^2\sigma^2/64\delta$ we obtain

$$|\phi_k^L(Lx)|\leqslant\left(\frac{eE_k^L}{\pi\nu}\right)^{\nu/4}[1+3(2^{\nu/2})]e^{-L\sigma/16\delta^{1/2}}$$

$$\leqslant c(E_k^L)^{\nu/4}/L^{\nu/2}.$$

Theorem 3: If $\bar{\rho}<\rho_c$ then

$$\lim_{L\to\infty}\nu^L(A)=\frac{\beta c_\nu}{\bar{\rho}}\int_0^\infty\frac{\bar{\xi}e^{\beta\eta}}{(e^{\beta\eta}-\bar{\xi})^2}F_A(\eta)\,d\eta,$$

where $\bar{\xi}$ is as in Theorem 1.

If $\bar{\rho}\geqslant\rho_c$, then

$$\lim_{L\to\infty}\nu^L(A)=\frac{\beta c_\nu}{\bar{\rho}}\int_0^\infty\frac{e^{\beta\eta}}{(e^{\beta\eta}-1)^2}F_A(\eta)\,d\eta$$

if $\bar{A}\cap A_0=\emptyset$,

$$\lim_{L\to\infty}\nu^L(A)=\left(1-\frac{\rho_c}{\rho}\right)+\frac{\beta c_\nu}{\bar{\rho}}\int_0^\infty\frac{e^{\beta\eta}}{(e^{\beta\eta}-1)^2}F_A(\eta)\,d\eta$$

if $A_0\subseteq A$.

Proof: If $\bar{\rho} < \rho_c$ the result follows immediately from Lemma 2. Suppose $\bar{\rho} > \rho_c$ and $\bar{A} \cap A_0 = \emptyset$ and let

$$\nu_\epsilon^L(A) = \frac{1}{\rho} \sum_{\eta_k^L > \epsilon} \frac{\zeta(L)}{e^{\beta\eta_k^L} - \zeta(L)} \int_A |\phi_k^L(Lx)|^2 \, d^\nu x.$$

Then again by Lemma 2 $\nu_\epsilon^L(A)$ tends to

$$\frac{1}{\bar{\rho}} \frac{\beta\tau_\nu}{\pi^\nu 2^{\nu/2}} \int_\epsilon^\infty \frac{e^{\beta\eta}}{(e^{\beta\eta} - 1)^2} F_A(\eta) \, d\eta.$$

However,

$$\nu^L(A) - \nu_\epsilon^L(A)$$

$$= \frac{1}{\bar{\rho}} \sum_{\eta_k^L < \epsilon} \frac{\zeta(L)}{e^{\beta\eta_k^L} - \zeta(L)} \int_A |\phi_k^L(Lx)|^2 \, d^\nu x$$

$$< c\epsilon^{\nu/4}$$

if ϵ is small enough and the required result follows.

If $A_0 \subset A$,

$$\nu^L(A) = \nu^L(\Lambda^1) - \nu^L(\Lambda^1 - A) = 1 - \nu^L(\Lambda^1 - A)$$

$$\to 1 - \frac{1}{\bar{\rho}} \frac{\beta\tau_\nu}{\pi^\nu 2^{\nu/2}} \int_0^\infty \frac{e^{\beta\eta}}{(e^{\beta\eta} - 1)^2} [F(\eta) - F_A(\eta)] \, d\eta$$

$$= 1 - \frac{\rho_c}{\bar{\rho}} + \frac{\beta c_\nu}{\bar{\rho}} \int_0^\infty \frac{e^{\beta\eta}}{(e^{\beta\eta} - 1)} F_A(\eta) \, d\eta.$$

Finally, under the conditions of Theorem 2 we prove that the condensate density is concentrated at the zeros of the potential

Corollary 1: Suppose $\nu \geqslant 3$ and $E_2^L/(E_2^L - E_1^L) < \kappa < \infty$ and let

$$\nu_k^L(A) = \frac{1}{\bar{\rho}} \frac{\zeta(L)}{e^{\beta\eta_k^L} - \zeta(L)} \int_A |\phi_k^L(Lx)|^2 \, d^\nu x;$$

then for $\bar{\rho} \geqslant \rho_c$,

$$\lim_{L\to\infty} \sum_{k=2}^\infty \nu_k^L(A) = \frac{\beta c_\nu}{\bar{\rho}} \int_0^\infty \frac{e^{\beta\eta}}{(e^{\beta\eta} - 1)^2} F_A(\eta) \, d\eta$$

and

$$\lim_{L\to\infty} \nu_1^L(A) = 0 \qquad \text{if } \bar{A} \cap A_0 = \emptyset,$$

$$= \left(1 - \frac{\rho_c}{\bar{\rho}}\right) \qquad \text{if } A_0 \subseteq A.$$

Proof: To prove this corollary it is clearly sufficient to show that

$$\int_0^\infty \frac{\zeta e^{\beta\eta}}{(e^{\beta\eta} - \zeta)^2} \left(F_A^L(\eta) - \frac{1}{c_\nu} \int_A |\phi_1^L(Lx)|^2 \, d^\nu x \right) d\eta$$

converges uniformly for $\zeta \in [0,1]$. Since $F_A^L(\eta) \leqslant F^L(\eta)$ this follows as in Theorem 2.

[1] M. Van den Berg, Phys. Lett. **78A**, 88 (1980).
[2] M. Van den Berg and J. T. Lewis, Commun. Math. Phys. **81**, 475 (1981).
[3] J. T. Lewis, J. V. Pulè, and M. Van den Berg, "Generalized Bose-Einstein condensation in non-interacting systems and the spectrum of the single-particle Hamiltonian" (to appear).
[4] L. J. Landau and I. F. Wilde, Commun. Math. Phys. **70**, 43 (1979).
[5] E. B. Davies, Commun. Math. Phys. **30**, 229 (1973).
[6] J. T. Lewis and J. V. Pulè, Commun. Math. Phys. **36**, 1 (1974).
[7] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Analysis of Operators* (Academic, New York, 1978), Vol. IV.
[8] D. Ray, Trans. Am. Math. Soc. **77**, 299 (1954).
[9] R. Arima, J. Math. Kyoto Univ. **4**, 207 (1964).
[10] S. Mizohata and R. Arima, J. Math. Kyoto Univ. **4**, 245 (1964).
[11] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1966), Vol. II.
[12] E. B. Davies, J. London Math. Soc. **7**, 483 (1973).

Renormalization group hypothesis for critical phenomena theory[a)]

George A. Baker, Jr.
Theoretical Division, Los Alamos National Laboratory, University of California, Los Alamos, New Mexico 87545

(Received 10 March 1981; accepted for publication 30 July 1982)

We give various "nonperturbative" results for strong coupling, ultraviolet cut-off removed limits of the bare mass in $g_0{:}\phi^4{:}_d$, lattice cut-off, boson field theory. We find that the renormalization-group, unique, strong-coupling, zero-lattice-spacing, double-limit hypothesis has some remarkable consequences, which seem difficult to reconcile with other available information.

PACS numbers: 11.10.Gh, 03.70. + k

The results[1] of the renormalization group approach[2] to the theory of critical phenomena have been in fairly close agreement with the predictions of the more traditional methods.[3] A close inspection,[4] however, shows that there may be small but persistent differences. These differences are most apparent[5] in a particular family of relations between the exponents which describe the rate of divergence (or vanishing) of the various physical quantities at the critical point of the system considered, e.g., a ferromagnet. The explicit appearance of the spatial dimension in such an exponent relation is the signature of this group of relations, and the family is called the hyperscaling relations.

Baker and Kincaid[6] have emphasized that the renormalization group theory of critical phenomena is based on a double limit hypothesis. The purpose of this paper is to begin an investigation of this hypothesis. To this end we will examine the behavior of the "bare mass" in various strong bare coupling limits and obtain several results. We will work in the context of a Euclidean, lattice cut-off, $g_0{:}\phi^4{:}_d$ boson field theory, which is equivalent to a continuous-spin Ising model. We use d to denote the spatial dimension.

In this paper we compute (based on a monotonicity property of the critical temperature) bounds on various strong-coupling continuum limits for the bare mass, and on the basis of the renormalization group hypothesis derive a remarkable formula for the amplitude of the correlation length. This formula implies that there is a particular value of the parameters at which there is singularity in this amplitude.

Although extension to other lattices is no problem, we will consider explicitly the hyper-simple-cubic family of lattices. The partition function is

$$Z(H) = M^{-1}\int_{-\infty}^{+\infty}\cdots\int\prod_{i=0}^{N-1}d\phi_i \exp\left[-\frac{1}{2}a^d\sum_{i=0}^{N-1}\left\{\sum_{\{\delta\}}\frac{(\phi_i-\phi_{i+\delta})^2}{a^2}+m_0^2{:}\phi_i^2{:}+\frac{2}{4!}g_0{:}\phi_i^4{:}\right\}+\sum_i H_i\phi_i\right], \tag{1}$$

where N is the number of lattice sites, a is the lattice spacing, $\{\delta\}$ is the set of unit vectors parallel to each of the d lattice directions, the : : imply the usual, field-theoretic, normal-ordered-product for fields of mass m_0 and H_i is the magnetic field at site i. The formal constant M is meant to impose the condition $Z(0) = 1$. This partition function is written in such a way as to be a lattice cut-off field-theory. The renormalization group hypothesis for this model can now be stated as *the double limit $g_0\to\infty$, $a\to 0$ of the field theory formalism exists and is independent of the order of approach.* More specifically, we are referring to what appears to be the calculationally most advanced version of the renormalization group; that is the Callan–Symanzik formalism expounded for example, by Brezin *et al.*[2] Examples of the unique double limit hypothesized are g^*, $\eta(g^*)$, $\eta_2(g^*)$, and $W(g^*)$ (also called $\beta(g^*)$ by many authors). Here in the calculations so far reported by them, some version of this hypothesis and in addition, a certain amount of smoothness and differentiability near this point, is required for the various quantities which connect the "bare" parameters g_0 and the renormalization constants with the renormalized ones. To complete the theory, the hypothesis will also be needed for various N point vertex functions, but serious calculation of these more elaborate quantities is currently more often a question of principle than practice. We will briefly discuss below one example and show that it relates to the bare mass. Although the bare mass is not "universal" its examination does begin the study of the key hypothesis and may perhaps suggest directions for further investigation.

If we perform the usual amplitude (Z_3) and mass renormalizations[6] ($m_0^2 = m^2 + \delta m^2$) then we can rewrite (1) as

$$Z(\tilde{H}) = \tilde{M}^{-1}\int_{+\infty}^{-\infty}\cdots\int\prod_{i=1}^{N}d\sigma_i\exp\left[\sum_i K\sum_{\{\sigma\}}\sigma_i\sigma_{i+\delta} - \tilde{g}_0\sigma_i^4 - \tilde{A}\sigma_i^2 + \tilde{H}\sigma_i\right], \tag{2}$$

where $\tilde{M}$ is a new normalization constant, and the relation between the field theory language of (1) and the statistical mechanical language of (2) is[6]

$$\tilde{g}_0 = g_0K^2a^{4-d}/24,$$
$$\tilde{A} = \tfrac{1}{2}K(2d + m^2a^2 + \delta m^2a^2 - \tfrac{1}{2}Ca^2g_0),$$
$$\tilde{H} = H_i[Ka^{2-d}]^{1/2}. \tag{3}$$

The constant C is the commutator $[\phi^-, \phi^+]$ which arises from the reduction of the normal ordered products $:\phi^p:$, and in the limit of an infinite system is given by[7]

a) Work performed under the auspices of the U.S. DOE.

 0022-2488/83/010143-05$02.50

$$C = \frac{1}{(\pi)^d} \int_{-\pi/2a}^{\pi/2a} \cdots \int \frac{d\,\mathbf{k}}{m_0^2 + 4a^{-2}\Sigma_{\{\delta\}} \sin^2(\mathbf{k}\cdot\delta a)}. \quad (4)$$

The normal ordering is performed with respect to the bare field, so m_0^2 appears in (4). We require $m_0^2 \geqslant 0$ to keep the integral in (4) from being singular; however, as we will see later, this restriction does not restrict $\tilde{A}$ as it can still range from $-\infty$ to $+\infty$, $d \geqslant 2$, and, of course, we will find no problem for $d < 2$. We have introduced a free parameter, K. We use this extra degree of freedom to impose the condition

$$\langle \sigma^2 \rangle_{H=K=0} = 1 = \frac{\int_{-\infty}^{+\infty} dx\, x^2 \exp(-\tilde{g}_0 x^4 - \tilde{A}x^2)}{\int_{-\infty}^{+\infty} dx \exp(-\tilde{g}_0 x^4 - \tilde{A}x^2)}, \quad (5)$$

which determines $\tilde{A}$ as a function of $\tilde{g}_0$. As long as $0 < \langle \phi^2 \rangle < \infty$ for the cut-off field theory we may always impose (5) by an amplitude renormalization. For large and small values of $\tilde{g}_0$, $\tilde{A}(\tilde{g}_0)$ has the expansions

$$\begin{aligned} \tilde{A}(\tilde{g}_0) &= \tfrac{1}{2} - 6\tilde{g}_0 + 48\tilde{g}_0^2 + O(\tilde{g}_0^3), \quad \tilde{g}_0 \ll 1, \\ &= -2\tilde{g}_0 - \tfrac{1}{2} - \tfrac{1}{4}\tilde{g}_0^{-1} - \tfrac{7}{16}\tilde{g}_0^{-2} + O(\tilde{g}^{-3}), \quad \tilde{g}_0 \gg 1. \end{aligned} \quad (6)$$

It is convenient to employ the following statistical mechanical notation (j lies on a lattice of unit spacing)

$$\begin{aligned} \chi &= \sum_{j=0}^{N-1} \langle \sigma_0 \sigma_j \rangle, \\ \xi^2 &= \sum_{j=0}^{N-1} j^2 \langle \sigma_0 \sigma_j \rangle / (2d\chi), \end{aligned} \quad (7)$$

where χ is the magnetic susceptibility and ξ is the dimensionless correlation length (number of lattice spacings), second-moment definition. Then the usual field-theoretic, renormalization condition,

$$\begin{aligned} &\Gamma_R^{(2)}(p, -p) \\ &\equiv \left\{ a^d \sum_{j=0}^{N-1} \frac{\partial^2 \ln Z(H)}{\partial H_0 \partial H_j}\bigg|_{H=0} \exp(-i\mathbf{p}\cdot\mathbf{j}a) \right\}^{-1} Z_3 \\ &\simeq m^2 + p^2 + \cdots, \quad \text{as } p \to 0, \end{aligned} \quad (8)$$

leads to the relations[6]

$$\begin{aligned} m^2 \xi^2 a^2 &= 1, \\ Z_3 &= K(\chi/\xi^2). \end{aligned} \quad (9)$$

As partial motivation for studying the bare mass, we note that one of the important critical indices can be derived directly from the bare mass

$$m_0^2 = a^{-2}\left[2\frac{\tilde{A}(\tilde{g}_0)}{K} - 2d + \tfrac{1}{2}Ca^2 g_0 \right], \quad (10)$$

which follows trivially from Eq. (3). The essential $g_0{:}\phi^4$: nature of the theory is now built explicitly into the functions $\tilde{A}(\tilde{g}_0)$ and $K(a, \tilde{g}_0)$. K actually depends on ma instead of a, but we will not write the m for simplicity of presentation, as m will usually be taken to be unity in what follows. If we follow Brezin *et al.*[2] and define

$$R_0(\tilde{g}_0, a) \equiv a^2 r_0 \equiv m_0^2 a^2 - \tfrac{1}{2}Ca^2 g_0 = 2\tilde{A}(\tilde{g}_0)/K - 2d, \quad (11)$$

by (10). Then it follows easily that if

$$\xi \propto (K_c - K)^{-\nu}, \quad K \to K_c^-, \quad (12)$$

then for

$$\begin{aligned} Q = Z_{(2)}/Z_3 &= a^{-2}[R_0(\tilde{g}_0, a) - R_0(\tilde{g}_0, 0)] \\ &= 2\tilde{A}(\tilde{g}_0)(K^{-1} - K_c(\tilde{g}_0)^{-1})a^{-2}, \end{aligned} \quad (13)$$

we have the result that ($\tilde{A}(\tilde{g}_0) \neq 0$),

$$\lim_{a \to 0} a \frac{\partial(\ln Q)}{\partial a}\bigg|_{\tilde{g}_0} = -2 + 1/\nu. \quad (14)$$

The renormalization constant $Z_{(2)}$ is that for ϕ^2 insertions (Brezin *et al.*[2]).

In Fig. 1 we show the various strong-coupling limits with which we will be concerned. Remember, in field theory, the natural mode of calculation is [(f) in Fig. 2] $g_0 m^{d-4}$ varying along the line $\xi^{-1} = a = 0$ ($\hat{\xi}_1 = 1$), and the goal of the field theoretic approach is to reduce any calculation to one of these types of calculation. In the mode of presentation in Fig. 1, this line is contracted to the upper left-hand corner and the field theory point $g_0 = \infty$, $a = 0$ is the whole top border. We have replotted in Fig. 2 the same picture in natural field theory variables. The abscissas are implicitly related by Eq. (3). In the statistical mechanics of critical phenomena, the natural mode of approach is to reach the line $\xi^{-1} = 0$ along a vertical path in Fig. 1, e.g., paths (a) or (b) in Figs. 1 and 2.

Now it is usual in the framework of the renormalization group approach[2] basically to project path (b) (a typical statistical mechanical one) on the top border in Fig. 2 in such a way as to be able to use the cut-off removed field theory. To clarify the relation of the approach to Eq. (14), we use the chain rule of partial differentiation to "turn" the direction of the derivative in (14). This procedure is claimed to be plausible in the renormalization group approach because it is correct order-by-order ($d < 4$) in perturbation theory, and so holds for g_0 sufficiently small, i.e., for path (e) of Figs. 1 and 2. Thus, writing by use of Eq. (3),

$$Q(g_0, a) = Q(24\tilde{g}_0 a^{d-4} K^{-2}, a), \quad (15)$$

we have

$$\begin{aligned} a\frac{\partial Q}{\partial a}\bigg|_{\tilde{g}_0} &= 24(d-4)\tilde{g}_0 a^{d-4} K^{-2}\left(\frac{\partial Q}{\partial g_0}\right)_a + a\left(\frac{\partial Q}{\partial a}\right)_{g_0} \\ &= (d-4)g_0\left(\frac{\partial Q}{\partial g_0}\right)_a \\ &\quad + a\left[\left(\frac{\partial Q}{\partial a}\right)_{g_0} - 2K^{-1}\frac{\partial K}{\partial a}\bigg|_{\tilde{g}_0} g_0\left(\frac{\partial Q}{\partial g_0}\right)_a\right]. \end{aligned} \quad (16)$$

Then, combining (13), (14), and (16) with the assumed, con-

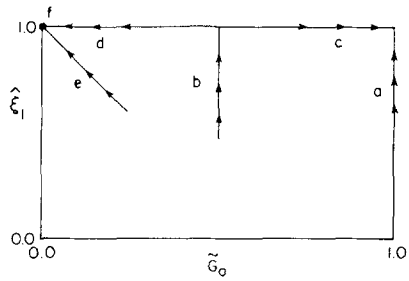


FIG. 1. Various different strong coupling, ultraviolet cut-off removed limits. Here $\hat{\xi}_1 = \xi^2/(1+\xi^2)$, $\tilde{G}_0 = \tilde{g}_0/(1+\tilde{g}_0)$. (a) is the Ising limit, $\tilde{g}_0 = g_0 = \infty$, $a \to 0$. (b) is a typical statistical mechanical problem, $\tilde{g}_0$ fixed, $a \to 0$. (c) is the case $a = 0$, $\tilde{g}_0 \to \infty$. (d) is the case $a = 0$, $\tilde{g}_0 \to 0$. (e) is a typical field theory problem $0 < g_0 < \infty$ fixed, $a \to 0$. It is illustrated here for $d = 2$, and g_0 determines the slope of the line. (f) $a = 0$, $g_0 \to \infty$ is completely contained in the upper left hand corner.

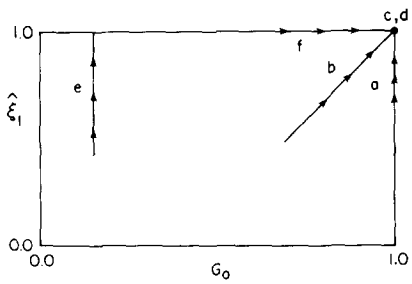


Fig 2

FIG. 2. Various different strong coupling, ultraviolet cut-off removed limits. Here $\hat{\xi}_1 = \xi^2/(1+\xi^2)$, $G_0 = g_0 m^{d-4}/(1+g_0 m^{d-4})$. (a) is the Ising limit $\tilde{g}_0 = g_0 = \infty$, $a \to 0$. (b) is a typical statistical mechanical problem $\tilde{g}_0$ fixed, $a \to 0$, illustrated here for $d = 2$. (c) is the case $a = 0$, $\tilde{g}_0 \to \infty$ and (d) is the case $a = 0, \tilde{g}_0 \to 0$. Both cases are contained in the upper right-hand corner of the figure. (e) is a typical field theory problem $0 < g_0 < \infty$ fixed and $a \to 0$. (f) is the usual field theoretic strong coupling limit $a = 0$, $g_0 \to \infty$.

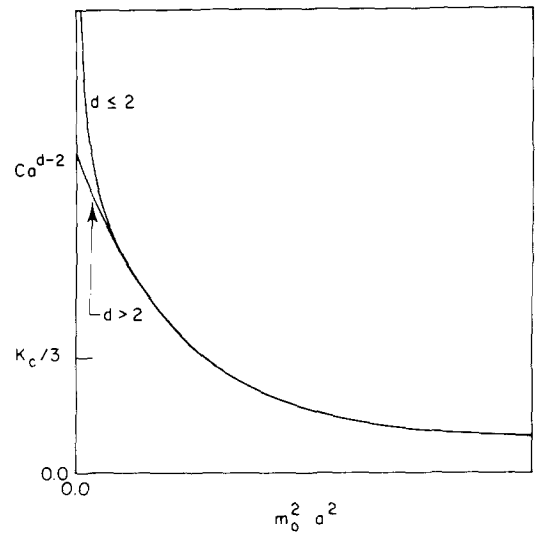


FIG. 3. Sketch of the behavior of $a^{d-2}C(m_0^2a^2)$ for the two cases $d \le 2$ where $C(0) = \infty$ and $d > 2$ where $C(0)$ is finite but always bigger than $K_c/3$, where K_c is the Ising model critical point.

tinuous differentiability at $a = 0, g_0 = \infty$, we have the renormalization group result

$$\lim_{g_0 \to \infty} (d-4) g_0 \frac{\partial \ln Q(g_0,0)}{\partial g_0}\bigg|_m = -2 + 1/\nu. \tag{17}$$

In other words, in Fig. 2, paths (f) and (b) will yield the same result as $(\partial Q/\partial a)_{g_0}$ is assumed finite in and at the upper right-hand corner of Fig. 2, so that the limit as $a \to 0$ of $a(\partial Q/\partial a)_{g_0}$ vanishes. Likewise the last term in (16) vanishes because $(\partial \ln K/\partial \ln a)$ vanishes as $a \to 0$.

Thus, under that hypothesis we get the same result for the field theory approach (f) and the statistical mechanical one (b) and even (a) which is assumed to be a uniform limit of paths of type (b).

We begin by exploring various strong-coupling limits of m_0^2. First the spin-$\frac{1}{2}$ Ising limit,[8] with $g_0 \to \infty$ and a fixed can be given. Here $\tilde{g}_0 \to \infty$; thus combining the expansion of $A(\tilde{g}_0)$, Eq (6), with Eq. (10), we have (keeping a fixed)

$$\lim_{g_0 \to \infty} \left[m_0^2 a^2 + K^{-1} + 2d + \tfrac{1}{2} g_0 a^{4-d} (\tfrac{1}{3} K - C a^{d-2}) \right] = 0, \tag{18}$$

m_0^2, K, and C are all functions of both a and g_0. Inspection of (18) shows that the coefficient of g_0 must be negative of zero. Physically, we must have $m_0^2 \ge 0$ so C can be defined by (4) in a nonsingular manner.

We need to consider in detail the behavior of C as a function of m_0^2 and dimension. Figure 3 shows a sketch of its behavior. Following the methods of Montroll and Weiss,[9] if we use the identities

$$\sin^2\theta = \tfrac{1}{2}(1 - \cos 2\theta), \quad I_0(z) = \frac{1}{\pi}\int_0^\pi e^{z\cos\theta}\, d\theta, \tag{19}$$

$$b^{-1} = \int_0^\infty e^{-by} dy,$$

then we can write

$$C = \tfrac{1}{2} a^{2-d} \int_0^\infty dy\, e^{-\frac{1}{2} m_0^2 a^2 y} [e^{-y} I_0(y)]^d. \tag{20}$$

Now as

$$I_0(z) \simeq 1 + \frac{\frac{1}{4}z^2}{(1!)^2} + \cdots \quad z \ll 1,$$

$$e_0^{-z} I_0(z) \simeq (2\pi z)^{-1/2}\left(1 + \frac{1}{8z} + \cdots\right) \quad z \gg 1, \tag{21}$$

we can give the estimates

$$C \simeq \frac{a^{2-d}}{m_0^2 a^2}, \qquad m_0^2 a^2 \to \infty, \tag{22}$$

$$\simeq a^{2-d} c(d), \qquad d > 2,\ m_0^2 a^2 \to 0,$$

$$\simeq \frac{a^{2-d}(m_0 a)^{d-2}}{4(\pi)^{d/2}} \Gamma(1 - \tfrac{1}{2}d), \qquad d < 2,\ m_0^2 a^2 \to 0, \tag{23}$$

$$\simeq \frac{-\ln(m_0 a)}{2\pi}, \qquad d = 2,\ m_0^2 a^2 \to 0,$$

where

$$c(d) = \frac{1}{2}\int_0^\infty (e^{-y} I_0(y))^d\, dy, \tag{24}$$

is a convergent integral for $d > 2$.

With these asymptotic results for C, we are in a position to analyze the limit (18). As $g_0 \to \infty$ for fixed $a > 0$, we expect and can prove over a limited region that $K(a, g_0) \to K(a, \tilde{g}_0 = g_0 = \infty)$, the Ising model case. Since[7] $3\,c(d) > K_c > K(a > 0)$ for $d = 2,3,\ldots$, we see from Fig. 3, [or more formally from Eqs. (22)–(24) and continuity] that there is a finite, nonzero value of $m_0^2 a^2$ which solves

$$C a^{d-2} - \tfrac{1}{3} K(a, \infty) = 0. \tag{25}$$

If we perturb this solution by an amount of order $(1/g_0 a^{4-d})$ then we can make the [ ] in Eq. (18) vanish. Thus the limit of $m_0^2 a^2$ as $g_0 \to \infty$ is given by the solution of Eq. (25). The limiting value, $\lim_{a \to 0} \lim_{g_0 \to \infty} m_0^2 a^2$, is then the solution of

$$\lim_{a \to 0} a^{d-2} C = \frac{1}{2}\int_0^\infty dy\, e^{-\frac{1}{2} m_0^2 a^2 y} (e^{-y} I_0(y))^d = \tfrac{1}{3} K_c, \tag{26}$$

where K_c is the critical value for the Ising model and has been determined numerically[10] to be 0.440 68···(exact),

0.221 71, 0.149 88, 0.114 03, and 0.09236, for $d = 2$–6. This result, Eqs. (25) and (26), gives the behavior of the bare mass as a function of lattice spacing in the Ising model or strong-coupling limit.[8]

The limit in the field-theory direction [path (e), Figs. 1 and 2] $a \to 0$, with g_0 fixed gives,

$$\lim_{a\to 0} m_0^2 = \lim_{a\to 0}\left(\frac{2\tilde{A}(\tilde{g}_0) - 2dK}{Ka^2}\right) + \frac{1}{2}g_0 \lim_{a\to 0} C. \tag{27}$$

Necessarily $\tilde{g}_0 \to 0$ in this limit and upon expansion, this formula must reduce to the usual, lattice-cut-off, bare coupling constant expansion for the bare mass, to which we do not make any further contribution. In fact, one could use the usual field theory perturbation expansions to provide an expansion of the $K(a = 0, \tilde{g}_0)$ in powers of g_0 by means of Eq. (27).

We can, however, compute some further $g_0 \to \infty$, $a \to 0$ limits. They are characterized by $\tilde{g}_0 = \text{const}$, $a \to 0$ [path (b) in Figs. 1 and 2]. If we rewrite Eq. (10) as

$$m_0^2 a^2 = (12/K^2)\tilde{g}_0(Ca^{d-2}) - 2d + 2\tilde{A}(\tilde{g}_0)/K, \tag{28}$$

then we may deduce that $\lim_{a\to 0} m_0^2 a^2$ is finite for $0 < \tilde{g}_0 < \infty$. If $\tilde{g}_0 \to \infty$, [path (c), Fig. 1], then, as in Eq. (18) we find that the coefficients of $\tilde{g}_0$ must cancel so we get

$$\lim_{g_0\to\infty}\left\{\left[\frac{1}{2}\int_0^\infty dy\, e^{-(1/2)m_0^2a^2y}(e^{-y}I_0(y))^d\right] - \frac{1}{3}K(\tilde{g}_0, a = 0)\right\} = 0 \tag{29}$$

determines the value of $m_0^2 a^2(\tilde{g}_0)$. As we have no reason to suppose that $\lim_{\tilde{g}_0\to\infty} K(\tilde{g}_0, a = 0)$ is not equal to K_c (Ising model critical point) we obtain, using that hypothesis, the same result here as in Eq. (26). A special case occurs in $d = 1$, as we know from general results that $\lim_{a\to 0} K(\tilde{g}_0, a) = \infty$ and $m_0 a \ll 1$, Eq. (28) becomes, according to the estimate (23),

$$m_0^2 a^2 \simeq \frac{3\tilde{g}_0}{K^2 m_0 a} - 2 + \frac{2A(\tilde{g}_0)}{K}. \tag{30}$$

Thus comparing orders of magnitude,

$$\lim_{a\to 0}\left(\tfrac{2}{3}m_0 a/\tilde{g}_0\right) = \lim_{a\to 0} K^{-2} = 0. \tag{31}$$

Result (31) holds uniformly in $0 < \tilde{g}_0 < \infty$, so that we can conclude

$$\lim_{\tilde{g}_0\to 0}\lim_{a\to 0}(m_0 a/\tilde{g}_0) = 0, \tag{32}$$

and also,

$$\lim_{\tilde{g}_0\to\infty}\lim_{a\to\infty}(m_0 a) = 0. \tag{33}$$

In order to extend our study of the limit as $\tilde{g}_0 \to 0$ to $d \geq 2$ dimensions we first observe

$$-2d + 2A(\tilde{g}_0)/K(\tilde{g}_0, 0) \leq 0 \tag{34}$$

uniformly in $\tilde{g}_0$. This remark is, of course, trivial in $d = 1$. It is also trivial for $\tilde{g}_0 \geq (\Gamma(3/4)/\Gamma(1/4))^2 \simeq 0.114\,236\,6452$, as[6] $A(\tilde{g}_0) \leq 0$ and $K \geq 0$ in the range. Since $A(\tilde{g}_0)$ is known[6] to be a monotonically decreasing function of $\tilde{g}$, and the left-hand side of (34) is exactly zero for $\tilde{g}_0 = 0$ by the well-known solution of the Gaussian model, (34) will follow from the monotonicity of $K(\tilde{g}_0, a)$ in $\tilde{g}_0$ near $a = 0$. This monotonicity is born out numerically.[6] We can see that it holds for very small $\tilde{g}_0$ analytically as follows. For a general hypercubic lattice, Baker and Kincaid's[6] combinatorial data yields for the Wortis method[11] in terms of the renormalized cumulants M_n,

$$\xi^2 = KM_2/(1 - 2dKM_2(K, \tilde{g}_0)) + O(\tilde{g}_0^2), \tag{35}$$

where use has been made of the results $M_4 \simeq -4!\tilde{g}_0$, and $M_n \simeq O(\tilde{g}_0^2)$, $n \geq 6$. Thus for $\tilde{g}_0$ small enough, we obtain

$$K(\tilde{g}_0, a) = [(2d + m^2a^2)M_2(K, \tilde{g}_0)]^{-1} + O(\tilde{g}_0^2). \tag{36}$$

We may expand

$$M_2(K, \tilde{g}_0) = 1 - 24\,\tilde{g}_0 S(K) + O(\tilde{g}_0^2), \tag{37}$$

where we have the high-temperature expansions[12]

$$\begin{aligned} S(K) &= K^2 + 3K^4 + 10K^6 + 35K^8 + \cdots, && d = 1,\\ &= 2K^2 + 18K^4 + 200K^6 + 2450K^8 + \cdots, && d = 2,\\ &= 3K^2 + 45K^4 + 930K^6 + 223\,65K^8 + \cdots, && d = 3,\\ &= 4K^2 + 84K^4 + 2560K^6 + 950\,60K^8 + \cdots, && d = 4. \end{aligned} \tag{38}$$

Now by examining the diagrams which lead to this series, we find at once that $1 + 2S(K)$ is the generating function of random walks which begin and end at the origin. Thus[13]

$$1 + 2S(K) = \frac{1}{(2\pi)^d}\int_{-\pi}^{\pi}\cdots\int\prod_{i=1}^{d} d\theta_i\left(1 - 2K\left(\sum_{i=1}^{d}\cos\theta_i\right)\right)^{-1}. \tag{39}$$

Manifestly, $S(K)$ is positive by (38) and (39), thus we see analytically that $K(\tilde{g}_0, a)$ does increase with $\tilde{g}_0$ for $\tilde{g}_0$ small enough. If we now consider (28) for $d \geq 3$, so that both $K_c(\tilde{g}_0)$ and $c(d)$ are finite, we have by (34)

$$m_0^2 a^2 \leq (12/K^2)\tilde{g}_0 c(d), \tag{40}$$

as $Ca^{d-2} \leq c(d)$. Thus, as $K \geq 2d$ we conclude

$$0 \leq \lim_{\tilde{g}_0\to 0}\lim_{a\to 0}(m_0^2a^2/\tilde{g}_0) \leq 3c(d)/d^2 < \infty, \tag{41}$$

as (41) holds uniformly for all $0 < \tilde{g}_0 < \infty$.

For the final case $d = 2$, (Ca^{d-2}) diverges weakly so we find from (28), using (23) and again (34)

$$0 \leq \lim_{\tilde{g}_0\to 0}\lim_{a\to 0}\left[-m_0^2a^2/((\ln\tilde{g}_0)\tilde{g}_0)\right] \leq \frac{3}{4\pi}. \tag{42}$$

These bounds (40)–(42) depend on the hypothesis that (34) is correct, which in turn follows from the idea that $K(\tilde{g}_0, 0)$ is monotonic in $\tilde{g}_0$ over the range $0 \leq \tilde{g}_0 \leq 0.1142$.

If we pick up from (16) and (13) the term which is set to zero in the derivation of (17) we have, along path (b), Figs. 1 and 2, which is appropriate to (14),

$$\lim_{\substack{a\to 0\\ \tilde{g}_0 \text{ fixed}}} a\frac{d}{da}\{\ln[2\tilde{A}(\tilde{g}_0)(K^{-1}(\tilde{g}_0, a) - K^{-1}(\tilde{g}_0, 0))a^{-2}]\}\big|_{\tilde{g}_0} = 0. \tag{43}$$

The last derivative term in (16) can be shown, by use of (45), to vanish in this limit [assuming (17)] and so is not included in (43). If further, in the neighborhood of $a = 0$, $\tilde{g}_0$ fixed we introduce the representation

$$\xi \simeq D_+(\tilde{g}_0)(1 - K/K_c)^{-\nu}, \tag{44}$$

then using (44) and (9), we may write

$$K(\tilde{g}_0,a)\simeq K(\tilde{g}_0,0)[1-(D_+(\tilde{g}_0)ma)^{1/\nu}]. \tag{45}$$

Thus, using (3), we obtain from (43), after taking the limit,

$$(4-d)\tilde{g}_0\left[\frac{\tilde{A}'(\tilde{g}_0)}{\tilde{A}(\tilde{g}_0)}-\frac{K'(\tilde{g}_0,0)}{K(\tilde{g}_0,0)}+\frac{1}{\nu}\frac{D'_+(\tilde{g}_0)}{D_+(\tilde{g}_0)}\right]+\frac{1}{\nu}-2=0. \tag{46}$$

If we integrate (46) over $\tilde{g}_0$, we obtain (assuming ν to be independent of $\tilde{g}_0$ as is contemplated in the renormalization group theory of critical phenomena)

$$[D_+(\tilde{g}_0)]^{(1/\nu)}=\Theta\frac{K(\tilde{g}_0,0)}{\tilde{A}(\tilde{g}_0)}\tilde{g}_0^{[(2-1/\nu)/(4-d)]}, \tag{47}$$

where Θ is a constant of integration which could depend, for example, on m, but not on $\tilde{g}_0$. For $d=1$, (45)–(47) are not expected to hold as $K(\tilde{g}_0,0)=\infty$. For $d\geqslant 2$, as $0<K(\tilde{g}_0,0)<\infty$ for all $\tilde{g}_0$ and $\tilde{A}(\tilde{g}_0)=0$ for $\tilde{g}_0=0.1142\cdots$ as remarked above, (47) implies a singular amplitude for $D_+(\tilde{g}_0)$. For $\tilde{g}_0\to 0$ the vanishing of D_+ is presumably related to the change of ν to $1/2$ in the Gaussian model ($g_0=0$) limit. The implication that $D_+(\tilde{g}_0)\to 0$ as $\tilde{g}_0\to\infty$ seems difficult to reconcile with the nonzero values determined numerically for this limit. Also, as $D_+(\tilde{g}_0)$ should be real and positive, the change of sign of the right-hand side of (47) at $\tilde{g}_0\simeq 0.1142\cdots$ is difficult to accomodate, unless, for example, ν changes there.

From the above analysis, using standard (nonrigorous) representations of some of the general features observed numerically for the quantities involved, we conclude that the full renormalization group hypothesis does not seem easy to accomodate over the whole range $\tilde{g}_0$. We remind the reader that although the hyperscaling index relations are known to hold for the two dimensional Ising model, neither the critical value of the renormalized coupling constant, g^*, nor the critical indices as predicted by the renormalization group theory[1] have been computed to any reasonable degree of accuracy to permit checking against the exact Ising model values. In three dimensions, where the renormalization group theory results seem to be more accurately determined,[1] as we remarked above, there may be small but persistent differences with the Ising model results for g^* and the critical indices. In one dimension, the Josephson relation, "$d\nu=2-\alpha$" is known[14] to fail although other hyperscaling relations hold.

ACKNOWLEDGMENTS

The author wishes to thank Daniel Bessis, Fred Cooper, M. E. Fisher, and B. Nickel for many helpful discussions and suggestions, for correspondence, and for a careful reading of earlier drafts of this manuscript.

[1] G. A. Baker, Jr., B. G. Nickel, and D. I. Meiron, Phys. Rev. B **17**, 1365 (1978); J. C. LeGuillou and J. Zinn-Justin, Phys. Rev. B **21**, 3976 (1980).
[2] K. G. Wilson, Phys. Rev. B **4**, 3174, 3184 (1971); E. Brezin, J. C. LeGuillou, and J. Zinn-Justin, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green(Academic, New York, 1976), Vol. 6, p. 125.
[3] C. Domb in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, New York, 1974), Vol. 3, p. 357.
[4] See, for example, J. Zinn-Justin, J. Phys. (Paris) **40**, 63 (1979); **42**, 783 (1981).
[5] G. A. Baker, Jr., Phys. Rev. B **15**, 1552 (1977).
[6] G. A. Baker, Jr. and J. M. Kincaid, Phys. Rev. Lett. **42**, 1431 (1979); G. A. Baker, Jr. and J. M. Kincaid, J. Statist. Phys. **24**, 469 (1981).
[7] G. A. Baker, Jr., J. Math. Phys. **16**, 1324 (1975).
[8] F. Constantinescu, Phys. Rev. Lett. **43**, 1632 (1979); C. M. Bender, F. Cooper, G. S. Garalnick, and D. Sharp, Phys. Rev. D **19**, 865 (1979); G. Caginalp, Rockefeller Univ. preprint (1979).
[9] E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).
[10] M. E. Fisher and D. S. Gaunt, Phys. Rev. **133**, A224 (1964).
[11] M. Wortis in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, New York, 1974), Vol. 3, p.114.
[12] J. M. Kincaid, G. A. Baker, Jr., and L. W. Fulterton, Los Alamos National Lab. Report No. LA-UR-79-1575.
[13] E. W. Montroll, in *Theory of Neutral and Ionized Gases*, edited by C. DeWitt and J.-F. Detoeuf (Wiley, New York, 1960), p.15.
[14] G. A. Baker, Jr. and J. C. Bonner, Phys. Rev. B **12**, 3741 (1975).

Existence and uniqueness of generalized vortices

M. A. Lohe
Department of Mathematical Physics, The University of Adelaide, Adelaide, 5000 South Australia
John van der Hoek
Department of Pure Mathematics, The University of Adelaide, Adelaide, 5000 South Australia

(Received 12 May 1981; accepted for publication 6 October 1981)

We investigate properties of the static noninteracting vortices determined by equations which generalize the first order Ginzburg–Landau equations. We prove that for each set of n points in the plane a unique solution exists to the first-order equations, with vortex number n. These n points mark the positions of the n vortices and are the only points at which the Higgs field $|\phi|$ vanishes. Regularity properties of the solution are related to those of an arbitrary non-negative function in the theory.

PACS numbers: 11.10.Np

I. INTRODUCTION

Vortex solutions of the abelian Higgs model have been the subject of detailed investigations in recent years. Properties of the static solutions, which are determined by the Ginzburg–Landau equations, depend on the value of a dimensionless parameter λ. Of particular interest is the critical value $\lambda = 1$, for then the static vortices do not interact and stable multivortex solutions exist. Only for $\lambda = 1$ do the masses of the Higgs and gauge mesons become equal, and intervortex forces cancel exactly. The mass of the multivortex configuration depends linearly on the vortex number n, and solutions can be obtained by solving certain first-order equations, as Bogomol'nyi[1] has shown. These first-order equations are of particular interest because their structure is related to that of the self-dual equations.[2,3] Following the analysis of Weinberg,[4] one expects the most general solution to depend on the $2n$ parameters which determine the vortex positions. Taubes[5] has recently verified this by proving that such an n-vortex solution exists and is unique.

The purpose of this paper is to show that the results of Taubes can be extended to apply to a much wider range of models. It has recently been shown[6] that the noninteracting phenomenon of vortices is not unique to the particular ϕ^4 interaction of the Higgs model, but is a general property for Hamiltonians of the form

$$\mathcal{H} = \tfrac{1}{4}(F_{ij})^2 + F(|\phi|)(D_i\phi^a)^2 + V(|\phi|). \tag{1.1}$$

Here (ϕ^1,ϕ^2) is a two-component real Higgs field, with $D_i\phi^a = \partial_i\phi^a - \epsilon^{ab}\phi^b A_i$, and we allow only two space dimensions. The expression (1.1) is the Hamiltonian for a static system in the gauge $A_0 = 0$. F and V are continuous real functions of $|\phi|$, to ensure gauge invariance, and are non-negative to ensure positive energy. This Hamiltonian allows vortex solutions provided V has a unique minimum at a non-zero value of $|\phi|$, to produce symmetry breaking. In order to obtain the noninteracting property we define V, for each F, according to

$$V = \tfrac{1}{2}w^2, \tag{1.2}$$

where

$$w(|\phi|) = \int_{|\phi|}^{1} sF(s)\,ds. \tag{1.3}$$

Evidently V is non-negative and has one minimum, which we have chosen by a suitable rescaling to lie at $|\phi| = 1$. A special case is $F \equiv 1$, for which $V = \frac{1}{8}(|\phi|^2 - 1)^2$, and we regain the Higgs model with the special coupling constant $\lambda = 1$ mentioned above.

The energy is bounded below by $2\pi n w(0)$ and this bound is attained if and only if[6]

$$F_{12} + w(|\phi|) = 0, \tag{1.4}$$

$$D_i\phi^a - \epsilon_{ij}\epsilon^{ab}D_j\phi^b = 0. \tag{1.5}$$

The vortex number n is defined by

$$n = \frac{1}{2\pi}\int_{\mathbb{R}^2} F_{21}\,dx, \tag{1.6}$$

and we have assumed $n > 0$ but the case $n < 0$ is treated analogously. The clue to the noninteracting property is that the energy depends linearly on n. Weinberg's analysis[4] generalizes to suggest that an n-vortex solution of Eqs. (1.4) and (1.5) depends on precisely $2n$ parameters, the vortex positions, which correspond to the zeros of the Higgs field $|\phi|$.

Equations (1.4) and (1.5) are a simple generalization of those obtained by Bogomol'nyi,[1] and as before[2,5] can be reduced to a single nonlinear equation. Writing $\phi^1 + i\phi^2 = |\phi|e^{i\alpha}$, we find from Eq. (1.5)

$$A_i = -\partial_i\alpha - \epsilon_{ij}\partial_j(\ln|\phi|), \tag{1.7}$$

and from Eq. (1.4)

$$\Delta\ln|\phi| + w(|\phi|) = [\partial_1,\partial_2]\alpha. \tag{1.8}$$

α is a gauge parameter, which we choose to be

$$\alpha = \sum_{i=1}^{n}\arctan\left(\frac{x_2 - a_2^i}{x_1 - a_1^i}\right), \tag{1.9}$$

where n is the vortex number, and the $2n$ parameters (a^i) are the vortex positions. Equation (1.8) becomes

$$\Delta\ln|\phi| + w(|\phi|) = 2\pi\sum_{i=1}^{n}\delta(x - a^i), \tag{1.10}$$

and is supplemented by the boundary conditions necessary for finite energy

$$\begin{aligned}&\lim_{|x|\to\infty}|\phi| = 1,\\ &\lim_{x\to a^i}\ln|\phi| = n_i\ln|x - a^i|,\end{aligned} \tag{1.11}$$

 0022-2488/83/010148-06$02.50

for $i = 1,...,n$, where n_i is the order of vanishing of $|\phi|$ at a^i.

For $F \equiv 1$ Taubes[5] has shown that each set of n points in $\mathbb{R}^2$ determines a unique classical solution to the first order Ginzburg–Landau equations. We obtain a similar result in general (Theorem I) provided F is restricted as follows:

(i) $F(s)$ is continuous and $F(s) \geqslant 0$ on $[0,\infty)$,

(ii) $F(1) > 0$,

(iii) $\{s \in [0,1]: F(s) = 0\}$ has Lebesgue measure zero.

The condition (ii), used in Propositions 3.7 and 4.1 below, states simply that the mass m of the elementary excitations, namely the Higgs and gauge mesons which are of equal mass, is nonzero since $m^2 = F(1)$. The condition $F(1) > 0$ also ensures that the minimum of V at $|\phi| = 1$ is unique. The condition (iii) is used in Proposition 3.10 to guarantee uniqueness of the solution.

We prove Theorem I by following the same strategy as in Ref. 5, defining on the appropriate Banach space a functional $\tilde{a}$ which is minimized by Eq. (1.10). Several steps in Ref. 5 require modification to accommodate general functions $F(|\phi|)$. In particular, we require the *a priori* result that all weak solutions of Eqs. (1.10) and (1.11) satisfy $|\phi| \leqslant 1$. Regularity properties of the solution then depend on those of F on $[0,1]$; in particular, the fields ϕ^a, A_i will be C^∞ provided F is C^∞ on $[0,1]$. In this case the arguments of Ref. 5 show that we have obtained all finite energy solutions of Eqs. (1.4) and (1.5).

A problem that remains concerns the equivalence of the first-order equations (1.4) and (1.5) and the second-order equations obtained by varying the Hamiltonian (1.1), assuming finite energy. Taubes[3] has demonstrated the equivalence of the first- and second-order formulations for $F \equiv 1$, and although the proof generalizes for functions F with certain restrictions, such as $2F + |\phi| F' \geqslant 0$ for all $|\phi|$, modifications appear to be necessary for a general proof.

II. DEFINITIONS

Following Taubes,[5] let

$$u_0 = -\sum_{i=1}^{n} \ln\left(1 + \frac{\lambda}{|x - a^i|^2}\right), \tag{2.1}$$

where λ is a suitably large real number to be chosen below (Remark 4.3). Define

$$g_0 = 4\sum_{i=1}^{n} \frac{\lambda}{(|x - a^i|^2 + \lambda)^2}, \tag{2.2}$$

so that on $\mathbb{R}^2 \setminus \cup_{i=1}^{n} \{a^i\}$, g_0 and $-\Delta u_0$ agree. We note that $u_0, g_0, 1 - e^{u_0} \in L^2(\mathbb{R}^2)$. Define the unknown function v by

$$u_0 + v = 2\ln|\phi|, \tag{2.3}$$

and if $|\phi|$ satisfies Eq. (1.10) with the boundary conditions (1.11) then v is a solution to

$$\Delta v + 2w(e^{(u_0+v)/2}) - g_0 = 0 \tag{2.4a}$$

with

$$\lim_{|x| \to \infty} v = 0. \tag{2.4b}$$

Define also the functional

$$a(v) = \int_{\mathbb{R}^2} \{\tfrac{1}{2}|\nabla v|^2 - v(2w(0) - g_0) + W(e^{(u_0+v)/2}) - W(e^{u_0/2})\}dx, \tag{2.5}$$

where

$$W(x) = 4\int_0^x \frac{dt}{t}\int_0^t sF(s)\,ds, \tag{2.6}$$

and where we have used the notation $\nabla v = (\partial_1 v, \partial_2 v)$. The variational equation of $a(\cdot)$ is formally Eq. (2.4a). However, we will find it useful to define another functional $\tilde{a}$ which can be more easily extended to a nonlinear functional on a suitable Sobolev space. Define $\tilde{F}$ by

$$\begin{aligned} \tilde{F}(s) &= F(s), \quad 0 \leqslant s \leqslant 1 \\ &= F(1), \quad s > 1 \end{aligned} \tag{2.7}$$

and then define $\tilde{w}$, $\tilde{W}$, and the functional $\tilde{a}$ in the same way as above, by simply replacing F by $\tilde{F}$.

Now let us mention the various spaces used below. $C^k(\mathbb{R}^2)$ denotes the space of functions with derivatives of order k continuous on $\mathbb{R}^2$. $C^\infty(\mathbb{R}^2)$ is the intersection of the spaces $C^k(\mathbb{R}^2)$ over all k, and $C_0^\infty(\mathbb{R}^2)$ is the space of infinitely differentiable functions with compact support. The Sobolev space $W^{m,p}(\mathbb{R}^2)$ is defined as the completion of $C_0^\infty(\mathbb{R}^2)$ in the norm

$$\|v\|_{m,p} = \sum_{\alpha_1+\alpha_2 \leqslant m} \left\| \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} v \right\|_p, \tag{2.8}$$

with α_1, α_2 non-negative integers. Here, the L^p norm $\|v\|_p$ is defined by

$$\|v\|_p = \left[\int_{\mathbb{R}^2} |v|^p dx\right]^{1/p}. \tag{2.9}$$

Properties of the spaces $W^{m,p}(\mathbb{R}^2)$ may be found in Adams.[7]

We will encounter the weak form of Eq. (2.4a). Define

$$\langle \text{grad } a(v), h \rangle = \int_{\mathbb{R}^2} \{\nabla v \cdot \nabla h + g_0 h - 2w(e^{(u_0+v)/2})h\}dx. \tag{2.10}$$

By a weak solution of Eq. (2.4) we mean a function $v \in W^{1,2}(\mathbb{R}^2) \cap C(\mathbb{R}^2)$ such that for all $h \in W^{1,2}(\mathbb{R}^2)$ $\langle \text{grad } a(v), h \rangle = 0$. We will prove (Lemma 3.9) that a weak solution satisfies $u_0 + v \leqslant 0$, implying that the solution of Eq. (2.4) minimizes both functionals $\tilde{a}$ and a.

Let us now state the main result of this paper:

Theorem I: For every point $(a_1,...,a_n)$ in $\mathbb{R}^2 \times \mathbb{R}^2 \times \cdots \times \mathbb{R}^2 = \mathbb{R}^{2n}$ and u_0 and g_0 defined by Eqs. (2.1) and (2.2), respectively, there exists a unique function $v \in W^{1,2}(\mathbb{R}^2) \cap C(\mathbb{R}^2)$ which is a weak solution of (2.4a) and (2.4b).

Regularity properties of v, which depend on F, are discussed in Sec. V. Since F is continuous we always have $v \in C^1(\mathbb{R}^2)$.

III. PROPERTIES OF THE FUNCTIONAL $\tilde{a}$

Proposition 3.1: $\tilde{a}$ defined on $C_0^\infty(\mathbb{R}^2)$ can be extended to a nonlinear functional with domain $W^{1,2}(\mathbb{R}^2)$.

Proof: We show that $\tilde{a}(v)$ is finite for each $v \in W^{1,2}(\mathbb{R}^2)$.

Now

$$\tilde{a}(v) = \int_{\mathbb{R}^2} \Big\{ \tfrac{1}{2}|\nabla v|^2 + vf + \tilde{W}(e^{(u_0+v)/2}) - \tilde{W}(e^{u_0/2}) - 2v\int_0^{e^{u_0/2}} s\tilde{F}(s)\,ds \Big\} dx,$$

where $f \equiv -2\tilde{w}(e^{u_0/2}) + g_0$. Let $A = \sup\{F(s): s\in[0,1]\}$ then as $|f| \leqslant g_0 + A(1-e^{u_0}), f\in L^2(\mathbb{R}^2)$ and hence $\int_{\mathbb{R}^2} vf\,dx$ is finite. Since

$$\tilde{W}(e^{(u_0+v)/2}) - \tilde{W}(e^{u_0/2}) - 2v\int_0^{e^{u_0/2}} s\tilde{F}(s)ds = \int_{e^{u_0/2}}^{e^{(u_0+v)/2}} 4\frac{dt}{t}\int_{e^{u_0/2}}^{t} s\tilde{F}(s)ds,$$

$$\left|\tilde{W}(e^{(u_0+v)/2}) - \tilde{W}(e^{u_0/2}) - 2v\int_0^{e^{u_0/2}} s\tilde{F}(s)ds\right| \leqslant Ae^{u_0}(e^v - v - 1). \tag{3.1}$$

The proof now continues as in Ref. 5, Proposition 4.1. ■

Proposition 3.2: The Gâteaux derivative of $\tilde{a}$, $\tilde{a}'(v;h)$, exists for all $v,h\in W^{1,2}(\mathbb{R}^2)$ and

$$\tilde{a}'(v;h) = \lim_{t\to 0}\frac{1}{t}[\tilde{a}(v+th) - \tilde{a}(v)] = \int_{\mathbb{R}^2}\{\nabla v\cdot\nabla h + g_0 h - 2\tilde{w}(e^{(u_0+v)/2})h\}dx. \tag{3.2}$$

Furthermore, for fixed v, $\tilde{a}'(v;\cdot)$ is a bounded linear functional on $W^{1,2}(\mathbb{R}^2)$. For fixed $h\in W^{1,2}(\mathbb{R}^2)$, $\tilde{a}'(\cdot;h)$ is a nonlinear functional with domain $W^{1,2}(\mathbb{R}^2)$.

Proof: Let v, $h\in W^{1,2}(\mathbb{R}^2)$, then

$$(1/t)[\tilde{a}(v+th) - \tilde{a}(v)] = \int_{\mathbb{R}^2}\{\nabla v\cdot\nabla h + g_0 h - 2\tilde{w}(e^{(u_0+v)/2})h\}dx + \frac{t}{2}\int_{\mathbb{R}^2}|\nabla h|^2 dx + \int_{\mathbb{R}^2}\{-2\tilde{w}(0)h + 2\tilde{w}(e^{(u_0+v)/2})h + \frac{1}{t}(\tilde{W}(e^{(u_0+v+th)/2}) - \tilde{W}(e^{(u_0+v)/2}))\}dx. \tag{3.3}$$

The integrand in the third term on the right-hand side of (3.3) equals

$$\frac{1}{t}\int_{e^{(u_0+v)/2}}^{e^{(u_0+v+th)/2}} 4\frac{d\xi}{\xi}\int_{e^{(u_0+v)/2}}^{\xi} s\tilde{F}(s)\,ds$$

and this in modulus, using arguments like those which lead to (3.1), does not exceed

$$(A/t)e^{u_0+v}(e^{th} - th - 1).$$

Equation (3.2) now follows from the arguments in Ref. 5, Proposition 4.2. We note also that for v, $h\in W^{1,2}(\mathbb{R}^2)$,

$$|\tilde{a}'(v;h)| \leqslant \left\{\left(\int_{\mathbb{R}^2}|\nabla v|^2 dx\right)^{1/2} + \left(\int_{\mathbb{R}^2} f^2 dx\right)^{1/2} + A\left(\int_{\mathbb{R}^2}(e^v - 1)^2 dx\right)^{1/2}\right\}\|h\|_{1,2},$$

where f and A are as in the proof of Proposition 3.1. ■

Remark 3.3: For each $v\in W^{1,2}(\mathbb{R}^2)$ we let grad $\tilde{a}(v)$ denote the bounded linear operator $\tilde{a}'(v;\cdot)$ defined by $\langle \text{grad } \tilde{a}(v), h\rangle = \tilde{a}'(v;h)$ for each $h\in W^{1,2}(\mathbb{R}^2)$. We will let $a'(v;h)$ denote the formal expression

$$\int_{\mathbb{R}^2}\{\nabla v\cdot\nabla h + g_0 h - 2w(e^{(u_0+v)/2})h\}dx, \tag{3.4}$$

where

$$w(t) = \int_t^1 sF(s)ds. \tag{3.5}$$

For suitable $v\in W^{1,2}(\mathbb{R}^2)$, e.g., when $v + u_0 \leqslant 0$, $a'(v;)$ will be a bounded linear functional on $W^{1,2}(\mathbb{R}^2)$. In this case we let grad $a(v)$ denote this bounded linear functional, in agreement with the definition of Eq. (2.10).

Proposition 3.4: $\tilde{a}$ is a convex functional on $W^{1,2}(\mathbb{R}^2)$.

Proof: For u, $v\in W^{1,2}(\mathbb{R}^2)$ and using the fact that $\tilde{F}\geqslant 0$,

$$\langle \text{grad } \tilde{a}(u) - \text{grad } \tilde{a}(v), u - v\rangle = a'(u,u-v) - a'(v,u-v) = \int_{\mathbb{R}^2}\{|\nabla u - \nabla v|^2 + 2(u-v)(\tilde{w}(e^{(u_0+u)/2}) - \tilde{w}(e^{(u_0+v)/2}))\}dx \geqslant 0. \tag{3.6}$$

Equation (3.6) implies that the Gâteaux derivative of $\tilde{a}$ is monotone. The convexity of $\tilde{a}$ follows from Ref. 8, Theorem 5.1. ■

Proposition 3.5: v_0 minimizes $\tilde{a}$ on $w^{1,2}(\mathbb{R}^2)$ if and only if grad $\tilde{a}(v_0) = 0$.

Proof: This follows from Propositions 3.2 and 3.4 and Ref. 8, Theorem 9.1. ■

Proposition 3.6: $\tilde{a}$ is weakly lower semicontinuous on $W^{1,2}(\mathbb{R}^2)$.

Proof: This follows from Proposition 3.2 and Ref. 8, Theorem 8.6. ■

Proposition 3.7: Let $v\in W^{1,2}(\mathbb{R}^2)\cap C(\mathbb{R}^2)$; then grad $\tilde{a}(v) = 0$ if and only if grad $a(v) = 0$.

We state without proof the following:

Lemma 3.8: Let Ω be a bounded open set in $\mathbb{R}^2$ and $v\in W^{1,1}(\Omega)\cap C(\bar{\Omega})$. Suppose that $v(x) = 0$ for all $x\in\partial\Omega$, then

$$\int_\Omega \nabla v\,dx = 0.$$

Lemma 3.9: Let $v\in W^{1,2}(\mathbb{R}^2)\cap C(\mathbb{R}^2)$ and suppose that either grad $\tilde{a}(v) = 0$ or grad $a(v) = 0$, then $v + u_0 \leqslant 0$.

Proof: We give a proof assuming grad $a(v) = 0$. Let $\psi\in C_0^\infty(\mathbb{R}^2)$ have the properties $\psi(x) = 1$ for $|x|\leqslant 1$, $\psi(x) = 0$ for $|x|\geqslant 2$, $0\leqslant\psi(x)\leqslant 1$, and $|\nabla\psi(x)|\leqslant N$, $N>0$, for all $x\in\mathbb{R}^2$. For $R>0$, define ψ_R by $\psi_R(x) = \psi(x/R)$ for $x\in\mathbb{R}^2$. Let $\Omega_+ = \{x\in\mathbb{R}^2: u_0(x) + v(x) > 0\}$. Ω_+ is an open set. Suppose that Ω_+ is nonempty. By continuity of v, $a_k\notin\Omega_+$ for $k = 1,2,\ldots,n$. Define

$$\eta_R = \begin{cases}(e^{(u_0+v)/2} - 1)\psi_R & \text{on } \Omega_+,\\ 0 & \text{on } \mathbb{R}^2\setminus\Omega_+.\end{cases} \tag{3.7}$$

Then $\eta_R\in W^{1,2}(\mathbb{R}^2)$, and has support in the closure of $D_{2R}(0) = \{x\in\mathbb{R}^2: |x| < 2R\}$. As grad $a(v) = 0$,

$$0 = a'(v,\eta_R) = \int_{\Omega_{2R}}\{\nabla v\cdot\nabla\eta_R + g_0\eta_R - 2w(e^{(u_0+v)/2})\eta_R\}dx, \tag{3.8}$$

where $\Omega_\rho^+ = \Omega_+ \cap D_\rho(0)$. Putting $2u = u_0 + v$ we obtain

$$\int_{\Omega_{2R}^+} \{ -2w(e^u)(e^u - 1)\psi_R + 2(e^u - 1)\nabla u \cdot \nabla \psi_R - (e^u - 1)\nabla u_0 \cdot \nabla \psi_R + 2|\nabla u|^2 e^u \psi_R - \nabla u \cdot \nabla u_0 e^u \psi_R + g_0(e^u - 1)\psi_R \} dx = 0. \tag{3.9}$$

Now

$$(e^u - 1)\frac{\partial u_0}{\partial x_i}\psi_R \in W^{1,1}(\Omega_{2R}^+) \cap C(\bar{\Omega}_{2R}^+)$$

and equals zero on $\partial\Omega_{2R}^+$ for $i = 1,2$; hence by Lemma 3.8,

$$\int_{\Omega_{2R}^+} \nabla \cdot ((e^u - 1)\nabla u_0 \psi_R)\, dx = 0. \tag{3.10}$$

This equation implies that

$$-\int_{\Omega_{2R}^+} (e^u - 1)\nabla u_0 \cdot \nabla \psi_R\, dx = \int_{\Omega_{2R}^+} \{ e^u \nabla u \cdot \nabla u_0 - (e^u - 1) g_0 \} \psi_R\, dx \tag{3.11}$$

since $\Delta u_0 = -g_0$ on Ω_+. Substituting (3.11) into (3.9) yields

$$\int_{\Omega_{2R}^+} \{ |\nabla u|^2 e^u - w(e^u)(e^u - 1) \} \psi_R\, dx = -\int_{\Omega_{2R}^+} (e^u - 1)\nabla u \cdot \nabla \psi_R\, dx. \tag{3.12}$$

Since F is continuous and $F(1) > 0$ there exists a constant $d > 0$ such that $-w(e^u)(e^u - 1) \geq d \min(e^u - 1, (e^u - 1)^2)$, so by (3.12) and the fact that $\psi_R \equiv 1$ on Ω_R^+,

$$\begin{aligned} &\int_{\Omega_R^+} \{ |\nabla u|^2 e^u + d \min(e^u - 1, (e^u - 1)^2) \}\, dx \\ &\leq \int_{\Omega_{2R}^+} (e^u - 1)|\nabla u||\nabla \psi_R|\, dx \\ &\leq \frac{N}{R} \int_{\Omega_{2R}^+} |\nabla u||e^u - 1|\, dx. \end{aligned} \tag{3.13}$$

As $e^u - 1 = e^{u_0/2}(e^{v/2} - 1) + (e^{u_0/2} - 1)$, $|e^u - 1| \leq |e^v - 1| + |e^{u_0} - 1| \in L^2(\mathbb{R}^2)$, and as $|\nabla u_0| \in L^2(\mathbb{R}^2)$ we also have $|\nabla u| \in L^2(\mathbb{R}^2)$. We conclude from Hölder's inequality that

$$\int_{\Omega_{2R}^+} (e^u - 1)|\nabla u|\, dx \leq \left(\int_{\mathbb{R}^2} (e^u - 1)^2 dx \right)^{1/2} \times \left(\int_{\mathbb{R}^2} |\nabla u|^2 dx \right)^{1/2} < \infty.$$

Now letting $R \to +\infty$ in (3.13) we obtain

$$\int_{\Omega} \min\{ e^u - 1, (e^u - 1)^2 \}\, dx = 0.$$

Since u is continuous on Ω_+ we conclude that Ω_+ has zero measure, a contradiction. ■

Proof of Proposition 3.7: The proposition follows from Lemma 3.9 and the fact that

$$w(e^{(u_0 + v)/2}) = \tilde{w}(e^{(u_0 + v)/2})$$

whenever $u_0 + v \leq 0$. ■

Proposition 3.10: There is a unique function $v \in W^{1,2}(\mathbb{R}^2) \cap C(\mathbb{R}^2)$ such that

$$\tilde{a}(v) = \inf\{ \tilde{a}(u) : u \in W^{1,2}(\mathbb{R}^2) \}. \tag{3.14}$$

Proof: Suppose to the contrary that there are two functions $v_1, v_2 \in W^{1,2}(\mathbb{R}^2) \cap C(\mathbb{R}^2)$ which satisfy (3.14). By Proposition 3.5 it follows that grad $\tilde{a}(v_1) =$ grad $\tilde{a}(v_2) = 0$, and so by Lemma 3.9 that $v_1 + u_0 \leq 0$ and $v_2 + u_0 \leq 0$, and that $a'(v_1, v_1 - v_2) - a'(v_2, v_1 - v_2) = 0$. We shall contradict this last statement. Since $v_1 \neq v_2$ there exists a neighborhood N which does not intersect $\{a_k : k = 1,2,\ldots,n\}$ on which $v_1 \neq v_2$. For $x \in N$

$$\begin{aligned} &2(v_1(x) - v_2(x))[\tilde{w}(e^{[u_0(x) + v_1(x)]/2}) - \tilde{w}(e^{[u_0(x) + v_2(x)]/2})] \\ &= 2(v_1(x) - v_2(x)) \times \int_{e^{[u_0(x) + v_2(x)]/2}}^{e^{[u_0(x) + v_1(x)]/2}} sF(s)ds > 0, \end{aligned} \tag{3.15}$$

since the measure of $\{ s \in [0,1] : F(s) = 0 \}$ is zero. By (3.15) it follows that in fact $a'(v_1, v_1 - v_2) - a'(v_2, v_1 - v_2) > 0$. ■

IV. EXISTENCE AND UNIQUENESS OF WEAK SOLUTIONS

It follows from Propositions 3.5 and 3.7 that the existence and uniqueness of a weak solution to Eqs. (2.4a) and (2.4b) follows from the existence and uniqueness of a function $v \in W^{1,2}(\mathbb{R}^2) \cap C(\mathbb{R}^2)$ such that $\tilde{a}(v) = \inf\{ \tilde{a}(u) : u \in W^{1,2}(\mathbb{R}^2) \}$. The uniqueness follows from Proposition 3.10. The rest of this section deals with existence.

Proposition 4.1: There exist constants $\alpha > 0$, b and $k > 0$, such that for all $v \in W^{1,2}(\mathbb{R}^2)$,

$$\tilde{a}'(v;v) \geq \frac{\alpha \|v\|_{1,2}^2}{(1 + k\|v\|_{1,2})} - b. \tag{4.1}$$

In order to prove this proposition we prove some properties about u_0, g_0, and $\tilde{w}$.

Lemma 4.2: There exists a positive constant c such that for all $x \leq 0$,

$$\tilde{w}(e^x) \geq c(1 - e^x). \tag{4.2}$$

Proof: This follows from the continuity of F and the fact that $F(1) > 0$. ■

Remark 4.3: From now on we assume that the constant λ in the definition of u_0 and g_0 satisfies $\lambda > 4n/c$, where c is the constant in (4.2).

Lemma 4.4: Let u_0 and g_0 be defined as in Eqs. (2.1) and (2.2), respectively. Let $M > 0$, then for $\lambda > 4n/M$

(i) there exists a constant $c_1 > 0$ such that for all $x \in \mathbb{R}^2$,

$$M - g_0(x) \geq c_1; \tag{4.3}$$

(ii) for all $x \in \mathbb{R}^2$

$$-g_0(x) + M(1 - e^{u_0(x)}) > 0. \tag{4.4}$$

Proof: This lemma follows from a minor modification of Ref. 5, Lemma 5.2. ■

Proof of Proposition 4.1: For $v \in W^{1,2}(\mathbb{R}^2)$,

$$\tilde{a}'(v;v) = \int_{\mathbb{R}^2} \{ |\nabla v|^2 + g_0 v - 2\tilde{w}(e^{(u_0 + v)/2})v \}\, dx.$$

Let $\Omega_1 = \{ x \in \mathbb{R}^2 : v(x) \leq 0 \}$, $\Omega_2 = \{ x \in \mathbb{R}^2 : 0 \leq v(x) \leq -u_0(x) \}$,

and $\Omega_3 = \{x \in \mathbb{R}^2 : v(x) \geqslant -u_0(x)\}$. On Ω_1, using Lemma 4.2,

$$\begin{aligned} g_0 v - 2\tilde{w}(e^{(u_0+v)/2})v &= |v|(-g_0 + 2\tilde{w}(e^{(u_0+v)/2}) \\ &\geqslant |v|(-g_0 + 2c(1 - e^{(u_0+v)/2})) \\ &\geqslant |v|(-g_0 + c(1 - e^{u_0+v})) \\ &= |v|(-g_0 + c(1 - e^{u_0})) \\ &\quad + c|v|e^{u_0}(1 - e^{-|v|}) \\ &\geqslant |v|(-g_0 + c(1 - e^{u_0})) \\ &\quad + \frac{c|v|^2 e^{u_0}}{1+|v|}, \end{aligned}$$

where we have used the inequality

$$1 - e^{-x} \geqslant x/(1+x)$$

for $x \geqslant 0$. By the remark following Lemma 4.2 and by Lemma 4.4 it follows that

$$g_0 v - 2\tilde{w}(e^{(u_0+v)/2})v \geqslant \frac{c_1|v|^2}{1+|v|}. \tag{4.5}$$

Let $\Lambda = \sup_{s \geqslant 0} \tilde{F}(s) < \infty$. Then on Ω_2,

$$\begin{aligned} -2\tilde{w}(e^{(u_0+v)/2}) &= -2\int_{e^{(u_0+v)/2}}^{1} sF(s)ds \\ &\geqslant -2\Lambda\int_{e^{(u_0+v)/2}}^{1} s\,ds \\ &= \Lambda(e^{u_0+v} - 1) \geqslant \Lambda(u_0 + v). \end{aligned}$$

Hence

$$v(g_0 - 2\tilde{w}(e^{(u_0+v)/2})) \geqslant v(g_0 + \Lambda u_0) + \Lambda v^2. \tag{4.6}$$

On Ω_3

$$\begin{aligned} -2\tilde{w}e^{(u_0+v)/2}) &= 2F(1)\int_1^{e^{(u_0+v)/2}} s\,ds \\ &= F(1)(e^{u_0+v} - 1) \\ &\geqslant F(1)(u_0 + v), \end{aligned}$$

whence

$$v(g_0 - 2\tilde{w}(e^{(u_0+v)/2})) \geqslant v(g_0 + F(1)u_0) + F(1)v^2. \tag{4.7}$$

The rest of the proof follows from (4.5), (4.6), (4.7), and minor modifications of Ref. 5, Proposition 5.1. ■

Proposition 4.5: There exists $v \in W^{1,2}(\mathbb{R}^2) \cap C(\mathbb{R}^2)$ such that

$$\tilde{a}(v) = \min\{\tilde{a}(u) : u \in W^{1,2}(\mathbb{R}^2)\}.$$

Before proving this proposition we establish the following.

Lemma 4.6: For each $v \in W^{1,2}(\mathbb{R}^2)$, $\tilde{w}(e^{(u_0+v)/2}) \in L^2(\mathbb{R}^2)$.

Proof: Let $\Omega_1 = \{x \in \mathbb{R}^2 : u_0(x) + v(x) \leqslant 0\}$, $\Omega_2 = \mathbb{R}^2 \sim \Omega_1$. On Ω_1,

$$\begin{aligned} 0 \leqslant \tilde{w}(e^{(u_0+v)/2}) &= \int_{e^{(u_0+v)/2}}^{1} s\tilde{F}(s)ds \\ &\leqslant (\Lambda/2)(1 - e^{u_0+v}), \end{aligned} \tag{4.8}$$

where Λ is as in the proof of Proposition 4.1. On Ω_2,

$$\begin{aligned} \tilde{w}(e^{(u_0+v)/2}) &\quad -F(1)\int_1^{e^{(u_0+v)/2}} s\,ds \\ &= -\frac{F(1)}{2}(e^{u_0+v} - 1). \end{aligned} \tag{4.9}$$

Combining (4.8) and (4.9)

$$|\tilde{w}(e^{(u_0+v)/2})| \leqslant \text{const}\{e^{u_0}|e^v - 1| + |e^{u_0} - 1|\} \in L^2(\mathbb{R}^2).$$

■

Lemma 4.7: Let $G(x)$ be a real Gâteaux differentiable functional defined on a real reflexive Banach space E, which is weakly lower semicontinuous and satisfies the condition

$$\langle \text{grad } G(x), x \rangle > 0 \tag{4.10}$$

for any vector $x \in E$, $\|x\| = R > 0$. Then there exists an interior point x_0 of the ball $\{x \in E : \|x\| \leqslant R\}$ at which $G(x)$ has a local minimum so that grad $G(x_0) = 0$.

Proof: See Ref. 8, Theorem 9.8.

Proof of Proposition 4.4: By Propositions 3.2, 3.6, and 4.1 the assumptions of Lemma 4.7 are satisfied for R sufficiently large. As $W^{1,2}(\mathbb{R}^2)$ is a reflexive Banach space, there exists $v \in W^{1,2}(\mathbb{R}^2)$ with grad $\tilde{a}(v) = 0$, and so v is a distributional solution of

$$\Delta v = -2\tilde{w}(e^{(u_0+v)/2}) + g_0.$$

By Lemma 4.6, $\Delta v \in L^2(\mathbb{R}^2)$ and hence $v \in W^{2,2}(\mathbb{R}^2)$. By the Sobolev imbedding theorem, see Ref. 7, p. 97, we conclude that $v \in C(\mathbb{R}^2)$. The proposition now follows from Proposition 3.5. ■

V. PROPERTIES OF THE SOLUTION

We remark that since u_0 depends on λ then the unique minimizer v of $\tilde{a}(\cdot)$ on $W^{1,2}(\mathbb{R}^2)$ will also depend on the choice of λ. By the same argument as in Ref. 5, $u_0 + v$ can be shown to be independent of the choice of λ, for λ sufficiently large.

We noted in Proposition 4.4 and 4.8 that v is continuous on $\mathbb{R}^2$. We can show more. We show also that Lemma 3.9 can be sharpened.

Proposition 5.1:

(i) If the first k derivatives of F are bounded on the interval $[0,1]$, then $v \in C^{k+1}(\mathbb{R}^2)$.

(ii) If F is C^∞ on $[0,1]$, then $v \in C^\infty(\mathbb{R}^2)$.

Proof: Clearly (i) implies (ii). v is a weak solution of

$$\Delta v = -2w(e^{(u_0+v)/2}) + g_0, \tag{5.1}$$

and $u_0 + v \leqslant 0$ on $\mathbb{R}^2$. Let $v_i \equiv \partial v/\partial x_i$; then on differentiating (5.1) with respect to x_i,

$$\Delta v_i = e^{u_0+v}F(e^{(u_0+v)/2})\left(v_i + \frac{\partial u_0}{\partial x_i}\right) + \frac{\partial g_0}{\partial x_i},$$

which implies that $\Delta v_i \in L^2(\mathbb{R}^2)$ and hence that $v_i \in W^{2,2}(\mathbb{R}^2)$ for $i = 1,2,\ldots,n$ whence $v_i \in C(\mathbb{R}^2)$ by the Sobolev imbedding theorem (see Ref. 7, p. 97) or $v \in C^1(\mathbb{R}^2)$. Repeating this argument gives (i). Since $u_0 + v \leqslant 0$, the regularity properties of v depend only on those of F on $[0,1]$.

Lemma 5.2: Let F have bounded first derivatives on the interval $[0,1]$; then the weak solution of (5.1) is $C^2(\mathbb{R}^2)$ and $u_0 + v < 0$.

Proof: By the continuity of F there exists $0 < r < 1$ such that $F(s) \geqslant \gamma > 0$ for $r \leqslant s \leqslant 1$. Let $\Omega = \{x \in \mathbb{R}^2 : u_0(x) + v(x) > 2\ln r\}$. Then $a_k \notin \Omega$ for $k = 1,2,\ldots,n$. It will be sufficient to show that $u_0 + v < 0$ on Ω. We have $w = w(\psi(x))$, where $\psi = e^{(u_0+v)/2}$ as above. Then from the strong equation (5.1), Proposition 5.1(i), and using

$$\Delta u_0 = -g_0,$$
$$[\Delta - b\cdot\nabla - \psi^2 F(\psi)]w = 0, \tag{5.2}$$

where

$$b = \left(\frac{2}{\psi} + \frac{F'(\psi)}{F(\psi)}\right)\nabla\psi.$$

Clearly b is bounded and $\psi^2 F(\psi) > 0$ on Ω. By Lemma 3.9, $w \geqslant 0$ on Ω and so by the maximum principle, see Ref. 9, Theorem 3.5, $w > 0$ on Ω and hence $u_0 + v < 0$ on Ω. ■

Proposition 5.3: If F is real analytic on an open interval containing $[0,1)$ then v is real analytic on $\mathbb{R}^2$.

Proof: Since e^{u_0} is real analytic on $\mathbb{R}^2$ and $u_0 + v < 0$ on $\mathbb{R}^2$ by Lemma 5.2, the proposition follows from the theory in Ref. 10, Sec. 5.8. ■

Remark 5.4: We have shown that if F is C^∞ then the solution v is C^∞ and it follows that the fields A_i and ϕ^a are C^∞. In this case we may quote Proposition A1.1 of Taubes,[5] the proof of which does not rely on the particular form of the function F, and which shows that the zeros of $|\phi|$ are discrete. A proof is also to be found in Ref. 11, p. 76. This implies that when F is C^∞ we have found all solutions to the first-order equations (1.4) and (1.5). Thus,

Proposition 5.5: Let A_i and ϕ^a be, respectively, C^∞ gauge and Higgs fields on $\mathbb{R}^2$ which satisfy Eqs. (1.4) and (1.5). Then $\{x \in \mathbb{R}^2 : |\phi|(x) = 0\}$ is discrete.

[1] E. B. Bogomol'nyi, Sov. J. Nucl. Phys. **24**, 449 (1976).
[2] M. A. Lohe, Phys. Lett. **70** B, 325 (1977).
[3] C. H. Taubes, Commun. Math. Phys. **75**, 207 (1980).
[4] E. Weinberg, Phys. Rev. D **19**, 3008 (1979).
[5] C. H. Taubes, Commun. Math. Phys. **72**, 277 (1980).
[6] M. A. Lohe, Phys. Rev. D **23**, 2335 (1981).
[7] R. Adams, *Sobolev Spaces* (Academic, New York, 1975).
[8] M. M. Vainberg, *Variational method and method of monotone operators in the theory of nonlinear equations* (Wiley, New York, 1973).
[9] D. Gilbarg and N. Trudinger, *Elliptic partial differential equations of second order* (Springer, New York, 1977).
[10] C. Morrey, *Multiple integrals in the calculus of variations* (Springer, New York, 1966).
[11] A. Jaffe and C. Taubes, *Vortices and monopoles* (Birkhäuser, Boston, 1980).

Compactificaton partly proven

Richard Stacey
Department of Physics and Astronomy, University College London, Gower Street, London WC1 6BT, England[a)]

(Received 18 September 1981, accepted for publication 30 October 1981)

For a restricted class of SU(2) gauge-field structures we show that only integral topological charges can occur, without making any assumptions about the asymptotic behavior of the fields

PACS numbers: 11.10.Np

1. INTRODUCTION

It is generally believed that in SU(2) gauge theories an arbitrary finite-action gauge-field structure *must* have integral topological charge. However; the usual "proof" of this[1] depends on properties of S^4, the one-point compactification of R^4. It *assumes* that all gauge-field structures can be extended from R^4 to S^4. This Compactification Assumption has been attacked by Crewther,[2] who has argued that nonintegral topological charges are important.

Recently[3] I showed that a proposed counterexample to compactification,[4] with topological charge $\frac{3}{2}$, was not valid. Reference 3 also quoted some results to the effect that all solutions in a wide class of gauge-field structures were instantons, hence compactifiable.

Here I will extend and prove those results. They reinforce our belief in the integral nature of topological charge, and perhaps provide the first stages of a full proof.

In Sec. 2 we define the class $C_{r,t}$ of gauge-field structures to be considered. It consists of self-dual fields A^i_μ defined by a familiar ansatz

$$A^i_\mu \sim \ln \rho_i$$

on patches P_i. All the ρ_i are taken to be functions of just two variables (r and t). This simplification allows us to prove some important crossing properties (Sec. 3). We use these in Sec. 4 to show that any $\rho_i(P_i)$ can be extended to all R^4 apart from isolated removable singularities (instantons). In Sec. 5 we find the complete solution for the class $C_{r,t}$. Section 6 contains some comments.

2. DEFINITION OF $C_{r,t}$

We consider the class ($\equiv C_{r,t}$) of self-dual gauge-field structures (or connections) defined by

$$A^i_\mu = \sigma_{\mu\nu}\frac{\partial}{\partial x_\nu}\ln \rho_i(r,t) \tag{1}$$

on regions P_i, where $\cup_i P_i = R^4$ and

$$x_0 = t, \quad r = \sqrt{x_1^2 + x_2^2 + x_3^2},$$
$$\sigma_{\mu\nu} = \begin{cases} \sigma_{ij} = (1/4i)[\sigma_i,\sigma_j] \\ \sigma_{0i} = \frac{1}{2}\sigma_i . \end{cases} \tag{2}$$

On $P_i \cap P_j \neq \phi$ the gauge fields must be related by a gauge transformation

$$A^i_\mu = \Omega A^j_\mu \Omega^{-1} + i\partial_\mu \Omega \cdot \Omega^{-1}, \tag{3}$$

where

$$\Omega = \exp\left\{i\alpha(r,t)\frac{\sigma\cdot x}{2r}\right\} \tag{4}$$

for some function $\alpha(r,t)$.

Note that if $\rho_i = Re^{i\phi}$ then

$$\partial_\mu \ln \rho_i = \partial_\mu \ln R + i\partial_\mu \phi \in \mathrm{Re} \tag{5}$$

implies that ϕ is a constant whose value does not affect A^i_μ. We can therefore take it to be zero without loss of generality, making ρ_i real and positive. Observe that ρ_i cannot change sign since that would imply a point with $\rho_i = 0$ in P_i, at which A^i_μ would be singular, i.e., undefined.

3. CROSSING PROPERTIES

We prove the following.

Theorem 1: (i) For an arbitrary function $h(r+it)$ which is analytic in $P_i \cap P_j$, if

$$r\rho_i = \mathrm{Re}(ch), \tag{6}$$

$$r\rho_j = \mathrm{Re}(c^*/h), \tag{7}$$

($c \in \mathbb{C}$) then Eqs. (3) and (4) are satisfied with

$$\alpha(r,t) = \pi + 2\arg(ch). \tag{8}$$

(ii) Conversely, any nontrivial solution of Eqs. (1)–(4) can be expressed as in Eqs. (6)–(8) for some $c \in \mathbb{C}$ and some analytic (in $P_i \cap P_j$) function $h(r+it)$.

Proof: A little algebra shows that Eqs. (1)–(4) are equivalent to

$$\frac{\partial\alpha}{\partial t} = \frac{\partial}{\partial r}[\ln(r\rho_i/r\rho_j)], \tag{9}$$

$$\frac{\partial\alpha}{\partial t} = \tan\left(\frac{\alpha}{2}\right)\frac{\partial}{\partial t}[\ln(r\rho_i\cdot r\rho_j)], \tag{10}$$

$$\frac{\partial\alpha}{\partial r} = -\frac{\partial}{\partial t}[\ln(r\rho_i/r\rho_j)], \tag{11}$$

$$\frac{\partial\alpha}{\partial r} = \tan\left(\frac{\alpha}{2}\right)\cdot\frac{\partial}{\partial r}[\ln(r\rho_i\cdot r\rho_j)]. \tag{12}$$

It is easily checked that Eqs. (6)–(8) satisfy this set of equations, and also the self-duality equations

$$\frac{1}{r\rho_k}\left(\frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial t^2}\right)(r\rho_k) = 0 \tag{13}$$

($k = i, j$). This proves part (i) of the theorem.

To prove part (ii) we need to find the most general solution of Eqs. (9)–(13). The most general solution of Eq. (13) with ρ_k real is

[a)] Present address: DAMTP, Univ. of Liverpool, P.O. Box 147, Liverpool, United Kingdom.

 0022-2488/83/010154-03$02.50

$$r\rho_k = \mathrm{Re}\, q_k(r+it) \quad (k=i,j), \tag{14}$$

where q_k is arbitrary except that it must be once-differentiable (hence analytic) to define A^k_μ.

Equations (9)–(12) lead to two relations involving just ρ_i and ρ_j (not α). Using Eq. (13) these reduce to just one:

$$\left(\frac{\partial^2}{\partial r^2} + \frac{\partial^2}{\partial t^2}\right)\ln\left(\frac{r\rho_i}{r\rho_j}\right) = 0. \tag{15}$$

The general solution of Eq. (15) with ρ_k real and positive $(k=i,j)$ is

$$r\rho_i / r\rho_j = |h(r+it)|^2$$

with h an arbitrary analytic function. Thus we have

$$\mathrm{Re}\, q_i(r+it) = \mathrm{Re}\, q_j(r+it)\cdot|h(r+it)|^2. \tag{16}$$

Next we will solve Eq. (16). If h is a constant then

$$A^i_\mu = A^j_\mu,$$

i.e., the solution is trivial. Otherwise we can substitute Eq. (16) into Eq. (13) (twice) to get

$$\mathrm{Re}\left\{q_j + q_j'\frac{h}{h'}\right\} = 0, \tag{17}$$

$$\mathrm{Re}\left\{q_i - q_i'\frac{h}{h'}\right\} = 0. \tag{18}$$

The general analytic solution of

$$\mathrm{Re}\, g(r+it) = 0$$

in a nonempty open region is

$$g(r+it) = \text{imaginary constant}.$$

Then Eqs. (17) and (18) imply

$$q_j = d/h + ia, \tag{19a}$$

$$q_i = ch + ib, \tag{19b}$$

where a, b ($\in$Re) and c, d ($\in\mathbb{C}$) are constants. Equation (16) implies

$$c = d^*$$

so that we have derived Eqs. (6) and (7). All we need to do now is to show that Eq. (8) follows. Equations (9) and (10) imply

$$\tan\frac{\alpha}{2} = \frac{\partial}{\partial r}[\ln(r\rho_i/r\rho_j)]\Big/\frac{\partial}{\partial t}[\ln(r\rho_i\cdot r\rho_j)], \tag{20}$$

which after a little manipulation gives

$$\tan\left(\frac{\alpha}{2}\right) = -\frac{\mathrm{Re}(ch)}{\mathrm{Im}(ch)} = \tan\left(\frac{\pi}{2} + \arg(ch)\right)$$

implying

$$\alpha = \pi + 2\arg(ch), \tag{21}$$

which is Eq. (8). This proves the theorem. ■

4. EXTENSION OF P_i

Theorem 2: For any solution in $C_{r,t}$, any of the $\rho_i(r,t)$'s can be extended to all R^4 apart from isolated points on the t axis. In the neighborhood of an excluded point—at $(r,t) = (0,\beta)$—we have

$$\rho \sim \frac{a}{r^2 + (t-\beta)^2} \quad (a>0). \tag{22}$$

Proof: Consider any two patches P_i and P_j with nonzero intersection. On $P_i \cup P_j$ we have two possibilities:

(i) $A^i_\mu = A^j_\mu$.

In this case

$$\rho_i = \rho_j \times c$$

on $P_i \cup P_j$ with c a positive constant. Then ρ_i and ρ_j can be extended to all $P_i \cap P_j$ by taking

$$\rho_i \equiv \rho_j \times c$$

everywhere.

(ii) $A^i_\mu \neq A^j_\mu$.

Here part (ii) of Theorem 1 applies. On $P_i \cap P_j$ there is an analytic function $h(r+it)$ such that

$$r\rho_i = \mathrm{Re}(ch), \tag{23}$$

$$r\rho_j = \mathrm{Re}(c^*/h).$$

Now c^*/h is analytic throughout P_j [cf. Eq. (14)]. It can therefore be used to define h (hence ρ_i) throughout P_j, except for poles where

$$c^*/h = 0.$$

The zeroes of a nonconstant analytic function are necessarily isolated, so the poles of ρ_i in P_i must also be isolated (in the complex $r+it$ plane).

Suppose that ρ_i has an Nth order pole at $r+it = \alpha + i\beta$, and

$$ch \sim \frac{a+ib}{(r+it-\alpha-i\beta)^N}. \tag{24}$$

Consider $\lambda \ll 1$ and take

$$r - \alpha = \lambda\cos\theta, \tag{25a}$$

$$t - \beta = \lambda\sin\theta. \tag{25b}$$

Now θ can take all values in $-\pi \to \pi$ if $\alpha \neq 0$, and all values in $-\pi/2 \to +\pi/2$ if $\alpha = 0$. Also

$$\arg(ch) = \arg(a+ib) - N\theta$$

must only take values in $-\pi/2 \to +\pi/2$, to keep ρ_i positive. This implies

$$N = \pm 1, \quad \alpha = 0, \qquad \arg(a+ib) = 0, \quad \downarrow \quad a>0,\ b=0.$$

In the neighborhood of the singularity, therefore,

$$ch \sim \frac{a}{r+i(t-\beta)},$$

which implies

$$\rho_i \sim \frac{a}{r^2 + (t-\beta)^2}$$

as asserted. Note that points on the t axis are unique on the (r,t) half-plane, in corresponding to points on R^4.

Putting (i) and (ii) together we see that ρ_i (ρ_j) can be extended to all $P_i \cup P_j$, and hence to all R^4, apart from point singularities of the type given by Eq. (22). This proves the theorem. ■

A simple corollary of Theorem 2 is that *at most two patches suffice to define any element of* $C_{r,t}$. One patch will be

defined except at isolated points in terms of a function $h(r+it)$; the other will be defined in terms of h^{-1} and will include *all* those points.

5. A GENERAL SOLUTION FOR $C_{r,t}$

Theorem 3: In the class $C_{r,t}$ only integral topological charge is possible.

Proof: For any element of $C_{r,t}$ Theorem 2 allows us to define a gauge field A^i_μ with only isolated singularities. Uhlenbeck[5] has shown that any such field can be extended from R^4 to S^4. This ensures integral topological charges.[1] ■

Theorem 3 can also be proven as a simple corollary to

Theorem 4: For any element of $C_{r,t}$, any $\rho_i(r,t)$ in that element takes the form

$$\rho_i(r,t) = a + \sum_{i=1}^{N} \frac{a_i}{r^2 + (t-\beta_i)^2} \tag{26}$$

with $a \geqslant 0$, $a_i > 0$, $\beta_i \neq \beta_j (i \neq j)$.

Proof: Consider $\rho_i(r,t)$ where P_i has been enlarged to all R^4 apart from singularities on the t axis. Then

$$\rho_i = \frac{1}{r} \operatorname{Re} q(r+it), \tag{27}$$

where $q(r+it)$ is analytic in the $r>0$ half-plane. Also

$$\begin{aligned} &\operatorname{Re} q > 0 \quad \text{for } r > 0, \\ &\operatorname{Re} q = 0 \quad \text{for } r = 0, \end{aligned} \tag{28}$$

except at singularities. We can use the reflection principle to extend q to a meromorphic function on the entire $r+it$ complex plane, using

$$q(r+it) \equiv -[q(-r+it)]^*. \tag{29}$$

Note that q is analytic for $r<0$, and

$$\operatorname{Re} q < 0 \quad \text{for } r < 0. \tag{30}$$

We have already found the possible singularities q can have (Theorem 2). The most general *analytic* function satisfying Eqs. (30) and (32) is

$$q = a(r+it) + ib \quad (a,b \in \operatorname{Re}, a > 0). \tag{31}$$

To see this take

$$\bar{q}(r+iz) = q(r+iz) - q(0), \tag{32}$$

which also obeys Eqs. (28) and (30) since $\bar{q}(0)$ is pure imaginary. Consider the change in argument of q around a circle of radius R centered at the origin. This is, $2\pi \times$ number of zeroes of $\bar{q}$ in the circle. The latter number is at least one, and must be exactly one $\forall R$ to satisfy Eqs. (28) and (30). Defining

$$g(z) \equiv \bar{q}/z \tag{33}$$

($z = r+it$) we see that g is analytic in $\mathbb{C}$, and has no zeroes. If $g(z)$ were a polynomial this would imply that g was a positive real constant, proving Eq. (31). To complete the proof we must show that g cannot be transcendental.[6] If it were, then the above arguments applied to

$$q \to q + \alpha z \quad (\alpha > 0)$$

would show that g cannot equal $-\alpha$ for all α real and positive. This contradicts Weierstrass' theorem, so g cannot be transcendental.

Combining Eq. (31) with Theorem 2 we find the most general solution

$$q = a(r+iz) + i\beta + \sum_{i=1}^{N} \frac{a_i}{r + i(t-\beta_i)}. \tag{34}$$

This proves the theorem. Note that

$$\left.\begin{aligned} \nu(\rho) &= N \ (a \neq 0) \\ &= N-1 \ (a = 0) \end{aligned}\right\}, \tag{35}$$

proving Theorem 3. For an arbitrary choice of the t axis, Eq. (34) recovers the full set of solutions of Jackiw *et al.*[7]

6. COMMENT

Without assuming compactification, we have proved that in the class of gauge-field structures $C_{r,t}$ the topological charge must be integral. A general proof would doubtless require more sophisticated techniques. Nevertheless, our central concept could be the basis of such a proof. It is the use of analytic techniques to show that only isolated singularities can occur.

In conclusion, we note that the compactifiable nature of the solutions in $C_{r,t}$ follows directly from the finiteness of the action. This implies that there are only finitely many singularities. On any patch P_i these must lie in a bounded interval ($< R$) beyond which everything is smooth, and can be extended to the point at infinity. With infinite action we could have an infinite string of singularities, and such a solution would not be compactifiable.

ACKNOWLEDGMENTS

I would like to thank Dr. J. M. Anderson (UCL) for his assistance in the proof of Theorem 4, and the United Kingdom Science and Engineering Research Council for financial support.

[1] M. F. Atiyah and R. S. Ward, Commun. Math. Phys. **55**, 117 (1977).
[2] R. J. Crewther, Phys. Lett. **70B**, 349 (1977).
[3] R. Stacey, Z. Phys. C **10**, 149 (1981).
[4] P. Forgacs, Z. Horvath, and L. Palla, Phys. Rev. Lett. **46**, 392 (1981).
[5] K. Uhlenbeck, Bull. Am. Math. Soc. **1**, 579 (1979).
[6] I would like to thank Dr. J. M. Anderson for his assistance in the completion of this proof.
[7] R. Jackiw, C. Nohl and C. Rebbi, Phys. Rev. D **15**, 1642 (1977).

Irreducible representations of SU(m/n)

J.-P. Hurni and B. Morel[a)]
Université Genève, Département de Physique Théorique, 1211 Genève 4, Switzerland

(Received 27 May 1981; accepted for publication 4 September 1981)

Typical and atypical representations of SU(m/n) are worked out, in an explicit and pedestrian way. Examples are given.

PACS numbers: 11.30.Pb

I. INTRODUCTION

Lie superalgebras (Refs. 1,2) are progressively becoming very used in physics: as algebras for super unification (Ref. 3), in nuclear physics (Ref. 4) and in supergravity (Ref. 5).

The irreducible representations of Lie superalgebras have already been studied to a certain extent (Refs. 6 and 7), but we feel that there is a need for some clarification in this problem. We hope to do it in this paper where we study the typical and atypical representations of SU(m/n).

Atypical representations are the ones which have been least studied, though they are obviously of importance: all the adjoint representations of SU(m/n) for $mn > 2$ are atypical; and they could be of importance also in physics.

In this work we rely heavily on the techniques of Ref. 6, perhaps the most powerful one. We have a pedestrian approach which we feel hopefully is pedagogical, simple, and operational.

In Sec. 2 we present the properties of SU(m/n) directly relevant to us. In Sec. 3 we explain simply how a typical representation is built. In Sec. 4 we study the atypical case; and in Sec. 5 we give several examples of SU(1/2), SU(2/2), SU(1/8), and SU(2/3). In this paper we state only the results and properties of unitary superalgebra relevant for us. For basic definitions and additional information we refer the reader to Refs. 1, 2, and 6.

2. SU(m/n)$\equiv A(m-1/n-1)$

The bosonic or even part $G_{\bar 0}$ of the Lie superalgebra SU(m/n) is $G_{\bar 0} = \mathrm{SU}(m) \otimes \mathrm{SU}(n) \otimes \mathrm{U}(1)$. [When $m = n$, the even part of $A(m-1/m-1)$ is $G_{\bar 0} = \mathrm{SU}(m) \otimes \mathrm{SU}(n)$, but in fact $A(m-1/m-1) \not\equiv \mathrm{SU}(m/m)$, cf. ultra.]

The fermionic or odd part of SU(m/n) is made of two parts G_{-1} and G_1 corresponding, respectively, to the irreducible representations $(m, \bar n)$ and $(\bar m, n)$ of $\mathrm{SU}(m) \otimes \mathrm{SU}(n)$.

The algebra SU(m/n) is then of dimension $(m^2-1)+(n^2-1)+1+2mn = (m+n)^2-1$, and its rank is $m+n-1$. $\mathrm{SU}(m/n) = G_{-1} \otimes G_{\bar 0} \otimes G_1$ satisfies

$$[G_{\bar 0}, G_{\bar 0}] = G_{\bar 0},$$
$$[G_{\bar 0}, G_{\pm 1}] \subseteq G_{\pm 1},$$
$$\{G_1, G_{-1}\} \subseteq G_{\bar 0},$$
$$\{G_1, G_1\} = \{G_{-1}, G_{-1}\} = 0,$$

and can be decomposed into:

(a) A Cartan subalgebra $\{h_i\}$, whose dimension is $m+n-1$. (In the case $m = n$, there is a linear combination of these $2m-1$ elements which gives an element I in the center of the algebra, so SU(m/n) is not simple and $A(m-1/m-1)$ is defined as the coset of the original superalgebra by this ideal:

$$A(m-1/m-1) \equiv \frac{S(m|m)}{kI}, \quad k \in \mathbb{C}.$$

We take the basis in which the simple roots are eigenvectors of the h_i as we will see later.

(b) The other generators are divided into positive/negative, even/odd roots:

—the positive/negative even roots $a_i^{\pm}$, $c_j^{\pm}$ correspond to the positive/negative roots of SU(m), SU(n), respectively.

—the positive/negative odd roots are called $b_j^{i\pm}$ and belong, respectively, to $(m/\bar n)$ and $(\bar m/n)$ of $\mathrm{SU}(m) \otimes \mathrm{SU}(n)$. One can define a system of simple roots in such a way that there is one simple positive odd root β^+, with $m-1$ and $n-1$ simple positive even roots of SU(m) and SU(n), α_1^+, $\ldots, \alpha_{m-1}^+$ and $\gamma_{m+1}^+, \ldots, \gamma_{m+n-1}^+$, respectively. (In the same way one can define simple negative roots.)

It is worth emphasizing that in terms of representation of $\mathrm{SU}(m) \otimes \mathrm{SU}(n)$, β^- corresponds to the highest weight of $(\bar m, n)$ and β^+ to minus that vector in the root space. [β^+ belongs to $(m, \bar n)$ but is not the highest weight.]

In terms of these simple roots, the algebra looks like (with the convention $\beta^+ = \alpha_m^+$; $\gamma_{m+j}^+ = \alpha_{m+j}^+$)

$$[\alpha_i^+, \alpha_j^-] = \delta_{ij} h_i, \quad i,j = 1 \cdots m-1;\ m+1 \cdots m+n-1,$$
$$\{\alpha_m^+, \alpha_m^-\} = h_m$$
$$[\alpha_m^+, \alpha_i^-] = [\alpha_i^+, \alpha_m^-] = 0, \quad i \neq m,$$
$$[h_i, \alpha_j^{\pm}] = \pm a_{ij} \alpha_j^{\pm} \quad \forall i,j,$$

where the a_{ij} are the elements of the Cartan matrix associated to the algebra

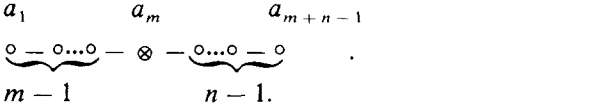
$$A = (a_{ij}) = \left(\begin{array}{c|c|cc|cc} \overbrace{\begin{array}{ccc} 2 & -1 & 0 \\ -1 & 2 & \\ & \ddots & -1 \\ & -1 & 2 \end{array}}^{m-1} & & & & & \\ \hline & & 0 & & & \\ & & -1 & 2 & -1 & 0 \\ \hline & & 0 & -1 & 0 & 1 \\ & & & 0 & -1 & 2 \\ & & & & -1 & \ddots \\ & & & & & 2 \end{array}\right) \begin{array}{l} \} m-1 \\ \\ \\ \\ \\ \} n-1 \end{array}$$

$$\underbrace{\qquad}_{n-1}$$

[a)]Supported in part by the Swiss National Science Foundation.

 0022-2488/83/010157-07$02.50

A contains the same information as the following Dynkin diagram: ($a_{mm} = 0$ implies $a_{m,m+1} = +1$; $a_{m,m-1} = -1$)



The a_i characterize the highest weight Λ of a given representation. a_m corresponds to the odd simple root, and to the zero in the diagonal of the Cartan matrix.

The Cartan subalgebra is defined in such a way to measure the projections of Λ along the simple positive roots α_i^+ giving the a_i. ($a_i \in \mathbb{Z}^+ \cup \{0\}$, $i \neq m$, $a_m \in \mathbb{C}$.)

Nonsimple roots are obtained by commutations between simple roots as in customary Lie algebra. These commutations are made more perceptible with the notation:

$$\beta^{\pm} = b_m^{m\pm}, \quad \alpha_{j<m}^{\pm} = \alpha_j^{\pm}, \quad \alpha_{j>m}^{\pm} = \gamma_j^{\pm}.$$

In the case of the odd negative roots it is worth while to be precise:

$$[b_l^{k-}, \alpha_i^-] = \delta_{i,k-1} b_l^{(k-1)-}, \tag{2.1}$$

$$[b_l^{k-}, \gamma_j^-] = -\delta_{j,l+1} b_{l+1}^{k-}. \tag{2.2}$$

This leads to a natural ordering of the odd roots $b_j^{i\pm}$ according to the values of the indices i (in decreasing order) and j (in increasing order).

The following relations are also easy to verify and useful:

$$[h_i, b_l^{k-}] = \begin{cases} 0 & \text{if } k > i+1 \text{ or } l < i, \\ & \quad k \leqslant i-1 \text{ or } l \geqslant i+1, \\ b_l^{k-} & \text{if } k = i+1 \text{ or } l = i-1, \\ -b_l^{k-} & \text{if } k = i \text{ or } l = i, \end{cases} \tag{2.3}$$

$$[h_m, b_l^{k-}] = \begin{cases} 0 & \text{if } k = l = m \text{ or } k < m,\ l > m, \\ b_l^m & \text{if } k = m,\ l \neq m, \\ -b_m^k & \text{if } l = m,\ k \neq m, \end{cases} \tag{2.4}$$

$$[b_l^{k-}, \alpha_i^+] = +\delta_{ik} b_l^{(k+1)-}, \tag{2.5}$$

$$[b_l^{k-}, \gamma_j^+] = -\delta_{jl} b_{l-1}^{k-},$$

$$\{b_l^{k+}, b_l^{k-}\} = h_m + \sum_{i=k}^{m-1} h_i - \sum_{j=m+1}^{l} j_j, \tag{2.6}$$

$$\{b_l^{k+}, b_l^{k'-}\} = [\alpha_{k+1}^+ \cdots [\alpha_{k'-1}^+, \alpha_{k'}^+] \cdots], \quad k < k', \\ = [\alpha_{k'+1}^- \cdots [\alpha_{k-1}^-, \alpha_k^-] \cdots], \quad k > k', \tag{2.7}$$

$$\{b_l^{k+}, b_{l'}^{k-}\} = [\gamma_{l-1}^+ \cdots [\gamma_{l'+1}^+, \gamma_{l'}^+] \cdots], \quad l > l', \\ = [\gamma_{l'-1}^- \cdots [\gamma_{l+1}^-, \gamma_l^-] \cdots], \quad l < l', \tag{2.8}$$

$$\{b_l^{k+}, b_{l'}^{k'-}\} = 0 \quad \text{if} \quad \begin{cases} k \neq k' \\ l \neq l' \end{cases}. \tag{2.9}$$

3. TYPICAL REPRESENTATIONS

In this section we describe the principle of construction of irreducible representations of SU(m/n). (This section is completely inspired from Kac, Ref. 6.)

(1) One chooses a highest weight Λ corresponding to a set of a_i appearing on the Dynkin diagram. Λ belong to the multiplet of SU(m) ⊗ SU(n) characterized by the $a_{i\neq m}$; a_m will characterize its "typicality." In this section we will suppose that a_m is a complex number or any number different from the one appearing in the next section. The representation will be called typical, that means that all the possible multiplets of SU(m) ⊗ SU(n) which could appear will appear.

(2) By definition of the highest weight,

$$\alpha_i^+ \Lambda = \gamma_j^+ \Lambda = \beta^+ \Lambda = b_j^{i+} \Lambda = 0.$$

The representation is exhibited by repeated applications of the negatives roots:

—the negative even roots α_i^- and γ_j^- make us migrate inside multiplets of SU(m) ⊗ SU(n)

—the odd roots b_j^{i-} [which are in a $(\bar{m}, n)$ of SU(m) ⊗ SU(n)] make us change the multiplet.

Since $\{b_j^{i-}, b_l^{k-}\} = 0 \ \forall i,j,k,l$, one can apply on Λ at most an antisymmetric combination of $m \times n$ odd roots. (We will say that the representation has a "ground floor" and $m \times n$ floors.)

The multiplicity of a typical representation is very easy to compute; it is the multiplicity of Λ times 2^{mn} [the multiplicity associated with the antisymmetric combinations of b_j^{i-} corresponds exactly to the binomial coefficients of $(1+1)^{mn}$]. In other words, the multiplicity can be written:

$$\dim V(\Lambda) = 2^{mn} \prod_{1 \leqslant i < j \leqslant m-1} \frac{a_i + \cdots + a_j + j - i + 1}{j - i + 1} \\ \times \prod_{m+1 \leqslant i < j \leqslant m+n} \frac{a_i \cdots + a_{j+j-i+1}}{j - i + 1}.$$

A few remarks are in order which will be useful in the next section:

—The zeroth floor corresponds to the SU(m) ⊗ SU(n) irreducible representation of which Λ is the highest weight. The first floor corresponds to the product $\Lambda \otimes (\bar{m}, n)$; the second floor to $\Lambda \otimes [(\bar{m}, n) \otimes (\bar{m}, n)]_A$; and so on···. So each floor, from the first, corresponds in general to a reducible representation of SU(m) ⊗ SU(n), which one has to disentangle; there are then "highest-highest weights" at each floor and sometimes also "lower-highest weights" for the other representations.

—In terms of Young diagrams, the antisymmetrized product of k times $(\bar{m}, n)$ is the direct sum of all possible pairs of Young diagrams made of k boxes with a maximum of m rows and n columns with respect to SU(n), and the contragradient of it is transposed for SU(m).

Proposition 3.1: The highest weight of the reducible representations at each floor (referred to as highest weight) is obtained by applying the negative odd roots b_j^{i-} in their "natural order."

Proof: A "naturally ordered" product of $\prod_{ij} b_j^{i-}$ is by definition such that it commutes with any positive even roots. That means that in the product, at the right of any b_j^{i-} there is either b_j^{i+1-} or b_{j+1}^{i-} or both. Then it is obvious that

$$\alpha_k^+ \prod_{ij} b_j^{i-} \Lambda = 0$$

if Λ is a highest weight.

To extract the SU(m) ⊗ SU(n) representations hidden

behind the highest-highest weight, one proceeds in the customary way: the lower-highest weights correspond to suitable linear combinations of the different operators leading to that weight; the orthogonal combinations corresponding to members of bigger representations.

If $V_{\bar{0}}$ and $V_{\bar{1}}$ denote the "bosonic," respectively, "fermionic" part of the representation, then dim $V_{\bar{0}}$ = dim $V_{\bar{1}}$ and either $V_{\bar{0}}$ or $V_{\bar{1}}$ is made of the elements of the even or odd floors by the following property of the binomial coefficients:

$$\sum_{\text{odd } n}\binom{N}{n}=\sum_{\text{even } n}\binom{N}{n}.$$

Given the biggest reduced representation and any little one, let λ_i be the number of times the even root α_i^- has to be applied anywhere in the product of odd roots in order to pass from the highest weight to the members of the linear combination of the lower-highest weight.

Let $\mathbf{a}_0$ be the weight vector of the biggest representation and $\mathbf{a}$ the one of the little; then $\boldsymbol{\lambda}=(a_{ij})^{-1}(\mathbf{a}_0-\mathbf{a})$, where (a_{ij}) is the Cartan matrix of SU(m) ⊗ SU(n). For SU(n),

$$(a_{ij})=\begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & \ddots & \ddots & \\ & & & \ddots & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix};$$

$$(a_{ij})^{-1}=\frac{1}{n}\begin{pmatrix} (n-1) & (n-2) & \cdots & 3 & 2 & 1 \\ (n-2) & & & & & 2 \\ \vdots & & & & & 3 \\ 3 & & & & & \vdots \\ 2 & & & & & (n-2) \\ 1 & 2 & 3 & \cdots & (n-2) & (n-1) \end{pmatrix}.$$

4. ATYPICAL REPRESENTATIONS

A pathology associated with irreducible representations of superalgebras is the notion of atypicality: it corresponds to the decoupling of some part of the representation appearing in the previous section; this phenomenon is associated with some values of a_m characterizing the highest weight Λ of the whole representation.

A. Decoupling of the "higher-highest weights"

We know from Proposition 3.1 that they correspond to $T_l^{k-}\Lambda=\Pi_{i=m-k}^{m}\Pi_{j=m}^{m+l}b_j^{i-}\Lambda$, where the b_j^{i-} appear in their natural order.

The decoupling of these "higher-highest" weights takes place when the operator T_l^{k-} is not invertible or when $T_l^{k+}T_l^{k-}\Lambda=P_l^k(h_i)|\Lambda\rangle=0$.

Proposition 4.1:

$$T_l^{k+}T_l^{k-}|\Lambda\rangle=\prod_{\substack{m-k\leqslant i\leqslant m\\ m\leqslant j\leqslant l+m}}\left(h_m-\sum_{t=m+1}^{j}h_t+\sum_{t=i}^{m-1}h_t+2m-i-j\right)|\Lambda\rangle$$

and (i,j) ordered.

Proof: There are two ways to obtain that result:

(a) by recursion on k and l using $T_l^{k-}|\Lambda\rangle=b_{m+l}^{(m-k)-}T_l^{(k-1)-}|\Lambda\rangle$, $T_0^{0-}|\Lambda=b_m^{m-}|\Lambda\rangle$ and using Eq. (2.6):

$$\begin{aligned}T_l^{k+}T_l^{k-}|\Lambda\rangle&=T_l^{(k-1)+}\left(h_m+\sum_{t=m-k+1}^{m-1}h_t-\sum_{t=m+1}^{m+l}h_t\right)T_l^{(k-1)-}|\Lambda\rangle\\&=\left(\Delta_l^{k-1}+h_m+\sum_{t=m-k+1}^{m-1}h_t-\sum_{t=m+1}^{m+l}h_t\right)T_l^{(k-1)+}T_l^{(k-1)-}|\Lambda\rangle.\end{aligned}\tag{4.1}$$

Using Eqs. (2.4) and (2.5),

$$\begin{aligned}\Delta_l^{k-1}&=N(i,m)-N(j,m)+\sum_{t=m-k-1}^{m-1}(N(t,i-1)-N(t,i))\\&\quad-\sum_{t=m+1}^{m+l}(N(t,j+1)-N(t,j)).\end{aligned}\tag{4.2}$$

$N(a,b)$ is the number of times that a is equal to b in the product $T_l^{k-}=\Pi b_j^{i-}$. In this case, $\Delta_l^k=k-l=2m-i-j$ ($i=m-k$ and $j=m+l$ are the natural values of the indices).

(b) Another way to get formula is to make a recursion on k and l in SU(k/l), noticing that when $k=m, l=n$, it corresponds exactly to the formula obtained by Kac (Ref. 6). These conditions of atypicality are expressed as

$$\begin{aligned}a_m&=\sum_{t=m+1}^{j}a_t-\sum_{t=i}^{m-1}a_t-2m+i+j,\\&\quad m-k\leqslant i\leqslant m\leqslant j\leqslant m+l.\end{aligned}\tag{4.3}$$

B. The decoupling of the highest-highest weight on a floor does not mean that the whole floor decouples

To see how this happens we need three more formulae:

Proposition 4.2: Let $\Lambda_l^k=T_l^{k-}\Lambda$ be any highest weight; then $\alpha_\tau\Lambda_l^k$ has at least the same decoupling conditions as Λ_l^k.

Proof: Let us take $m-k\leqslant i\leqslant m$. This comes from: $T_l^{k+}\alpha_i^+\alpha_i^-T_l^{k-}\Lambda=(h_i+d_l^k)P_l^k(h_i)\Lambda$ where using (2.3)

$$d_l^k=\sum_{m-k\leqslant t\leqslant m}(N(t,i-1)-N(t,i)).$$

So the zeroes of $P_l^k(h_i)$ remain decoupling conditions.

This proposition implies essentially that the members of a SU(m) ⊗ SU(n) irreducible multiplet have the same decoupling conditions as their highest weight.

Proposition 4.3: $\Lambda_l^k=T_l^{k-}\Lambda$ and $X_l^k=T_l^k\alpha_\tau\Lambda$ have not necessarily the same decoupling conditions.

Proof:

$$\alpha_i^+T_l^{k+}T_l^{k-}\alpha_i^-\Lambda=P_l^k(h_j-a_{ij})h_i\Lambda,\tag{4.4}$$

but

$$T_l^{k+}\alpha_i^+T_l^{k-}\alpha_i^-\Lambda=P_l^k(h_j)h_i\Lambda,\tag{4.5}$$

and

$$\alpha_i^+T_l^{k+}\alpha_i^-T_l^{k-}\Lambda=h_iP_l^k(h_j)\Lambda.\tag{4.6}$$

Q.E.D.

Equation (4.4) implies others conditions of decoupling for some values of i.

Proposition 4.4: If $\Lambda_I^k = T_I^{k-}\Lambda$ is a "highest-highest weight," there are only two meaningful ways to insert an even root α_i^-; it is either $\alpha_i^- T_I^{k-}\Lambda$ or $T_I^{k-}\alpha_i^-\Lambda$.

Proof: From Propositions 4.2 and 4.3 we know that they lead to inequivalent decoupling conditions.

If one inserts α_i^- anywhere inside $T_I^{k-} = \Pi b_j^{i-}$, using Eq. (2.1) one can let it migrate to either end of the product. If we let it towards the right end (where lie the lower in order b_j^{i-}), because $(b_j^{i-})^2 = 0$ and because Λ_I^k is a highest-highest weight, that will not introduce any new terms.

C. Consequences on the atypical representations

From Proposition 4.2 we know that an atypical representation is made of irreducible multiplets of SU(m) ⊗ SU(n).

In the case where the highest weight Λ of the superrepresentation is a singlet of SU(m) ⊗ SU(n), Propositions 4.3 and 4.4 are irrelevant. But in the case where Λ is not a singlet of SU(m) ⊗ SU(n), lower-highest weight of smaller representations of SU(m) ⊗ SU(n) are hidden behind the representation of the highest-highest weight.

From Proposition 4.3 we know that they have not the same decoupling conditions as the highest-highest weights. The corresponding conditions can be deduced from Eq. (4.4) and Eq. (4.3). From Proposition 4.4 we know that Eq. (4.4) is the only new set of conditions to consider. It is important to realize that these new conditions are not new typicality conditions!

They really are decoupling conditions if they coincide with the conditions given by Eq. (4.3). This occurs because of Eq. (4.5) which tells us that it is also possible from $T_I^{k-}\alpha_i^-\Lambda$ to come back to Λ by applying $T_I^{k+}\alpha_i^+$ on $T_I^{k-}\alpha_i^-\Lambda$.

Equation (4.6) means that there is no reduction in the number of decoupling conditions for members of the multiplet of SU(m) ⊗ SU(n) whose highest weight is the highest-highest weight Λ_I^k. In the next section we see examples on how things work.

5. SOME EXAMPLES

Example 1: SU(1/2): $\overset{a_1}{\otimes}\!\!-\!\!-\overset{a_2}{\circ}$

The two negative odd roots are $\beta^- = \alpha_1^-$ and $\beta_1^- = [\beta^-, \alpha_2^-]$, α_2^- being the simple negative root associated with the bosonic subalgebra SU(2). The Cartan subalgebra is made of $h_1 = \{\beta^+, \beta^-\}$ and $h_2 = [\alpha_2^+, \alpha_2]$; the Cartan matrix is $\begin{pmatrix} 0 & \frac{1}{2} \\ -1 & 2 \end{pmatrix}$ and its elements a_{ij} appear in: $[h_i, \alpha_j^\pm] = \pm a_{ij}\alpha_j^\pm$, $i,j = 1,2$.

A representation is characterized by the highest weight Λ whose behavior under SU(2) is given by a_2. The corresponding typical representation can be written as $\Lambda\cdot(1 + 2 + 1)$ in terms of SU(2); in particular,

—If Λ is a singlet ($a_2 = 0$) we obtain a $(1 + 2 + 1)$.

—If Λ is a doublet ($a_2 = 1$) we have $2(1 + 2 + 1) = (2_1 + (1 + 3) + 2_2)$.

The highest weight of the 3 corresponds to $\beta^-\Lambda$ which is decoupled by $\beta^+\beta^-\Lambda = h_1\Lambda = a_1\Lambda = 0$, i.e., if $a_1 = 0$.

The highest weight of the 2_2 corresponds to $\beta_1^-\beta^-\Lambda$, and using Eqs. (2.6) and (2.3) one gets $\beta^+\beta_1^+\beta_1^-\beta\Lambda = (h_1 - h_2 - 1)h_1\Lambda$; the decoupling conditions are $a_1 = 0$ and $a_2 + 1$, i.e., 2 in the present case. These are exactly the decoupling conditions we would get from (4.3).

Because $\beta^-\alpha_2^-\Lambda$ does not decouple when $a_1 = 0$, $\alpha_2^+\beta^+\beta\alpha_2\Lambda = (h_1 - 1)h_2\Lambda$ [which is equivalent to Proposition (4.3)], we see that the 1 does not decouple for $a_1 = 0$ or $a_1 = 2$ so the representation is, if

$a_1 = 0$ $\quad 2 + 1$ of SU(2),

$a_1 = 2$ $\quad 2 + 1 + 3$ of SU(2),

$a_1 \neq 0,2$ $\quad 2 + (1 + 3) + 2$

(it corresponds to the only case where the adjoint is a typical representation).

Example 2: SU(2/2)

$\overset{a_1}{\circ}\!\!-\!\!\otimes\!\!-\!\!\overset{a_2}{\circ}\!\!-\!\!\overset{a_3}{\circ}$

The bosonic subalgebra is SU(2) ⊗ SU(2) ⊗ U(1). [The bosonic subalgebra of A(1,1) is SU(2) ⊗ SU(2) only, but our method applies for SU(m/n) $\forall m, n$; not for A($m - 1, n - 1$) when $m = n$.]

The Cartan matrix is

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}$$

Let $\alpha_1^\pm$, $\alpha_3^\pm = \gamma_3^\pm$ denote the simple roots, respectively, of the two SU(2)'s. The simple negative root β^- is b_2^{2-} in the notation of Sec. 2; other negative roots are $b_2^{1-} = [b_2^{2-}, \alpha_1^-]$; $b_3^{2-} = [b_2^{2-}, \alpha_3^-]$ and $b_3^{1-} = [b_2^{1-}, \alpha_3^-] = [b_3^{2-}, \alpha_1^-]$.

A representation of SU(2/2) with highest weight Λ in terms of representation SU(2) ⊗ SU(2) is part or totality of

$$\Lambda\cdot[(1,1) + (2,2) + [(3,1) + (1,3)] + (2,2) + (1,1)].$$

If Λ is not a singlet of SU(2) ⊗ SU(2), in general each floor will correspond to a reducible representation. The higher-highest weight will correspond to

$\Lambda^{(1)} = b_2^{2-}\Lambda$,	for the first floor [where we have (2,2)],
$\Lambda^{(2)} = b_2^{1-}b_2^{2-}\Lambda$	$\Lambda^{(2)}$ associated with (1,3),
$\Lambda^{(3)} = b_3^{2-}b_2^{2-}\Lambda$	for the second floor $\Lambda^{(3)}$ associated with (3,1),
$\Lambda^{(4)} = b_3^{2-}b_2^{1-}b_2^{2-}\Lambda$	for the third floor,
$\Lambda^{(5)} = b_3^{1-}b_3^{2-}b_2^{1-}b_2^{2-}\Lambda$	for the fourth floor,

leading to the decoupling conditions

$\Lambda^{(1)}$: $a_2 = 0$,

$\Lambda^{(1)}$: $a_2 = 0$, $a_2 = -(a_1 + 1)$,

$\Lambda^{(3)}$: $a_2 = 0$, $a_2 = a_3 + 1$,

$\Lambda^{(4)}$: $a_2 = 0$, $a_2 = -(a_1 + 1)$ or $a_2 = +a_3 + 1$,

$\Lambda^{(5)}$: $a_2 = 0$, $-(a_1 + 1)$, $(a_3 + 1)$, $a_3 - a_1$,

which are exactly those predicted by Eq. (4.3).

Suppose Λ is (1,2) of SU(2) ⊗ SU(2), namely that the representation corresponded to a:

$\overset{0}{\circ}\!\!-\!\!\otimes\!\!-\!\!\overset{a_1}{\circ}\!\!-\!\!\overset{1}{\circ}$

We will just look at what would happen to the first floor. It will be reducible and made of

$$(1,2)\times(2,2) = (2,1) + (2,3).$$

$b_2^{2-}\Lambda$ corresponds to the highest weight of the (2,3). $\chi_1 = \gamma_3^- b_2^{2-}\Lambda$ and $\chi_2 = b_2^2\gamma_3^-\Lambda$ are independent; the (2,1) corresponds to the linear combination $(\chi_1 - \frac{1}{2}\chi_2)$; and the orthogonal combination is a member of the (2,3).

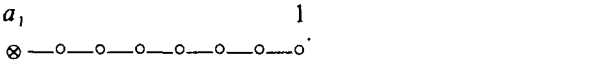
There are two ways from χ_1 and χ_2 to get back to the highest weight Λ of the representations: by applying $\gamma_3^+ b_2^{2+}$ and $b_2^{2+}\gamma_3^+$, all the combinations lead to $a_2 = 0$ as decoupling condition (for example $b_2^{2+}\gamma_3^+\gamma_3^- b_2^{2-}\Lambda$) except for one, $\gamma_3^+ b_2^{2+} b_2^{2-}\gamma_3^-\Lambda$, which leads to $a_2 = 1$. That means that when $a_2 = 0$ the (2,1) is not decoupled; it is certainly coupled as well when $a_s \neq 0$ so it is always part of the irreducible representation.

Let us define I_0 such that $\chi_1 - \frac{1}{2}\chi_2 = I_0\Lambda$. If we call "norm" of the (2,1) the expression $I_0^+ I_0\Lambda$, when $a_2 = 0$, this norm is zero. (This is connected with the decoupling of part of the next floor for that value of a_2.)

It certainly does not mean that the (2,1) is decoupled, though all the states which decouple have a zero "norm." (This "norm" is by no means a norm in fact; in particular it is not necessarily positive, and could be complex.)

Example 3: SU(1/8).

In Ref. 8 an attempt has been made toward superunification by studying the spectrum of particles for O(8) — extended supergravity. In fact, these states fall into the representation of SU(1/8); here we study the relevant representation of SU(1/8) and see whether the "trace condition" used in Ref. 8 to decouple some of the states from the physical spectrum correspond to use of an atypical representation; we find that it is not so. The bosonic subgroup is SU(8) ⊗ U(1). We are interested in studying the representation where the highest weight is in a $\bar{8}$, corresponding to



The odd negative roots b_j^- form an 8 of SU(8); therefore the corresponding typical representations (according to Sec. 3) will correspond to a $2^8\times 8$ dimensional supermultiplet:

$$\begin{aligned}\bar{8}\ &\times(1 + 28 + 56 + 70 + \overline{56} + \overline{28} + \bar{8} + 1)\\ &= \bar{8} + (63 + 1) + (\overline{216} + 8) + (420 + 28) + (504 + 56)\\ &\quad + (378 + 70) + (168 + \overline{56}) + (\overline{36} + \overline{28}) + \bar{8}.\end{aligned}$$

From Proposition 3.1,

$$\begin{aligned}\Lambda_{63} &= b_1^-\Lambda_{\bar{8}},\\ \Lambda_{216} &= b_2^- b_1^-\Lambda_{\bar{8}},\\ \Lambda_{420} &= b_3^- b_8^- b_1^-\Lambda_{\bar{8}},\\ \Lambda_{504} &= b_4^- b_3^- b_2^- b_1^-\Lambda_{\bar{8}},\\ \Lambda_{378} &= b_5^- b_4^- b_3^- b_2^- b_1^-\Lambda_{\bar{8}},\\ \Lambda_{168} &= b_6^- b_5^- b_4^- b_3^- b_2^- b_1^-\Lambda_{\bar{8}},\\ \Lambda_{36} &= b_7^- b_6^- b_5^- b_4^- b_3^- b_2^- b_1^-\Lambda_{\bar{8}},\\ \Lambda_{\bar{8}} &= b_8^- b_7^- b_6^- b_5^- b_4^- b_3^- b_2^- b_1^-\Lambda_{\bar{8}},\end{aligned}$$

where $b_j^- = [b_{j-1}^-, \gamma_j^-]$.

The decoupling of these highest-highest weight is given [Eq. (4.3)] by the zeroes of

$$\prod_{j=1}^{k}\left(h_1 - \sum_{t=2}^{j} h_t + 1 - j\right)\Lambda,$$

$k = 1$ corresponding to Λ_{63}, $k = 2$ to Λ_{216}, etc., and the conditions for the decoupling of the corresponding states are

$$a_5 = 0,1,2,3,4,5,6,8.$$

The conditions of decoupling of the lower-highest weight can be deduced by remarking

$$\begin{aligned}\Lambda_1 &= \gamma_2^-\gamma_3^-\gamma_4^-\gamma_5^-\gamma_6^-\gamma_7^-\gamma_8^-\Lambda_{63},\\ \Lambda_8 &= \gamma_3^-\gamma_4^-\gamma_5^-\gamma_6^-\gamma_7^-\gamma_8^-\Lambda_{216},\\ \Lambda_{28} &= \gamma_4^-\gamma_5^-\gamma_6^-\gamma_7^-\gamma_8^-\Lambda_{420},\\ \Lambda_{56} &= \gamma_5^-\gamma_6^-\gamma_7^-\gamma_8^-\Lambda_{504},\\ \Lambda_{70} &= \gamma_6^-\gamma_7^-\gamma_8^-\Lambda_{378},\\ \Lambda_{56} &= \gamma_7^-\gamma_8^-\Lambda_{168},\\ \Lambda_{28} &= \gamma_8^-\Lambda_{36}.\end{aligned}$$

It is a bit messy but straightforward to find the following result:

a_2	irreducible representation
typical	$\bar{8} + (63 + 1) + (216 + 8) + (420 + 28) + (504 + 56) + (378 + 70) + (168 + 56) + (26 + 28) + \bar{8}$
$= 8$	$\bar{8} + (63 + 1) + (216 + 8) + (420 + 28) + (504 + 56) + (378 + 70) + (168 + \overline{56}) + (36 + 28)$
$= 6$	$\bar{8} + (63 + 1) + (216 + 8) + (420 + 28) + (504 + 56) + (378 + 70) + (168 + 56) + 28$
$= 5$	$\bar{8} + (63 + 1) + (216 + 8) + (420 + 28) + (504 + 56) + (378 + 70) + 56$
$= 4$	$\bar{8} + (63 + 1) + (216 + 8) + (420 + 28) + (504 + 56) + 70$
$= 3$	$\bar{8} + (63 + 1) + (216 + 8) + (420 + 28) + 56$
$= 2$	$\bar{8} + (63 + 1) + (216 + 8) + 28$
$= 1$	$\bar{8} + (63 + 1) + 8$ ←adjoint of SU(1/8)
$= 0$	$\bar{8} + 1$

Notice in particular that one does not get the trace condition of Ref. 8.

Notice also that at each floor, in general the conditions of decoupling of the lower-highest weight are different from those of the highest-highest weight.

Example 4: SU(2/3).

Finally we look at SU(2/3) with highest weight taken in a $(2, \bar{3})$; namely the following representation:

$$\begin{array}{ccccccc} 1 & & a_2 & & 1 & & \\ \circ & - & \otimes & - & \circ & - & \circ \end{array}$$

The corresponding typical representation would be

$$\begin{aligned}(2,\bar3)\times\{&(1,1)+(2,3)+[(1,6)+(3,\bar3)]+[(2,8)+(4,1)]\\&+[(1,\bar6)+(3,3)]+(2,\bar3)+(1,1)\}\\&=(2,\bar3)+(1+3,\ 1+8)+[(2,\ 3+15)\\&\quad+(2+4,\ \bar6+3)]+[(1+3,\ \overline{15}+6+\bar3)\\&\quad+(5+3,\ \bar3)]+[(2,\ \overline{10}+8)+(4+2,\ 1+8)]\\&\quad+(1+3,\ \bar6+3)+(2,\bar3)\\&=\underline{(2,\bar3)}+[\underline{(1,1)}+\underline{(3,1)}+\underline{(1,8)}+(3,8)]+[\underline{(2,3)}\\&\quad+\ +(2,15)+\cdots].\end{aligned}$$

(We underlined the atypical adjoint representation)

$$\begin{array}{ccccccc} 1 & & 0 & & 0 & & 1 \\ \circ & - & \otimes & - & \circ & - & \circ \end{array}$$

The odd roots are

$$b_2^{2-};\ b_j^{1-}=[b_j^{2-},\ \alpha_1^-],\quad j=2,3,4,$$
$$b_3^{i-}=[b_2^{i-},\ \gamma_3^-],\quad b_4^{i-}=[b_3^{i-},\ \gamma_4^-],\quad i=1,2.$$

The other conditions of decoupling are obtained using

$$\begin{aligned}&\Lambda_1^1=\alpha_1^-\gamma_3^-\gamma_4^-\Lambda_8^3, && \Lambda_3^2=\gamma_4^-\gamma_3^-\Lambda_{15}^2,\\&\Lambda_1^3=\gamma_3^-\gamma_4^-\Lambda_8^3, && \Lambda_6^2=\alpha_1^-\Lambda_6^4,\\&\Lambda_8^1=\alpha_1^-\Lambda_8^3, && \Lambda_3^2=\alpha_1^-\gamma_4^-\Lambda_6^4,\\& && \Lambda_3^4=\gamma_4^-\Lambda_6^4,\\&\Lambda_5^3=\alpha_1^-\Lambda_3^5, && \Lambda_1^4=\gamma_3^-\gamma_4^-\Lambda_8^4,\\&\Lambda_8^2=\gamma_4^-\gamma_3^-\Lambda_{\overline{10}}^2, && \Lambda_8^2=\alpha_1^-\Lambda_8^4,\\& && \Lambda_1^2=\alpha_1^-\gamma_3^-\gamma_4^-\Lambda_8^4,\end{aligned}$$

Since it is a bit tedious to extract all the conditions of typicality, we will do one example in detail, the (1, 16) of the third floor. Its highest weight Λ_6^1 is hidden behind $\Lambda_{\overline{15}}^3$ From $\Lambda_{15}^3=b_2^{1-}b_3^{2-}b_2^{2-}\Lambda_3^2$ one extracts the decoupling condition from $h_2(h_2+h_1+1)(h_2-h_3-1)\Lambda_3^2$ which is equivalent to the previous formula (5.1). To get to Λ^1 we have to plug α_1^- and γ^- anywhere in the product (5.2): there are 4 inequivalent configurations:

$$\begin{aligned}A_6^1&=b_2^{1-}b_3^{2-}b_2^{2-}\alpha_1^-\gamma_4^-\Lambda_3^2,\\B_6^1&=\alpha_1^-b_2^{1-}b_3^{2-}b_2^{2-}\gamma_4^-\Lambda_3^2,\\C_6^1&=\gamma_4^-b_2^{1-}b_3^{2-}b_2^{2-}\gamma_4^-\Lambda_3^2,\\D_6^1&=\alpha_1^-\gamma_4^-b_2^{1-}b_3^{2-}b_2^{2-}\Lambda_3^2.\end{aligned}$$

A, B, C, D span a four-dimensional space in which lies one member of (3, 15), (3, 6), (1, 115), (1, 6).

The highest-highest weights are

$$\begin{aligned}&\text{1st floor:}\quad \Lambda_8^3=b_2^{2-}\Lambda_3^2,\\&\text{2nd floor:}\quad \Lambda_{15}^2=b_2^{1-}b_2^{2-}\Lambda_3^2,\ \Lambda_6^4=b_3^{2-}b_2^{2-}\Lambda_3^2,\\&\text{3rd floor:}\quad \Lambda_{\overline{15}}^3=b_2^{1-}b_3^{2-}b_2^{2-}\Lambda_3^2,\\&\qquad\qquad \Lambda_3^5=b_4^{2-}b_3^{2-}b_2^{2-}\Lambda_3^2,\\&\text{4th floor:}\quad \Lambda_{\overline{10}}^2=b_3^{1-}b_2^{1-}b_3^{2-}b_2^{2-}\Lambda_3^2,\\&\qquad\qquad \Lambda_8^4=b_2^{1-}b_4^{2-}b_3^{2-}b_2^{2-}\Lambda_3^2\\&\text{5th floor:}\quad \Lambda_6^3=b_3^{1-}b_2^{1-}b_4^{2-}b_3^{2-}b_2^{2-}\Lambda_3^2,\\&\text{6th floor:}\quad \Lambda_3^2=b_4^{1-}b_3^{1-}b_2^{1-}b_4^{2-}b_3^{2-}b_2^{2-}\Lambda_3^2.\end{aligned}$$

If a highest-highest weight corresponds to $(\Pi_{k\leqslant i\leqslant 2\leqslant j\leqslant l}b_j^{i-})\Lambda_3^2$, the corresponding conditions of decoupling are [Eq. (4.1)]

$$\prod_{k\leqslant i\leqslant 2\leqslant j\leqslant l}\Big(h_2-\sum_{t=3}^{j}h_t+\sum_{t=i}^{1}h_t+4-i-j\Big). \tag{5.1}$$

The case of Λ_8^4 for example gives $\alpha_2=0,\ -2,1,3$.

$$\begin{aligned}&\Lambda_{15}^1=\alpha_1^-\Lambda_{\overline{15}}^3,\\&\Lambda_6^3=\gamma_4^-\Lambda_{\overline{15}}^3,\\&\Lambda_6^1=\alpha_1^-\gamma_4^-\Lambda_{\overline{15}}^3,\\&\Lambda_3^3=\gamma_3^-\gamma_4^-\Lambda_{\overline{15}}^3,\\&\Lambda_3^1=\alpha_1^-\gamma_5^-\gamma_4^-\Lambda_{\overline{15}}^3,\\&\Lambda_3^1=\alpha_1^-\gamma_4^-\Lambda_6^3,\\&\Lambda_3^3=\gamma_4^-\Lambda_6^3,\\&\Lambda_6^1=\alpha_1^-\Lambda_6^3.\end{aligned}$$

To apply the results of the previous sections we will consider instead the equivalent four states obtained by shifting $\alpha_1^-\gamma^-$ to the left:

$$\begin{aligned}E_6^1&=b_2^{1-}b_4^{2-}b_2^{2-}\Lambda_3^2,\\F_6^1&=b_3^{1-}b_2^{1-}b_2^{2-}\Lambda_3^2,\\G_6^1&=b_4^{1-}b_2^{1-}b_2^{2-}\Lambda_3^2.\end{aligned}$$

From Proposition (4.2) we see that the decoupling conditions of D_6^1 are the same as Λ_{15}^3 [Eq. (5.3)], i.e., $a_2=2,\ 0,\ 2$.

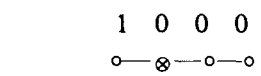
It is easy to find [cf. Eqs. (4.4) and (4.5)] that the corresponding conditions for

$$\begin{aligned}E_6^1&:\ a_2=-2,0,2,\\F_6^1&:\ a_2=-2,0,\\G_6^1&:\ a_2=-2,0,1.\end{aligned}$$

So,

$a_2=$	decouple	remain
$-2,0$	$(3,15)+(1,15)+(1,6)+(3,6)$	nothing
1	nothing	$(3,6)+(1,6)+(1,15)+(3,15)$
2	$(3,15)+(3,6)$	$(1,15)+(1,6)$

The decoupling condition $a_2 = 0$ is compatible with the adjoint representation and with the fundamental representation



which is the underlined part of

$$(2,1)\times\{(1,1) + (2,3) + [(1,6) + (3,\bar{3})] + [(2,8) + \cdots] + \cdots\}$$
$$= \underline{(2,1)} + \underline{(1,3)} + (3,3) + (2,6) + (4,\bar{3}) + \ldots$$

6. REMARKS AND CONCLUSION

In this paper we give a recipe to build explicitly typical and atypical representations of a superalgebra SU(m/n).

A typical representation is naturally connected to the following expansion in terms of Grassman variables ξ^i_j where $i = 1,\ldots,m, j = 1,\ldots,n$.

$$f^a(x,\xi) = \sum_{\substack{k=1,\ldots,n \\ l=1,\ldots,m}} \sum_{i_1\ldots i_k \; j_1\ldots j_l} f^{a j_1\cdots j_l}_{i_1\cdots i_k}(x)\xi^{i_1}_{j_1}\cdots\xi^{i_k}_{j_l},$$

is an SU(m)$\times$SU(n) group index which characterizes the SU(m)$\times$SU(n) representation to which the highest weight Λ of the superrepresentation belongs.

This expansion can (cf. Berezin, Ref. 7 and references therein) be seen as an expansion on a supermanifold.

In the case when Λ is a singlet of SU(m)$\times$SU(n), an atypical representation corresponds to the case where only a certain number of the ξ^i_j are linearly dependent, which could correspond to the manifestation of some constraints. That means atypical representations would correspond to nontrivial supermanifolds.

In the case $f^a(x,\xi)$ is a tensor field on the supermanifold, we saw that the decoupling scheme is more complicated and this should reflect here.

All this is related to the understanding of the representation of the supergroups which is still preliminary.

Local and global invariance under supergroups, when it is understood, should be closely related to extended supergravity and super Yang–Mills, should provide for a more systematic approach to them, and then allow a deeper understanding of these kind of theories.

ACKNOWLEDGMENTS

We gratefully acknowledge a very useful criticism from V. Kac and correspondence from J. Thierry-Mieg as well as many profitable conversations with colleagues in the University of Geneva, in particular with T. Schücker and H. Ruegg.

[1]V. Kac; (a) Adv. in Math. **26**, 8 (1977); (b) Commun. Math. Phys. **53**, 31 (1977).
[2]V. Rittenberg, *Lecture Notes in Physics* **79** (Springer, Berlin, 1978); M. Scheunert, *Lecture Notes in Mathematics* **716** (Springer, Berlin, 1979) and references therein.
[3]P. H. Dondi and P. D. Jarvis, Phys. Lett. B **84**, 75 (1980); J. G. Taylor, *ibid.* **83**, 331 (1979); E. J. Squires, *ibid.* **82**, 395 (1979); D. B. Fairlie, *ibid.* **82**, 97 (1979); Y. Neeman, *ibid.* **81**, 190 (1979).
[4]F. Iachello, Phys. Rev. Lett. **44**, 772 (1980), I. Bars and A. Balantekin, Yale University Preprint YTP 80-06; A. Balantekin, I. Bars, and F. Iachello, Yale University Preprint YTP 81-10; A. Balantekin and I. Bars, Yale University Preprint YTP 80-36 to be published in J. Math. Phys.
[5]J. Thierry-Mieg and B. Morel; (a) CMP to be published; (b) Prep. UGVA-DPT 1981/01-277.
[6]V. Kac, *Lecture Notes in Mathematics* **676** (Springer, Berlin, 1978).
[7]F. A. Berezin; (a) Sov. J. Nucl. Phys. **29**, 857 (1979); (b) **30**, 605 (1979).
[8]J. Ellis, H. K. Gaillard, and B. Zumino, Nucl. Phys. Lett. B **94**, 345 (1980).

Representations of graded Lie groups

B. R. Sitaram and K. C. Tripathy
Department of Physics and Astrophysics, University of Delhi, Delhi-110007, India

(Received 26 June 1981; accepted for publication 13 November 1981)

In this paper we establish the existence of a faithful matrix representation of finite type for every connected simply connected graded Lie group. We also show the 1-1 correspondence between finite-dimensional representations of a graded Lie algebra and the representations of finite type of the corresponding connected simply connected graded Lie group.

PACS numbers: 11.30.Pb, 02.20.Qs, 02.40. – k

I. INTRODUCTION

Recently, Kostant[1] has given a very elegant formulation of the theory of graded Lie groups. Mathematically the formulation seems to be more attractive than the one given by Kac and Berezin,[2] although few physical applications of Kostant's formulation have been made. In view of the quite detailed knowledge we have regarding the representation theory of Kac–Berezin graded Lie groups,[3] it would be necessary for us to study the representation theory of the Kostant graded Lie groups so as to make the similarities between the two formulations more transparent. Our results in this direction can be summarized by

Theorem 1: There exists a faithful matrix representation of finite type for every connected simply connected (csc) graded Lie group.

Theorem 2: The finite type representations of a csc graded Lie group are in 1-1 correspondence with the finite-dimensional representations of the corresponding graded Lie algebras.

In Sec. II, we give a brief resume of Kostant's formulation, while in Sec. III, we prove the above theorems.

II. RESUME OF THE KOSTANT THEORY[1]

Let $\mathbf{g} = \mathbf{g}_0 + \mathbf{g}_1$ be a Z_2 graded Lie algebra over $K = R$ or C. Let G be the unique csc Lie group with Lie algebra $\mathbf{g}_0$.

Definition 1[4]: The K-group ring of G, $K(G)$, is the free abelian group generated by elements of the form (r, g), $r \in K$, $g \in G$. Explicitly, $K(G)$ is an algebra over K, with the properties

$$\begin{aligned} &(r_1, g) + (r_2, g) = (r_1 + r_2, g), \\ &r(r_1, g) = (r_1, g)r = (rr_1, g), \\ &(r_1, g)\cdot(r_2, g') = (r_1 r_2, gg'), \quad r, r_i \in K,\ g, g' \in G. \end{aligned} \tag{1}$$

In the sequel, we denote the element (r, g) by rg.

Let $U(\mathbf{g})$ be the universal enveloping algebra over $\mathbf{g}$, i.e. $U(\mathbf{g}) = T(\mathbf{g})/J$ where $T(\mathbf{g})$ is the tensor algebra over $\mathbf{g}$, and J is the two-sided ideal of $T(\mathbf{g})$ defined by elements of the form

$$X \otimes Y - (-1)^{|X|\cdot|Y|} Y \otimes X - [X, Y], \quad X, Y \in \mathbf{g}, \tag{2}$$

where, e.g., $|X| = Z_2$ degree of X. Note, of course, that $U(\mathbf{g})$ is bigraded w.r.t. $Z_2 \otimes Z$. Both $K(G)$ and $U(\mathbf{g})$ are in fact Hopf algebras. We recollect

Definition 2[5]: A Hopf algebra over K is a triple $(H, \Delta, 1_K)$ where H is a graded algebra over K, $\Delta: H \to H \times H$ (the coproduct) and $1_K: H \to K$ (the counit) are homomorphisms of graded K algebras such that the diagram

$$\begin{array}{ccccc} & & & 1_H \times 1_K & H \times K \\ & \overset{\Delta}{\rightarrow} & H \times H & & \\ H & & & & H \\ & & & 1_K \times 1_H & K \times H \end{array} \tag{3}$$

commutes.

The map Δ can be explicitly given for $K(G)$ and $U(\mathbf{g})$: if $g \in K(G)$, then $\Delta(g) = g \otimes g$ while if $X \in \mathbf{g}$, $\Delta(X) = 1 \otimes X + X \otimes 1$, Δ being defined over the rest of $U(\mathbf{g})$ by the fact that Δ is an algebra homomorphism.

Finally, let ad: $\mathbf{g}_0 \times \mathbf{g} \to \mathbf{g}$, $(X, Y) \to [X, Y]$ be adjoint mapping restricted to $\mathbf{g}_0$. We know that ad exponentiates to give a unique map π: $G \times \mathbf{g} \to \mathbf{g}$ such that the diagram

$$\begin{array}{ccc} \mathbf{g}_0 \times \mathbf{g} & \overset{\text{ad}}{\rightarrow} & \mathbf{g} \\ \exp \times 1_{\mathbf{g}} \downarrow & & \downarrow 1_{\mathbf{g}} \\ G \times \mathbf{g} & \overset{\pi}{\rightarrow} & \mathbf{g} \end{array} \tag{4}$$

commutes, where exp: $\mathbf{g}_0 \to G$ is the usual exponential map.

Definition 3: The csc graded Lie group $E(G, \mathbf{g})$ with graded Lie algebra $\mathbf{g}$ is defined to be

$$E(G, \mathbf{g}) = K(G) \# U(\mathbf{g}), \tag{5}$$

where $\#$, the smash product, is taken w.r.t. π. Explicitly, $E(G, \mathbf{g})$ is a Hopf algebra generated by elements of the form (g, X), $g \in G$, $X \in \mathbf{g}$ with the properties

$$\begin{aligned} &\text{(i)} \quad (g, X)\cdot(g', Y) = (gg', X \cdot \pi(g, Y)), \\ &\text{(ii)} \quad \Delta(g, X) = (\Delta g, \Delta X), \quad g, g' \in G,\ X, Y \in \mathbf{g}. \end{aligned} \tag{6}$$

If $\mathbf{g}$ is trivially graded ($\mathbf{g} = \mathbf{g}_0$) than $E(G, \mathbf{g})$ can be shown to be isomorphic to the set of all distributions on $C^\infty(G)$ with finite support. Kostant has shown that if $\mathbf{g}$ is nontrivially graded, then a similar interpretation can be given in terms of distributions with finite support on a certain sheaf of commutative graded K algebras.

III. PROOFS OF THE THEOREMS

Let $E(G, \mathbf{g})$ be a csc graded Lie group.

Definition 4: A representation of $E(G, \mathbf{g})$ in $E'(G', \mathbf{g}')$ is a map $\Sigma: E(G, \mathbf{g}) \to E(G', \mathbf{g}')$ which preserves the Hopf algebra structure.

Definition 5: A representation $\Sigma: E(G, \mathbf{g}) \to E(G', \mathbf{g}')$ is said to be of finite type if $\mathbf{g}'$ is finite dimensional (as a vector

 0022-2488/83/010164-02$02.50

space over K).

Theorem 3: There exists a faithful matrix representation of finite type for $E(G, \mathbf{g})$.

Proof: By the generalized Ado Theorem,[6] we have the existence of an isomorphism

$$\sigma: \mathbf{g} \to \mathbf{g}', \tag{7}$$

where $\mathbf{g}'$ is a finite-dimensional matrix graded Lie algebra, such that $\sigma_0 = \sigma|_{\mathbf{g}_0}$ is an isomorphism $\sigma_0: \mathbf{g}_0 \to \mathbf{g}'_0$. We know that σ_0 exponentiates to define an isomorphism $\exp \sigma_0$: $G \to G'$, where G, G' are the csc Lie groups with Lie algebras $\mathbf{g}_0, \mathbf{g}'_0$. Further $\exp \sigma_0$ extends to a unique isomorphism

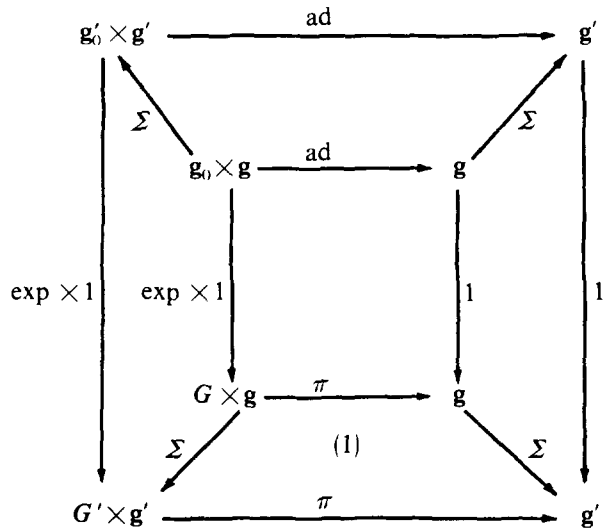
$$\exp \sigma_0: K(G) \to K(G') \tag{8}$$

of Hopf algebras.

Also, by the universality of $U(\mathbf{g})$, σ defines an isomorphism $U(\sigma): U(\mathbf{g}) \to U(\mathbf{g}')$ of Hopf algebras. Consider the map

$$\Sigma \equiv (\exp \sigma_0 \times U(\sigma)): \; K(G) \times U(\mathbf{g}) \to K(G') \times U(\mathbf{g}'). \tag{9}$$

We now show that Σ is in fact an isomorphism $\Sigma: E(G, \mathbf{g}) \to E(G', \mathbf{g}')$ of Hopf algebras. It is clear from the above that Σ preserves the coproduct and the counit in $K(G) \times U(\mathbf{g})$. It is therefore sufficient to prove that Σ commutes with π. Consider therefore the following diagram:



(10)

The outer diagram and all the subdiagrams except the subdiagram (1) commute, hence the diagram (1):

$$\begin{array}{ccc} & \overset{\pi}{\to} & \\ G \times \mathbf{g} & \to & \mathbf{g} \\ \exp \sigma_0 \times \sigma \downarrow & & \downarrow \sigma \\ G' \times \mathbf{g}' & \underset{\pi}{\to} & \mathbf{g}' \end{array} \tag{11}$$

also commutes which shows that Σ commutes with π. Hence the theorem.

Lemma 1: Let $\sigma: \mathbf{g} \to \mathbf{g}'$ be a representation with $\dim \mathbf{g}' < \infty$. Then σ defines a unique representation $\Sigma: E(G, \mathbf{g}) \to E(G', \mathbf{g}')$.

Proof: Σ is constructed as in the proof of Theorem 1, i.e., $\Sigma = \exp \sigma_0 \times U(\sigma)$. To prove that Σ commutes with π, we note that $E(G, \mathbf{g})$ and $E(G', \mathbf{g}')$ can be assumed to be matrix graded Lie groups. For such groups, we know that $\pi(g, X) = gXg^{-1}$. Assume that $g = \exp Z, Z \in \mathbf{g}_0$. Then,

$$\pi(g, X) = \sum_{n=0}^{\infty} \frac{(\operatorname{ad} Z)^n}{n!} X.$$

Now,

$$\begin{aligned} \Sigma\{(g, X), (g', y)\} &= \Sigma(gg', X \cdot gYg^{-1}) \\ &= (\exp \sigma_0(gg'), \sigma(X) \cdot \sigma(gYg^{-1})). \\ & \qquad g, g' \in G, \; X, \; Y \in \mathbf{g}. \end{aligned} \tag{12}$$

Also

$$\begin{aligned} \sigma(gYg^{-1}) &= \sigma\left(\sum_{n=0}^{\infty} \frac{(\operatorname{ad} Z)^n}{n!} Y\right) = \sum_{n=0}^{\infty} \frac{(\operatorname{ad} \sigma_0(Z))^n}{n!} \sigma(Y) \\ &= \pi(\exp \sigma_0(Z), \sigma(X)), \end{aligned} \tag{13}$$

where we have made use of the fact that $E(G, \mathbf{g})$ is a matrix graded Lie group in the second step. Now, if $g = \exp Z \cdot g'$, $g' \in G$, then we have,

$$\sigma(gYg^{-1}) = \exp \sigma_0(Z) \sigma(g' Yg'^{-1}) \exp(-\sigma_0(Z)). \tag{14}$$

Hence, $\sigma(gYg^{-1}) = (\exp \sigma_0(g)) \cdot Y \cdot (\exp \sigma_0(g^{-1})) \forall g \in G$. proving that Σ commutes with π. Hence the lemma.

Lemma 2: Every representation $\Sigma: E(G, \mathbf{g}) \to E(G', \mathbf{g}')$, $E(G, \mathbf{g}), E(G', \mathbf{g}')$ csc graded Lie groups of finite type, defines a unique representation $\sigma: \mathbf{g} \to \mathbf{g}'$.

Proof: Obviously $\sigma = \Sigma|_{\mathbf{g}}$ is a representation, $\sigma: \mathbf{g} \to U(\mathbf{g}')$. The fact that $\operatorname{Im} \sigma \subseteq \mathbf{g}'$ follows from the fact that $U(\sigma) = \Sigma|_{U(\mathbf{g})}$ preserves the Z degree.

As an immediate consequence, we have

Theorem 4: There is a 1-1 correspondence between the finite-dimensional representations of a graded Lie algebra and the representations of finite type of the corresponding csc graded Lie group.

Finally, let $\mathcal{GLA}$ be the category of finite-dimensional graded Lie algebras and let $\mathcal{GLG}$ be the category of csc graded Lie groups of finite type.

Let $\mathcal{K}_{os}: \mathcal{GLA} \to \mathcal{GLG}$ be defined by

$$\begin{aligned} &\mathcal{K}_{os}(\mathbf{g}) = E(G, \mathbf{g}), \\ &\mathcal{K}_{os}(\sigma: \mathbf{g} \to \mathbf{g}') = \Sigma': E(G, \mathbf{g}) \to E(G', \mathbf{g}'), \end{aligned} \tag{15}$$

using the notation of Theorem 1. Then,

Corollary: $\mathcal{K}_{os}$ is a covariant functor from $\mathcal{GLA}$ to $\mathcal{GLG}$. Further $\mathcal{K}_{os}$ is invertible, i.e., there exists a functor $\mathcal{K}_{os}^{-1}: \mathcal{GLG} \to \mathcal{GLA}$ such that $\mathcal{K}_{os} \circ \mathcal{K}_{os}^{-1} = 1_{\mathcal{GLG}}$ and $\mathcal{K}_{os}^{-1} \circ \mathcal{K}_{os} = 1_{\mathcal{GLA}}$, where $1_{\mathcal{GLA}}$ and $1_{\mathcal{GLG}}$ are the identity functors on $\mathcal{GLA}$, $\mathcal{GLG}$, respectively.

We hope to report on more investigations in this direction in a future paper.

[1] B. Kostant, "Graded Manifolds, Graded Lie Theory and Prequantization" in *Differential Geometric Methods in Mathematical Physics, Bonn 1975*, edited by K. Bleuler and A. Reetz, *Lecture Notes in Mathematics*, Vol. 576 (Springer, Berlin, 1977).

[2] F. A. Berezin, Preprint IETF-76 (1977), IETF-77 (1977), IETF-78 (1978); F. A. Berezin and G. I. Kac, Mat. Sb. **82**, 343 (1976).

[3] F. A. Berezin and V. N. Tolstoy, Commun. Math. Phys. **78**, 409 (1981); P. Fayet and S. Ferrara, Phys. Rep. **32**, 249 (1977).

[4] S. Mac Lane, *Homology, Grundlehren der Matematischen Wissenschaften*, Vol. 114 (Springer, Berlin, 1963).

[5] M. Sweedler, *Hopf Algebras* (Benjamin, New York, 1969).

[6] M. Scheunert, *Graded Lie Algebras, Lecture Notes in Mathematics* (Springer, Berlin, 1979).

Aspects of modular quantization

R. Kleeman
Department of Mathematical Physics, University of Adelaide, GPO Box 498, Adelaide, 5001 South Australia

(Received 6 April 1982; accepted for publication 20 August 1982)

A recent method of quantization is examined. A condition is found to isolate suitable Fock representations. Using these representations and a generalized Klein transform the quantization is compared with the normal quantization with U(m) symmetry. Finally the connection of this work with the color superalgebras is shown.

PACS numbers: 12.35.Ht

1. INTRODUCTION

Generalized methods of quantization were first proposed in 1953 by Green.[1] These schemes of quantization were known as para-Fermi and para-Bose statistics and had Fermi and Bose statistics, respectively, as special cases.

These quantizations remained somewhat of a curiosity until 1964 when Greenberg[2] applied them to the recently formulated quark model. He postulated that quarks were actually parafermions of order 3 rather than fermions (which are parafermions of order 1). This model allowed baryons to be symmetric with respect to interchange of quarks—a property which seemed to be required by experiment. This, in fact, was the first introduction of color into a quark theory. It was then shown[3,4] that the Greenberg model is essentially equivalent to the three-triplet color model. In other words, by replacing fermions by parafermions of order 3 one is really introducing an SU(3) [to be strictly correct U(3)] symmetry into the quark model. This symmetry is now known as color.

Attempts to pursue this "algebraic" notion of color further have run into difficulty with the "cluster property."[5] Basically what this says is that the creation and annihilation operators for quarks must always remain "confined" to the same baryon or meson. In order to resolve this difficulty Green in 1975[6] introduced a different generalized quantization which satisfied the cluster-property. This is known as modular quantization, the name modular deriving from the "clustering" of creation and annihilation operators into "modules." The aim of this paper is to examine the relationship between this new method of quantization and an ordinary quantization with U(m) symmetry. This comparison has already been carried out in some detail for para-Fermi statistics by Drühl *et al.*[4]

We begin by reviewing briefly the basics of the representation theory for para-Fermi quantization. The basic relations satisfied by the creation and annihilation operators are

$$[a_m,\tfrac{1}{2}[a_k^*,a_l]_-]_- = \delta_{mk}a_l,$$
$$[a_m,[a_k,a_l]_-]_- = 0. \tag{1.1}$$

Solutions to these equations are given by the Green's ansatz

$$a_k = \sum_{\alpha=1}^{p} b_k^{(\alpha)}, \tag{1.2}$$

where the $b_k^{(\alpha)}$ satisfy

$$\begin{aligned}
&[b_k^{*(\alpha)},b_l^{(\alpha)}]_+ = \delta_{kl},\\
&[b_k^{(\alpha)},b_l^{(\alpha)}]_+ = 0,\\
&[b_k^{*(\alpha)},b_l^{(\beta)}]_- = 0, \quad \alpha\neq\beta,\\
&[b_k^{(\alpha)},b_l^{(\beta)}]_- = 0, \quad \alpha\neq\beta.
\end{aligned} \tag{1.3}$$

Greenberg[7] showed that if one takes a Fock representation of the Von Neumann algebra of the a_k which satisfies

$$a_k a_l^*|\rangle = p\delta_{kl}|\rangle$$
with
$$a_k|\rangle = 0 \quad \forall_k, \tag{1.4}$$

then all irreducible representations (up to unitary equivalence) are given by the Fock representation of the Von Neumann algebra of the $b_k^{(\alpha)}$ [through (1.2) naturally]. It should be noted that there are possibilities for irreducible Fock representations not satisfying (1.4). These are the so-called reservoir states of Govorkov.[8,9] It would appear, however, that these result from choosing a nonvacuum state in the representation of the $b_k^{(\alpha)}$'s as a vacuum state for the a_k's.

We now move on to consider Modular quantization.[6] In the original paper the relations for the creation and annihilation operators are given with the aid of a "color" superscript

$$\begin{aligned}
&a_j^{(r)}a_k^{(s)} + a_k^{(s-1)}a_j^{(r+1)} = 0,\\
&a_j^{*(r)}a_k^{(s)} + a_k^{(s+1)}a_j^{*(r+1)} = \delta_{rs}\delta_{jk}.
\end{aligned} \tag{1.5}$$

The color superscript being defined with the aid of a unitary operator u which satisfies

$$u^m = 1, \quad m \text{ integral}, \tag{1.6}$$

and defines the color superscript through

$$a_k^{(r)} = u^{-r}a_k^{(0)}u^r. \tag{1.7}$$

It is possible to obtain relations not involving the superscripts,

$$\begin{aligned}
&a_{k_1}a_{k_2}\cdots a_{k_m}a_{k_{m+1}} + a_{k_{m+1}}a_{k_2}a_{k_3}\cdots a_{k_m}a_{k_1} = 0,\\
&a_{k_1}^*a_{k_2}a_{k_3}\cdots a_{k_{m+1}} + a_{k_1}a_{k_3}\cdots a_{k_{m+1}}a_{k_2}a_{k_1}^*\\
&\quad = \delta_{k_1k_2}a_{k_3}a_{k_4}\cdots a_{k_{m+1}},\\
&a_ja_k^*a_l + a_la_k^*a_j = \delta_{jk}a_l + \delta_{lk}a_j,
\end{aligned} \tag{1.8}$$

if we set $a_j \equiv a_j^{(r)}$ for any r.

 0022-2488/83/010166-07$02.50

The natural question to now ask is how the Fock representations of the a_k corresponds to those of the full $a_k^{(s)}$ algebra? Denote by $\mathscr{A}$[10] the former algebra satisfying (1.8) and by $\mathscr{B}$ the latter satisfying (1.5), then we have the following theorem.

Theorem: If a Fock representation of $\mathscr{A}$ satisfies

$$n<m, \quad a_{k_1}a_{k_2}\cdots a_{k_n}a_{j_n}^*a_{j_{n-1}}^*\cdots a_{j_1}^*|\rangle = \delta_{k_1j_1}\delta_{k_2j_2}\cdots\delta_{k_nj_n}|\rangle, \tag{1.9}$$

then it is unitarily equivalent to the subspace of the Fock representation of $\mathscr{B}$ generated by $a_k^{(s)}$ (s fixed). □

We first prove the following

Lemma: if

$$\phi \equiv a_{k_1}^*a_{k_2}^*\cdots a_{k_r}^*|\rangle \quad r<m,$$

then

$$a_ja_k^*\phi = \delta_{jk}\phi. \quad \square$$

Proof: For $r=m-1$ the result is immediate due to the second equation of (1.8). For $r<m-1$ consider firstly $j=k$. The third of (1.8) shows

$$a_j^*a_ja_j^*\phi = a_j^*\phi$$

$$\Rightarrow \|a_ja_j^*\phi\|^2 = \|a_j^*\phi\|^2 = (\phi, a_ja_j^*\phi). \tag{1.10}$$

However, (1.9) means that $\|a_j^*\phi\|^2 = 1$ (providing $|\rangle$ is normalized) and $\|\phi\|^2 = 1$ so (1.10) shows immediately that

$$a_ja_j^*\phi = \phi \quad \text{as required.}$$

For $j\neq k$ we have

$$a_j^*a_ja_k^* + a_k^*a_ja_j^* = a_k^*$$

$$\Rightarrow \|a_ja_k^*\phi\|^2 + (a_ja_j^*\phi, a_ka_k^*\phi) = (\phi, a_ka_k^*\phi)$$

$$\Rightarrow a_ja_k^*\phi = 0 \quad \text{as required.}$$

Corollary: The lemma together with the second equation of (1.8) shows that any state in a Fock representation of $\mathscr{A}$ can be written as a sum of terms such as

$$U = a_{k_1}^*a_{k_2}^*\cdots a_{k_s}^*|\rangle, \quad s \text{ arbitrary.} \quad \square$$

A fact that we will have cause to use later. Now the first and second of (1.8) together with (1.9) show that for any term of the above form *which is nonzero* there exists an operator $V\in\mathscr{A}$ s.t. $VU|\rangle = |\rangle$, namely

$$V = a_{k_s}a_{k_{s-1}}\cdots a_{k_1},$$

which means that any $A\in\mathscr{A}$ has a $B\in\mathscr{A}$ s.t. $BA|\rangle = |\rangle$. This enables us to say that $\mathscr{A}|\rangle$ is irreducible. In order to show that $\mathscr{B}|\rangle$ is irreducible it is necessary to carry out a linear transformation on the color indices

$$b_k^{(\alpha)} = m^{-1/2}\sum_{\beta=1}^{m-1}\epsilon^{-\alpha\beta}a_k^{(\beta)}, \tag{1.11}$$

where $\epsilon^m = 1$ (ϵ is the mth primitive root of unity). This transformation will later be seen to be central to the modular quantization. The $b^{(\alpha)}$ satisfy

$$b_k^{(\alpha)}b_l^{(\beta)} + \epsilon^{(\alpha-\beta)}b_l^{(\beta)}b_k^{(\alpha)} = 0,$$

$$b_k^{*(\alpha)}b_l^{(\beta)} + \epsilon^{(\beta-\alpha)}b_l^{(\beta)}b_k^{*(\alpha)} = \delta_{\beta\alpha}\delta_{kl}, \tag{1.12}$$

$$u^{-r}b_k^{(\alpha)}u^r = \epsilon^{r\alpha}b_k^{(\alpha)}.$$

Now if $\mathscr{B}'$ is the algebra generated by the $b^{(\alpha)}$ then because the transformation given by (1.10) is invertible it is easy to see that $\mathscr{B}'|\rangle = \mathscr{B}|\rangle$. Further, the operators $b_j^{*(\alpha)}b_j^{(\alpha)}$ are number operators in the usual sense and consequently the vacuum projection operator for $\mathscr{B}'$ (and $\mathscr{B}$) is

$$\Lambda = \prod_{k,\alpha}\frac{\sin \pi b_k^{*(\alpha)}b_k^{(\alpha)}}{\pi b_k^{*(\alpha)}b_k^{(\alpha)}}.$$

This shows that $\mathscr{B}|\rangle$ is irreducible.

Finally, with the lemma above, it is possible to calculate any vacuum expectation value of $\mathscr{A}$. So (1.9) and (1.8) together define an irreducible representation up to unitary equivalence. It is easy to check that (1.5) implies that (1.9) holds on any subspace of $\mathscr{B}|\rangle$ generated by $a^{(s)}$ (s fixed) and so the theorem is proved. It is worth noting, in passing, that for $m=2$, (1.8) and (1.9) become the defining relations for parastatistics of order 2, apart from a numerical factor.

It could still be asked whether the conditions (1.9) are necessary to obtain the relevant Fock representation. We remark only that this indeed is the case if we assume that there exists number operators n_k with the properties

$$P_\mu = \sum_k k_\mu n_k,$$

$$[a_k, n_l] = \delta_{kl}a_k, \tag{1.13}$$

$$\text{s.t.} \quad n_k|\rangle = 0,$$

where P_μ is the energy-momentum tensor. $P_0 = H$ the Hamiltonian, can have no negative eigenvalues which implies that n_k has none. It is not hard to see now that for $n\leqslant m$

$$W \equiv a_{k_1}a_{k_2}\cdots a_{k_n}a_{j_n}^*a_{j_{n-1}}^*\cdots a_{j_1}^*|\rangle$$

$$= \alpha(k_1,k_2,\ldots k_n,j_n,j_{n-1},\ldots,j_1)|\rangle,$$

where α is a numerical factor. We have then

$$(|\rangle, W|\rangle) = \alpha(k_1,k_2,\ldots,k_n,j_n,\ldots,j_1).$$

So unless this vacuum expectation value has the value

$$\delta_{k_1j_1}\delta_{k_2j_2}\cdots\delta_{k_nj_n},$$

then we have a unitarily inequivalent representation. As a final observation we see that

$$(|\rangle, W|\rangle) = (a_{j_n}^*a_{j_{n-1}}^*\cdots a_{j_1}^*|\rangle, a_{k_n}^*a_{k_{n-1}}^*\cdots a_{k_1}^*|\rangle),$$

which shows that there are $n!$ independent "n-particle" states for modular quantization, in the usual Fock representation given by (1.9). This contrasts with the situation in para-Fermi quantization, see Ref. 11, where there are in general, less.

2. HEISENBERG'S PRINCIPLE

As Green pointed out in his original paper, if one defines P_μ as

$$P_\mu = \int\left(\sum_{r=0}^{m-1} i\psi^{*(r)\alpha}\psi_{\alpha,\mu}^{(r)}\right)d^3x, \tag{2.1}$$

and one assumes that the $\phi_\alpha^{(r)}, \phi^{*(r)\alpha}$ (the spatial operators corresponding to the a_k, a_k^*) satisfy the equal-time relations [corresponding to (1.5)].

$$\psi^{*(r)\alpha}\psi_\beta^{(s)} + \psi_\beta^{(s+1)}\psi^{*(r+1)\alpha} = \delta_{r,s}\delta_{\alpha,\beta}\delta(\underset{\sim}{x}_\alpha - \underset{\sim}{x}_\beta),$$

$$\psi_\alpha^{(r)}\psi_\beta^{(s)} + \psi_\beta^{(s-1)}\psi_\alpha^{(r+1)} = 0, \tag{2.2}$$

then Heisenberg's principle is satisfied.

Now what is the interpretation to be given to the color superscripts? In this paper we shall adopt the following interpretation. The only physical states shall be those given by applying fields with a fixed color index to the vacuum. The color superscripts can be regarded as a mathematical convenience.

In this light, and given that the P_μ given in (2.1) is unique (see for example Takahashi[12]), we are faced with the following possibilities.

1. Assume P_μ has the form (2.1) and that the appearance of the superscripts is needed[13] to construct physical observables but is not needed in constructing physical states from the vacuum. This is somewhat analogous to the color singlet hypothesis of Q.C.D. We adopt this approach below.

2. Drop Heisenberg's principle (!). This is not as severe as it sounds for the following reason: suppose we take the ϕ_α to be the fields for "unobservable" quarks, then we would not expect that Heisenberg's principle should necessarily hold for the individual quark but merely for the meson $\psi^{*\alpha}\psi_\alpha, \psi_\alpha\psi^{*\alpha}$ composites and the baryon $\psi^{*\alpha}\psi^{*\alpha}\psi^{*\alpha}\cdots\psi^{*\alpha}$ (m factors) composites. If we were to take P_μ as

$$P_\mu = \int i[\psi^{*\alpha}, \psi_{\alpha,\mu}]_- d^3x, \tag{2.3}$$

then the following would hold [using (2.2)]:

$$[P_\mu, [\psi^{*\alpha}, \psi_\alpha]_+]_- = -i\partial_\mu([\psi^{*\alpha}, \psi_\alpha]_+), \tag{2.4}$$

$$[P_\mu, \underbrace{\psi^{*\alpha}\psi^{*\alpha}\cdots\psi^{*\alpha}}_{m \text{ factors}}]$$
$$= -i[\psi^{*\alpha}_{,\mu}\psi^{*\alpha}\cdots\psi^{*\alpha} + \psi^{*\alpha}\psi^{*\alpha}\cdots\psi^{*\alpha}_{,\mu}\psi^{*\alpha}]. \tag{2.5}$$

Equation (2.4) evidently has the correct form. However, (2.5) appears somewhat different to what one would expect. This difference can be explored if one introduces the unitary operator $U(a^\mu)$ corresponding to space-time translations

$$U(a^\mu) \equiv \exp(ia^\mu P_\mu). \tag{2.6}$$

Equation (2.5) can then be used to show that

$$U^*(a^\mu)\psi^{*\alpha}(x)\psi^{*\alpha}(x')\cdots\psi^{*(\alpha)}(x^m)U(a^\mu)$$
$$= \psi^{*\alpha}(x + a^\mu)\psi^{*\alpha}(x')\cdots\psi^{*\alpha}(x^{m-1})\psi^{*\alpha}(x^m + a^\mu).$$

This has the naive interpretation that when a baryon is subjected to a space-time translation only two out of the m quarks it contains are translated and the others are "left behind." Clearly this indicates that the x in $\phi^{*\alpha}(x)$ cannot have the straightforward interpretation that it does in usual field theories. For this reason we consider the possibility of P_μ given by (2.3) as somewhat questionable. It is interesting, however, that objects not of the usual $\phi^*\phi$ or $\phi^*\phi^*\cdots\phi^*$ (m factors) form (or products thereof) fail to satisfy relations of the form (2.4) and (2.5) and could be considered to be as "unphysical' as quarks.

3. TRANSFORMATION TO FERMI FIELDS

We begin, as previously, by reviewing briefly the Para-Fermi situation.[4] This is summarized in Fig. 1 to which the following comments are addressed.

The ansatz fields are the spatial analogs of the operators given in (1.3). As a result of Greenberg's work one may regard the parafields as a sub-algebra of these fields, providing one is taking the usual Fock representation. The transformation to the Fermi fields is achieved by the nonlocal Klein transformation. Explicitly we have

$$\begin{aligned}\phi^{(r)} &= \psi^{(r)}K_{r+1}, \quad r \text{ odd},\\ &= -i\psi^{(r)}K_r, \quad r \text{ even},\end{aligned} \tag{3.1}$$
$$K_r = \exp\left[i\pi \sum_{s=r}^{p} \int \psi^{*(s)}\psi^{(s)} d^3x\right],$$
$$K_r = K_r^* = K_r^{-1}.$$

For the Modular Fields the situation is slightly more complicated as has been intimated in the previous section. The reader is referred to Fig. 2 and the following comments are appropriate.

The restricted fields are the fixed color indice fields mentioned previously and can be taken to satisfy the spatial analogs of Eqs. (1.8). Provided we have a Fock space satisfying (1.9) they may be regarded as a subalgebra of the expanded fields.

The ansatz fields are the spatial analogs of the operators in (1.12) and can be obtained from the expanded fields via the

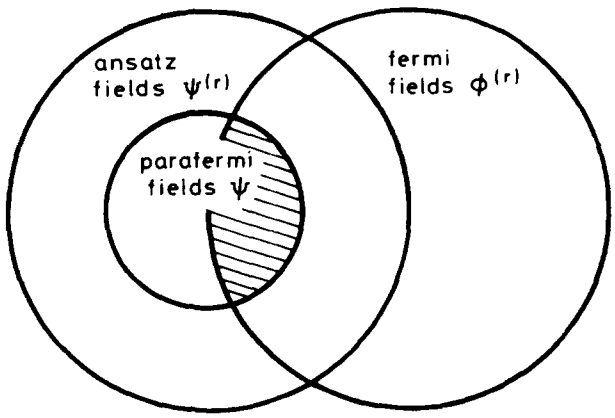


FIG. 1. The algebras associated with para-Fermi quantization. The shaded area represents algebraic elements which are possible physical variables, for example P_μ.

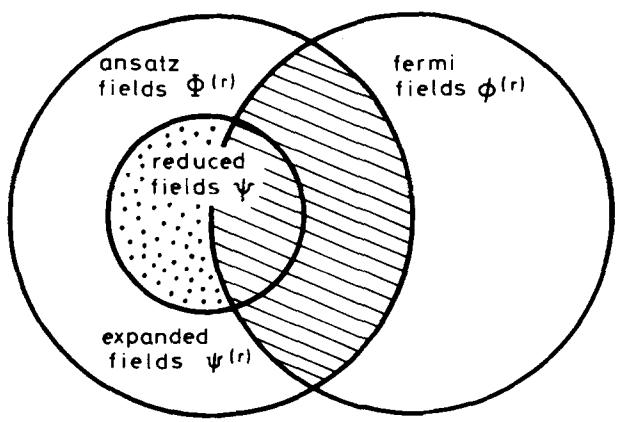


FIG. 2. The algebras associated with modular quantization. The shaded area represents possible physical variables; the P_μ which satisfies Heisenberg's principle is included. The dotted area shows where the "non-Heisenberg" P_μ is located.

unitary transformation given in (1.11). The transformation to Fermi fields can be achieved by the following "Generalized Klein transformation"

$$\phi^{(r)} = u^{r-1}\Phi^{(r)}. \tag{3.2}$$

As Green has pointed out, u may be written as

$$u = \exp(i\pi I_3),$$

with

$$I_3 \equiv \frac{1}{2\sqrt{m}} i \sum_{r=0}^{m-1} (A_{r\,r+1} - A_{r+1\,r}), \tag{3.3}$$

$$A_{rs} \equiv \int \psi^{*(r)}\psi^{*(s)}\, d^3x,$$

which is similar in form to the third equation of (3.1).

4. THE CONNECTION WITH GAUGE INVARIANT THEORIES

Doplicher *et al.*[14] have considered field theories (with normal commutation relations) in which a global gauge group "selects" out an "observable" algebra from a field algebra, in the sense that the observable algebra is the subalgebra of the field algebra left invariant by the gauge group. It also must obey the usual condition for observables—that of local commutativity. They show that in such a theory the super-selection quantum numbers are in one to one correspondence with the equivalence classes of irreducible representations of the gauge group.

If a para-Fermi theory is considered,[4] then providing certain obvious conditions are met then the Fermi fields in Fig. 1 can be regarded as above. As the above authors point out it is possible to identify various subalgebras of the parafield algebra which obey local commutativity. When these are written in terms of the "normal" Fermi fields they become "observable" algebras in the sense described above. Now providing one assumes that $K_r|\rangle = |\rangle$ then the Fock space H_p generated by the parafield algebra is contained in that generated by the Fermi fields. So in order to decide which gauge group is the appropriate one for parafields it is necessary to observe whether all super-selection numbers are possible in H_p (or equivalently, from above, whether all irreducible tensors of the gauge group are included in H_p). This turns out to be the case only for the group $U(p)$.

The conclusion then is that the Fock-like para-Fermi field theory is essentially equivalent to a normal Fermi theory with $U(p)$ symmetry except that the degeneracy associated with a particular set of super-selection numbers is less in the former case. In the case of Modular field theory the situation is somewhat different since when Heisenberg's principle is assumed (as we shall do) P_μ lies outside the reduced algebra (the analog of the para-Fermi algebra). Certainly then, "observable" algebras cannot be constructed purely from this algebra. In fact, we shall consider constructions from the expanded algebra. Following the philosophy put forward in Sec. 2 we shall regard the Fock space H_R generated by the reduced algebra as the physical Hilbert space.

We shall concern ourselves with the algebra[15] generated by elements of the form (x_1, x_2 belonging to some region V of space-time)

$$U(x_1,x_2) \equiv \sum_{r=0}^{m-1} \psi^{*(r)}(x_1)\psi^{(r)}(x_2), \tag{4.1}$$

upon transformation to the Fermi fields we obtain

$$U(x_1,x_2) = \sum_{r=0}^{m-1} \phi^{*(r)}(x_1)\phi^{(r)}(x_2). \tag{4.2}$$

As has been observed,[14] this algebra is the subalgebra of the Fermi-field algebra which is invariant under the gauge group $U(m)$. The gauge group being implemented through the automorphisms

$$a_g(\phi^{(r)}(x_1)) \equiv \sum_{s=0}^{m-1} A^{r}{}_{s}\phi^{(s)}(x_1), \tag{4.3}$$

where the matrix $A^{r}{}_{s}$ being a representation of $U(m)$ and $g \in U(m)$. It is worth pointing out that P_μ is also invariant under $U(m)$, a fact which is also true in the para-Fermi case.

Consider now the space H_R mentioned above. This space will become a subspace of the Fock space $\mathcal{F}$ generated by the Fermi fields provided that $u|\rangle = |\rangle$ which we assume. (The G.N.S. construction of the Hilbert spaces precedes in a nearly identical manner as para-field theory[4]). It remains to be seen then, whether H_R contains all inequivalent irreducible tensors of $U(m)$. We begin by proving the following theorem.

Theorem: Let δ be the Young symmetrizer corresponding to an arbitrary Young tableau with no more than m columns. Let the permutation group S_n be implemented as follows:

$$\gamma \in S_n \;\; \gamma(\psi^*(x_1)\psi^*(x_2)\cdots\psi^*(x_n)) = \psi^*(x_{\gamma(1)})\psi^*(x_{\gamma(2)})\cdots\gamma^*(x_{\gamma(n)}).$$

Then

$$\delta(\psi^*(x_1)\cdots\psi^*(x_n)) \neq 0. \tag{4.4}$$

Proof: It clearly suffices to show that

$$\delta(\psi^*(x_1)\cdots\psi^*(x_n))|\rangle = \delta(\psi^* x_1)\cdots\psi^*(x_n)|\rangle) \neq 0. \quad \square$$

Now

$$\psi^*(x_1)\cdots\psi^*(x_n) = \sum_{r_1,r_2,\ldots r_n=0}^{m-1} \psi^{*(r_1)}(x_1)\psi^{*(r_2)}(x_2)\cdots\psi^{*(r_n)}(x_n)|\rangle$$

$$= \sum_{r_1,r_2,\ldots,r_n=0}^{m-1} f(r_1,r_2,\ldots,r_n)\phi^{*(r_1)}(x_1)\phi^{*(r_2)}(x_2)\cdots\phi^{*(r_n)}(x_n)|\rangle, \tag{4.5}$$

with

$$f(r_1,r_2,\ldots,r_n)\equiv \epsilon^{\sum_{i=1}^{n}(i-1)r_i-\sum_{i<j}r_ir_i}\neq 0, \tag{4.6}$$

where (3.2) and (1.12) were used.

For notational purposes we write each term of (4.5) as

$$f(r_1,r_2,\ldots,r_n)|(r_1)(r_2)\cdots(r_n)|\rangle,$$

and the sum as

$$f(r_1,r_2,\ldots,r_n)|[r_1][r_2]\cdots[r_n]|\rangle, \tag{4.7}$$

the order of the brackets indicates which variable x_i they refer to.

Consider now an arbitrary $\gamma\in S_n$, then

$$\gamma(|\psi^*(x_1)\cdots\psi^*(x_n)|\rangle)=\gamma(f(r_1,r_2,\ldots,r_n))|[r_1][r_2]\cdots[r_n]|\rangle, \tag{4.8}$$

$$\gamma(f(r_1,r_2,\ldots,r_n))\equiv \mathrm{sign}(\gamma)f(r_{\gamma(1)},r_{\gamma(2)},\ldots,r_{\gamma(n)}).$$

Now let $\delta=\theta\eta$ where η is the symmetrizer and θ the antisymmetrizer for the Young tableau in Fig. 3. We have then

$$\delta(|\psi^*(x_1)\cdots\psi^*(x_n)|\rangle)=\delta(f(r_1,r_2,\ldots,r_n))|[r_1][r_2]\cdots[r_n]|\rangle. \tag{4.9}$$

Since any term $(r_0)(r_1)\cdots(r_n)$ is linearly independent of any other term $(s_0)(s_1)\cdots(s_n)$ unless $s_0=r_0,\ldots,s_n=r_n$ (see, for example, Ref. 16), it clearly suffices to show that

$$A\equiv\delta(f(0,1,\ldots,s(1)-1,0,1,\ldots,s(2)-s(1)-1,\ldots,0,1,\ldots,n-s(t)-1))\neq 0.$$

Consider first

$$\eta(f(0,1,\ldots,s(1)-1,0,1,\ldots,s(2)-s(1)-1,\ldots,0,1,\ldots,n-s(t)-1))$$
$$=d\eta_1(f(0,1,\ldots,s(1)-1))\eta_2(f(0,1,\ldots,s(2)-s(1)-1))\cdots\eta_t(0,1,\ldots,n-s(t)-1),$$

where

$$d=\epsilon^{(\frac{1}{2}\sum_{l=1}^{t}s(l)[s(l+1)-s(l)][s(l+1)-s(l)-1])}$$

and where the η_i are the symmetrizers for the ith rows of the Young tableau. Defining

$$m_i\equiv s(i)-s(i-1)\leqslant m,$$

we have

$$\eta_i(0,1,\ldots,m_i)=\epsilon^{-c}\sum_{\text{perm}}\mathrm{sign}(\lambda)\epsilon^{g(\lambda)}, \tag{4.10}$$

where

$$c\equiv\sum_{0<i<j}^{m_i-1}ij,$$

and

$$g(\lambda)\equiv\sum_{i=0}^{m_i-1}i\lambda(i),$$

and λ is an arbitrary permutation of $(0,1,\ldots,m_i-1)$. Now the right-hand side of (4.10) is just

$$\epsilon^{-c}\det(S),$$

where S is the $m_i\times m_i$ Sylvester matrix

$$S_{ij}=\epsilon^{ij}.$$

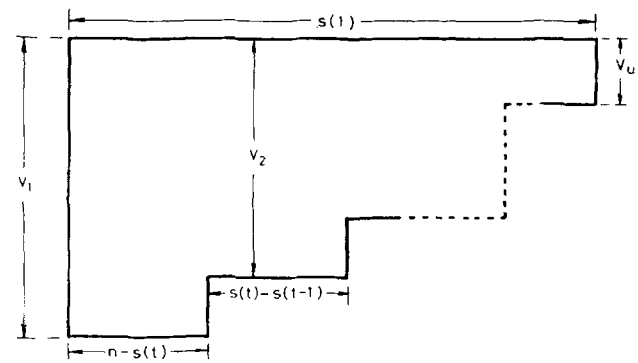


FIG. 3. The Young tableau referred to in the text.

S is (apart from a factor of $1/\sqrt{m}$) the inverse of the transformation given in (1.11). It determinant is given by

$$\det(S)=\prod_{i>j>0}^{m_i-1}(\epsilon^i-\epsilon^j) \tag{4.11}$$

(see, for example, Ref. 17), which is clearly nonzero for $m_i\leqslant m$ (and zero for $m_i>m$ which is why the Young tableau may have no more than m columns.) Thus

$$A=\theta[h(0,1,\ldots,s(1)-1,0,1,\ldots,s(2)-2(1)-1,\ldots,0,1,\ldots,n-s(t)-1)],$$

where $h\neq 0$. It is obvious from the definition (4.8) and the character of θ that

$$A=\prod_{i=1}^{u}V_i!\,h\neq 0,$$

where V_i is the number of boxes in the ith column of the Young tableau. Thus the theorem is proved. As a corollary to the theorem we note that

$$\delta(f(r_1,r_2,\ldots,r_n)[r_1][r_2]\cdots[r_n])\neq 0. \tag{4.12}$$

Consider now the algebra χ generated by elements such as

$$\phi^{*(r_1)}(x_1)\phi^{*(r_2)}(x_2)\cdots\phi^{*(r_n)}(x_n).$$

Clearly this carries a representation of U(m) equivalent to the n-fold tensor product of U(m) and following Ref. 4 we may decompose it into irreducible components using Young symmetrizers $\bar{\delta}\in S_n$ where S_n is implemented on χ as

$$\gamma(\phi^{*(r_1)}(x_1)\cdots\phi^{*(r_n)}(x_n))$$
$$=\phi^{*(r_1)}(x_{\gamma(1)})\phi^{*(r_2)}(x_{\gamma(2)})\cdots\phi^{*(r_n)}(x_{\gamma(n)}). \tag{4.13}$$

Comparing this implementation with that of the above theorem and using (4.12) we conclude that for every Young symmetrizer $\bar{\delta}\in S_n$ which has no more than m *rows* there exists an $X\in\chi$ ($X=f(r_1,r_2,\ldots,r_n)[r_1][r_2]\cdots[r_n]$) such that

$$\bar{\delta}(X)\neq 0.$$

Furthermore

$$\bar{\delta}(X)|\rangle\in H_R,$$

as can be seen from the proof of the theorem above.

To show that all Young tableau occur, let $\mathscr{P}$ be an ordinary symmetrizer. Then it is not hard to show, using (4.10), (4.5), and (4.6) that

$$P \equiv \mathscr{P}(\psi^*(x_1)\cdots\psi^*(x_m))$$

$$= \epsilon^{-c} \det(S) \sum_{\lambda\{i\}} \operatorname{sign}(\lambda)\phi^{*\{\lambda\{0\}\}}(x_1)\cdots\phi^{*\{\lambda\{m-1\}\}}(x_m) \neq 0,$$

where we are using the notation following (4.10) with $m_i = m$. As was noted in Ref. 4 we have for $g \in U(m)$

$$\alpha_g(P) = (\det g)^{-1} P. \tag{4.14}$$

Now P corresponds to a single column Young tableau in the implementation of S_n in (4.13). By multiplying a suitable product of P's by $\bar{\delta}(X)$ we obtain a $U(m)$ tensor corresponding to an arbitrary Young tableau. We have not, however, produced all "physically relevant" $U(m)$ tensors and in fact it is fairly easy to see that it is impossible to construct the meson singlet. This contrasts with the para-Fermi case where this is given by $[\psi^*(x_1),\psi(x_2)]_-$.

It is the view of this author that this "problem" may be solved with the introduction of reservoir states. These are vacuumlike states contained in the Fock space of the expanded algebra. An example would be the state $|k\rangle = a_j^{*(1)} b_l^{*(1)}|\rangle$ which clearly satisfies $b_l|k\rangle = a_j|k\rangle = 0$ (the a_j^* and b_l^* being, respectively, particle and antiparticle creation operators). The author hopes to pursue these matters further in a future paper.

5. COLOR SUPERALGEBRAS

Rittenberg *et al.*[18] have considered a generalization of a graded Lie algebra which has the generalized Lie product

$$\langle X_{\underset{\sim}{\alpha}}, X_{\underset{\sim}{\beta}} \rangle \equiv X_{\underset{\sim}{\alpha}} X_{\underset{\sim}{\beta}} - (-1)^{(\underset{\sim}{\alpha},\underset{\sim}{\beta})} X_{\underset{\sim}{\beta}} X_{\underset{\sim}{\alpha}} = C^{\underset{\sim}{\alpha}+\underset{\sim}{\beta}}_{\underset{\sim}{\alpha},\underset{\sim}{\beta}} X_{\underset{\sim}{\alpha}+\underset{\sim}{\beta}}, \tag{5.1}$$

where the $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ belong in general to an n-dimensional complex "grading" space and $(\underset{\sim}{\alpha},\underset{\sim}{\beta})$ is a mapping which is required to satisfy various properties so that the symmetry of (5.1) is maintained and a generalized Jacobi relation is satisfied. The usual graded Lie algebra is obtained by considering the vector space Z_2. Rittenberg, however, considered the more general vector space of $Z_r \oplus Z_s \oplus \cdots \oplus Z_t$. The relevance to this paper of these algebras becomes apparent when one realizes that the ansatz operators of both the para-Fermi and modular schemes form color superalgebras with grading spaces of $Z_2 \oplus Z_2 \oplus \cdots Z_2$ and $Z_m \oplus Z_m \oplus Z_2$, respectively. The para-Fermi case has already been discussed by Rittenberg. For the modular case one uses the mapping

$$(\underset{\sim}{\alpha},\underset{\sim}{\beta}) = (2/m)(\alpha_1\beta_2 - \alpha_2\beta_1) + \alpha_3\beta_3, \tag{5.2}$$

with $\alpha_1, \beta_1 \in Z_m$, $\alpha_2, \beta_2 \in Z_m$ and $\alpha_3, \beta_3 \in Z_2$ and sets, for example,

$$X_{(\alpha_i,1,1)} = b^{\{\alpha_i\}}, \quad X_{(-\alpha_i,-1,1)} = b^{*\{\alpha_i\}}, \quad X_{(0,0,0)} = 1,$$

and all other elements equal to zero.

In his paper Rittenberg claims that to every color superalgebra with grading space $Z_2 \oplus Z_2 \oplus \cdots \oplus Z_2$ (n factors) and mapping $(\underset{\sim}{\alpha},\underset{\sim}{\beta}) = \Sigma_i \alpha_i \beta_i$ there corresponds an ordinary superalgebra with identical structure constants. The correspondence being given by

$$Y_{\underset{\sim}{\alpha}} = \Gamma_1^{\alpha_1}\Gamma_2^{\alpha_2}\cdots\Gamma_n^{\alpha_n} \otimes X_{\underset{\sim}{\alpha}}, \tag{5.3}$$

where the Γ are Clifford matrices of dimension 2^v ($n = 2v, 2v+1$) which satisfy

$$\Gamma_i\Gamma_j + \Gamma_j\Gamma_i = 2\delta_{ij}1. \tag{5.4}$$

The Lie bracket for the $Y_{\underset{\sim}{\alpha}}$ is given by

$$\langle Y_{\underset{\sim}{\alpha}}, Y_{\underset{\sim}{\beta}} \rangle = Y_{\underset{\sim}{\alpha}} Y_{\underset{\sim}{\beta}} - (-1)^{\Sigma_i \alpha_i \cdot \Sigma_i \beta_i} Y_{\underset{\sim}{\beta}} Y_{\underset{\sim}{\alpha}}. \tag{5.5}$$

Unfortunately the structure constants are not identical and a short calculation will show that

$$C'^{\underset{\sim}{\alpha}+\underset{\sim}{\beta}}_{\underset{\sim}{\alpha},\underset{\sim}{\beta}} = (-1)^{\Sigma_{i>j}\beta_i\alpha_j} C^{\underset{\sim}{\alpha}+\underset{\sim}{\beta}}_{\underset{\sim}{\alpha},\underset{\sim}{\beta}}. \tag{5.6}$$

(The C' being the $Y_{\underset{\sim}{\alpha}}$ structure constants.) The lack of symmetry between $\underset{\sim}{\beta}$ and $\underset{\sim}{\alpha}$ in the constant of proportionality indicates that one may not overcome this problem by inserting some factor in the right-hand side of (5.3).

The existence of the generalized Klein transformation (3.2) suggests that a generalization of the explicit correspondence (5.3) should be possible. (An implicit generalization has been given by Scheunert.[19]) We shall confine ourselves here to the grading space $Z_m \oplus Z_m \oplus Z_2$ and mapping (5.2) and simply remark that obvious extensions exist. The correspondence is

$$Y_{\underset{\sim}{\alpha}} = E_2^{\alpha_1} E_1^{\alpha_2} \otimes X_{\underset{\sim}{\alpha}}, \tag{5.7}$$

where the E are the so-called generalized Clifford matrices (see for example, Ref. 20) which are m-dimensional and satisfy

$$E_1E_2 = \epsilon E_2E_1, \qquad E_i^m = 1. \tag{5.8}$$

The Lie bracket for the $Y_{\underset{\sim}{\alpha}}$ is

$$\langle Y_{\underset{\sim}{\alpha}}, Y_{\underset{\sim}{\beta}} \rangle = Y_{\underset{\sim}{\alpha}} Y_{\underset{\sim}{\beta}} - (-1)^{\alpha_1\beta_1} Y_{\underset{\sim}{\beta}} Y_{\underset{\sim}{\alpha}}, \tag{5.9}$$

and the new structure constants are given by

$$C'^{\underset{\sim}{\alpha}+\underset{\sim}{\beta}}_{\underset{\sim}{\alpha},\underset{\sim}{\beta}} = \epsilon^{\alpha_2\beta_1} C^{\underset{\sim}{\alpha}+\underset{\sim}{\beta}}_{\underset{\sim}{\alpha},\underset{\sim}{\beta}}. \tag{5.10}$$

Finally, we note that the new structure constants in (5.10) and (5.6) mean that the ordinary graded Lie algebras $Y_{\underset{\sim}{\alpha}}$ have the required symmetries in their Lie brackets and satisfy the usual Jacobi identities.

6. CONCLUSIONS

A comparison between modular quantization and the usual quantization with $U(m)$ symmetry has shown that the theories differ in that modular quantization does not produce a meson singlet state. It does, however, produce all the $U(m)$ states for baryons, the degeneracy with respect to the symmetry being greater than the para-Fermi theory but less than the usual quantization. It should be stressed that these conclusions depend on the condition (1.9) which selects out a particular Fock representation among many possibilities. Evidently other physical models may be constructed with other representations. This has in fact been done in the case of para-Fermi quantization by Bracken and Green.[21]

It appears that the color superalgebras provide the basis for the most fruitful generalization of the results presented in this paper. It would seem to this author that one cannot really consider the elements of these algebras as physical fields but one needs to consider linear transformations of same [such as that given by the Sylvester matrix in (1.10)]. This is to ensure that a suitable reduced algebra may be identified and a time ordering (see Ref. 6) defined. The usefulness

of the color superalgebras lies in the possibility of Klein transformations to ordinary Fermi fields. This allows us to examine whether the generalized quantizations correspond to any "usual" theory. In this light it is worth pointing out that the "non-Heisenberg" P_μ given in (2.3), when transformed to usual Fermi fields, involves the nonlocal u operator. This is a strong indication of its peculiarity.

Finally it should be pointed out that the above discussion could be easily altered to deal with Bose-like fields rather than the Fermi-like fields considered. In this case the ansatz algebra would be a color algebra (rather than superalgebra) with the grading space $Z_m \oplus Z_m$.

ACKNOWLEDGMENTS

I would like to thank Professor H. S. Green for his encouragement and for many fruitful discussions. Also P. Broadbridge who helped to clarify a number of hazy notions concerning the Klein transformations.

[1]H. S. Green, Phys. Rev. **90**, 270 (1953).
[2]O. W. Greenberg, Phys. Rev. Lett. **13**, 598 (1964).
[3]H. Fritzsch and M. Gell-Mann, "Light Cone Current Algebra," Tel-Aviv International Conference on Duality and Symmetry, 1971.
[4]K. Drühl, R. Haag, and J. E. Roberts, Commun. Math. Phys. **18**, 204–226 (1970).
[5]D. A. Gray, Progr. Theor. Phys. **47**, 1400 (1973).
[6]H. S. Green, Aust. J. Phys. **28**, 115–125 (1975).
[7]O. W. Greenberg and A. M. L. Messiah, Phys. Rev. **138**, B1155–B1167 (1965).
[8]A. B. Govorkov, Zh. Eksp. Teor. Fiz. **54**, 1785–1798 (1968) [Sov. Phys. JETP **27**, 1960–1966 (1968)].
[9]A. B. Govorkov, Int. J. Theor. Phys. **7**, 49–55 (1973).
[10]We shall confine ourselves here to the polynomial algebra. The extension to a Von Neumann algebra presents only technical problems which we avoid for sake of clarity.
[11]Y. Ohnuki and S. Kamefuchi, Ann. Phys. **51**, 337–358 (1969).
[12]Y. Takahashi, *An Introduction to Field Quantization* (Pergamon, Oxford, 1969), p. 160.
[13]It is easy to convince oneself that except for $m = 2$ it is impossible (without the introduction of field algebra inverses) to construct P_μ with a fixed color index.
[14]S. Doplicher, R. Haag, and J. E. Roberts, Commun. Math. Phys. **13**, 1–23 (1969).
[15]Strictly, we should talk about "nets." However, it is quite simple, in principle, to make the above discussion rigorous so we leave out this detail.
[16]A. I. Akhiezer and V. B. Berestetski, *Quantum Electrodynamics* (Wiley, New York, 1965), Chap. III.
[17]P. H. Hanus, *Theory of Determinants* (Ginn, Boston, 1903), p. 187.
[18]V. Rittenberg and D. Wyler, Nucl. Phys. B **139**, 189–202 (1978).
[19]M. Scheunert, J. Math. Phys. **20**, 712–720 (1979).
[20]A. Ramakrishnan, *L-Matrix Theory* (Tata-McGraw-Hill, Bombay, 1972), p. 92.
[21]A. J. Bracken and H. S. Green, J. Math. Phys. **14**, 1784–1793 (1973).

Mathematical properties and asymptotic expansion of the generalized quark statistical function

K. C. Chang, C. K. Chew, T. Y. Liang, and K. K. Phua
Department of Physics, National University of Singapore, Singapore 0511, Republic of Singapore

H. B. Low
IBM, Singapore 0106, Republic of Singapore

(Received 14 November 1980; accepted for publication 4 September 1981)

An asymptotic expansion for the generalized quark statistical distribution function in which quarks are introduced into Chao–Yang statistics is derived. Mathematical properties of the function are also examined.

PACS numbers: 14.80.Dg

I. INTRODUCTION

The Chao–Yang statistics[1] was first introduced in 1974 to determine the statistical charge distributions of nucleons and pions in "violent collisions."

The well-known quark structure of hadrons was later incorporated into Chao–Yang statistics to give the so-called quark statistics.[2–4] Such a scheme was strongly suggested by us because of our firm belief that quarks are the basic constituents of hadrons.

When two colliding systems impart violent impulse on each other either by transferring large transverse momentum or arresting each other completely in a small region of space before disintegration the physical assumption in quark statistics is that the quarks of the colliding systems must participate fully together with the quark-antiquark pairs created at short distances within the small central region. They are therefore asymptotically free and their mutual interactions may be neglected. Using quark statistics, we are able to calculate the particle ratios and dihadron spectra in the final state of a violent hadron–hadron or hadron–nucleon collision. The results obtained are in good qualitative agreement with experiments.

II. STATISTICAL DISTRIBUTION OF QUARKS

Consider a collection of l quarks of n types $q_1, q_2,...,q_n$ and their associated antiquarks $\bar{q}_1, \bar{q}_2,...,\bar{q}_n$. We define n_{q_i} to be the number of q_i quarks in the collection, and $n_{\bar{q}_i}$, n_{q_2}, $n_{\bar{q}_2}$ etc., are similarly defined. The quantum state of the collection is given by $(m_1, m_2,...,m_n)$ which is equivalent to a state of $m_1 q_1$ quarks, $m_2 q_2$ quarks, and so on. From these definitions, we have

$$\sum_{i=1}^{n} (n_{q_i} + n_{\bar{q}_i}) = l \tag{1}$$

and

$$n_{q_i} - n_{\bar{q}_i} = m_i, \quad i = 1,2,...,n. \tag{2}$$

As an illustration, the "quark quantum state" of $\pi^- p$ consisting of $2u$ quarks, $1\bar{u}$ quark, and $2d$ quarks is represented by (1,2,0,...,0).

Let $N^l_{m_1,m_2,...,m_n}$ be the number of possible ways of distributing $(m_1, m_2,...,m_n)$ over a collection of l quarks. The generating function of $N^l_{m_1,m_2,...,m_n}$ is defined as

$$\left[\sum_{i=1}^{n}\left(x_i + \frac{1}{x_i}\right)\right]^l = \sum_{m_1,m_2,...,m_n} N^l_{m_1,m_2,...,m_n} x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}, \tag{3}$$

where x_i and $1/x_i$ are variables for q_i quark and $\bar{q}_i$ antiquark, respectively. In order to investigate the symmetry properties of the distribution function we have assumed equal probability for the creation of all kinds of quark-antiquark pairs in the central region.

To obtain an explicit expression for $N^l_{m_1,m_2,...,m_n}$, we multiply both sides of Eq. (3) by $(1/x_1^{k_1+1})(1/x_2^{k_2+1})\cdots(1/x_n^{k_n+1})$ and perform the contour integral by means of Cauchy's integral formula.

We obtain

$$N^l_{m_1,m_2,...,m_n} = \frac{1}{(2\pi i)^n} \oint \frac{dx_1}{x_1} \oint \frac{dx_2}{x_2} \cdots \oint \frac{dx_n}{x_n} \cdot \frac{\{\Sigma_i (x_i + 1/x_i)\}^l}{x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}}. \tag{4}$$

III. MATHEMATICAL PROPERTIES AND THE PHYSICAL IMPLICATIONS OF THE DISTRIBUTION FUNCTION $N^l_{m_1,m_2,...,m_n}$

(a) $$\sum_{m_1,m_2,...,m_n} N^l_{m_1,m_2,...,m_n} = (2n)^l. \tag{5}$$

Proof: This follows by putting $x_i = 1$ in Eq. (3). It should be noted that the number of possible quark combinations increase much faster than the exponential increase e^l.

(b) $N^l_{m_1,m_2,...,m_n}$ is invariant under any permutation on the n symbols $(m_1, m_2,...,m_n)$.

Proof: This follows from the symmetry of Eq. (3) under interchange of the variables x_i's. This implies that $N^l_{m_1,m_2,...,m_n}$ depends only on the number of quark-antiquark combinations but not on their ordering.

(c) $N^l_{m_1,m_2,...,m_n}$ is invariant under any change of sign on the n symbols $(m_1, m_2,...,m_n)$.

Proof: This follows from Eq. (3) which is symmetric under the interchange of the variables x_i and $1/x_i$. Physically this means that the number of possible quark combinations $N^l_{m_1,m_2,...,m_n}$ depends only on the absolute differences between the quarks and the antiquarks of the same type.

 0022-2488/83/010173-04$02.50

(d) $N^{l}_{m_1,m_2,\ldots,m_n}=0$ if $(l+\Sigma_i m_i)$ is odd.

Proof: On replacing x_i by $-x_i$ in Eq. (3), we have

$$(-1)^{l}\left[\sum_i\left(x_i+\frac{1}{x_i}\right)\right]^{l}=\sum_{m_1,m_2,\ldots,m_n}(-1)^{\Sigma_i m_i}N^{l}_{m_1,m_2,\ldots,m_n}\cdot x_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}. \tag{6}$$

Combining Eqs. (3) and (6), we obtain

$$\sum_{m_1,m_2,\ldots,m_n}[(-1)^{l+\Sigma_i m_i}-1]N^{l}_{m_1,m_2,\ldots,m_n}x_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}=0,$$

and hence

$$N^{l}_{m_1,m_2,\ldots,m_n}=0 \quad \text{if}\left(l+\sum_i m_i\right)\text{is odd.}$$

(e)
$$\begin{aligned}N^{l}_{m_1,m_2,\ldots,m_n}&=N^{l-1}_{m_1-1,m_2,\ldots,m_n}+N^{l-1}_{m_1+1,m_2,\ldots,m_n}\\&\quad+N^{l-1}_{m_1,m_2-1,\ldots,m_n}+N^{l-1}_{m_1,m_2+1,\ldots,m_n}+\cdots\\&\quad+N^{l-1}_{m_1,m_2,\ldots,m_i-1,\ldots,m_n}\\&\quad+N^{l-1}_{m_1,m_2,\ldots,m_i+1,\ldots,m_n}.\end{aligned} \tag{7}$$

Proof: This is immediate from the definition of $N^{l}_{m_1,m_2,\ldots,m_n}$ if we note that a collection of l quarks is obtained from a collection of $l-1$ quarks by adding either a q_i or $\bar{q}_i$ quark.

(f)
$$\frac{m_i}{m_j}=\frac{N^{l-1}_{m_1,m_2,\ldots,m_i-1,\ldots,m_n}-N^{l-1}_{m_1,m_2,\ldots,m_i+1,\ldots,m_n}}{N^{l-1}_{m_1,m_2,\ldots,m_j-1,\ldots,m_n}-N^{l-1}_{m_1,m_2,\ldots,m_j+1,\ldots,m_n}} \tag{8}$$

Proof: By differentiating Eq. (3) with respect to x_i, we obtain

$$m_i N^{l}_{m_1,m_2,\ldots,m_i,\ldots,m_n}=l(N^{l-1}_{m_1,m_2,\ldots,m_i-1,\ldots,m_n}-N^{l-1}_{m_1,m_2,\ldots,m_i+1,\ldots,m_n}).$$

Similarly if we differentiate Eq. (3) with respect to m_j, we will obtain the above ratio for m_i over m_j.

IV. MODIFIED STATISTICAL DISTRIBUTION FUNCTION AND THE DEFINITION OF PROBABILITIES

In Sec. III for simplicity we assumed that the quark combinations resulting from violent collision are independent of the types of quarks involved in the reaction. This implies that for a fixed l, each of the states $N^{l}_{m_1,m_2,\ldots,m_n}$ is equally probable. Now, in order to satisfy the experimental results that the production of strange hadrons is suppressed in non strange hadron–hadron collisions, the equation of the generating function may be modified as follows:

$$\left\{\sum_i a_i\left(x_i+\frac{1}{x_i}\right)\right\}^{l}=\sum_{m_1,m_2,\ldots,m_n}N^{l}_{m_1,m_2,\ldots,m_n}\times x_1^{m_1}x_2^{m_2}\ldots x_n^{m_n}, \tag{9}$$

where a_i's can be interpreted as the relative strength in producing the ith type of quark among the n types.

The probabilities of finding the individual quarks and antiquarks, respectively, are defined as follows:

$$P_{q_i}=\frac{N^{l-1}_{m_1,m_2,\ldots,m_i-1,\ldots,m_n}}{N^{l}_{m_1,m_2,\ldots,m_n}}, \tag{10}$$

$$P_{\bar{q}_i}=\frac{N^{l-1}_{m_1,m_2,\ldots,m_i+1,\ldots,m_n}}{N^{l}_{m_1,m_2,\ldots,m_n}}. \tag{11}$$

These definitions can be easily generalized to a collection of k quarks.

To obtain the particle ratios in terms of the distribution function $N^{l}_{m_1,m_2,\ldots,m_n}$, we consider a violent collision of the hadron h_1 and h_2 having quark states $(p_1,p_2,\ldots,p_n)$ and $(q_1,q_2,\ldots q_n)$ and consisting of k_1 and k_2 quarks, respectively. It is reasonable to assume that the probability of producing h_1 is proportional to

$$\frac{N^{l-k_1}_{m_1-p_1,m_2-p_2,\ldots,m_n-p_n}}{N^{l}_{m_1,m_2,\ldots,m_n}}, \tag{12}$$

and similarly the probability of producing h_2 is proportional to

$$\frac{N^{l-k_2}_{m_1-q_1,m_2-q_2,\ldots,m_n-q_n}}{N^{l}_{m_1,m_2,\ldots,m_n}}. \tag{13}$$

Hence the ratio of the two hadrons is

$$\frac{h_1}{h_2}\propto\frac{N^{l-k_1}_{m_1-p_1,m_2-p_2,\ldots,m_n-p_n}}{N^{l-k_2}_{m_1-q_1,m_2-q_2,\ldots,m_n-q_n}}. \tag{14}$$

V. ASYMPTOTIC EXPANSION OF $N^{l}_{m_1,m_2,\ldots,m_n}$

We shall now derive the asymptotic expansion for the modified distribution function $N^{l}_{m_1,m_2,\ldots,m_n}$ by substituting $x_j=e^{i\theta_j}$ into Eq. (9). We obtain

$$\begin{aligned}&N^{l}_{m_1,m_2,\ldots,m_n}\\&=\frac{2^l}{(2\pi)^n}\int_{-\pi}^{\pi}d\theta_1\int_{-\pi}^{\pi}d\theta_2\cdots\int_{-\pi}^{\pi}d\theta_n\left(\sum_{i=1}^{n}a_i\cos\theta_i\right)^{l}\cdot\prod_{j=1}^{n}\cos(m_j\theta_j)\\&=\frac{2^l}{\pi^n}\int_0^{\pi}d\theta_1\int_0^{\pi}d\theta_2\cdots\int_0^{\pi}d\theta_n\left(\sum_i a_i\cos\theta_i\right)^{l}\cdot\prod_j\cos(m_j\theta_j),\end{aligned} \tag{15}$$

if we made use of the fact that cosine is an even function and sine is an odd function.

We can further split the integral into quadrant as

$$\int_0^{\pi}d\theta_i=\int_0^{\pi/2}d\theta_i+\int_0^{\pi/2}d(\pi-\theta_i), \tag{16}$$

and hence

$$\begin{aligned}&\int_0^{\pi}d\theta_1\int_0^{\pi}d\theta_2\cdots\int_0^{\pi}d\theta_n\\&=\int_0^{\pi/2}d\theta_1\int_0^{\pi/2}d\theta_2\cdots\int_0^{\pi/2}d\theta_n\\&\quad+\int_0^{\pi/2}d(\pi-\theta_1)\int_0^{\pi/2}d\theta_2\cdots\int_0^{\pi/2}d\theta_n\\&\quad+\int_0^{\pi/2}d\theta_1\int_0^{\pi/2}d(\pi-\theta_2)\cdots\int_0^{\pi/2}d\theta_n\\&\quad+\int_0^{\pi/2}d(\pi-\theta_1)\int_0^{\pi/2}d(\pi-\theta_2)\cdots\int_0^{\pi/2}d\theta_n\\&\quad+\cdots+\int_0^{\pi/2}d(\pi-\theta_1)\int_0^{\pi/2}d(\pi-\theta_2)\cdots\int_0^{\pi/2}d(\pi-\theta_n).\end{aligned} \tag{17}$$

Note that the products $\prod_i \cos(m_i\theta_i)$ in each term of Eq. (17) are equal in magnitude except for a difference in sign. Furthermore, the summation $(\sum_i a_i \cos\theta_i)^l$ is constructive for the first and last terms of Eq. (17) only and for the other terms in which the summation bears opposite sign they are destructive. As $l\to\infty$, the destructive terms are in orders of magnitude smaller than the constructive terms and we are therefore left with terms, i.e.,

$$\lim_{l\to\infty}\int_0^{\pi} d\theta_1\cdots\int_0^{\pi} d\theta_n = \int_0^{\pi/2} d\theta_1\int_0^{\pi/2} d\theta_2\cdots\int_0^{\pi/2} d\theta_n + \int_0^{\pi/2} d(\pi-\theta_1)\int_0^{\pi/2} d(\pi-\theta_2)\cdots\int_0^{\pi/2} d(\pi-\theta_n). \quad (18)$$

Hence

$$N^l_{m_1,m_2,\ldots,m_n} \sim \{1+(-1)^{l+\Sigma_j m_j}\}\frac{2^l}{\pi^n}\int_0^{\pi/2} d\theta_1\int_0^{\pi/2} d\theta_2\cdots\int_0^{\pi/2} d\theta_n \cdot\left(\sum_{i=1}^n a_i\cos\theta_i\right)^l\prod_{j=l}^n\cos(m_j\theta_j). \quad (19)$$

The integrand can be expressed in series as follows:

$$\prod_i\cos(m_i\theta_i) = \prod_i\left(1-\frac{m_i^2\theta_i^2}{2}+\frac{m_i^4\theta_i^4}{24}-\cdots\right) = 1-\frac{1}{2}\sum_i m_i^2\theta_i^2+\frac{1}{24}\sum_i m_i^4\theta_i^4 + \frac{1}{8}\sum_{i\neq j} m_i^2m_j^2\theta_i^2\theta_j^2+\cdots, \quad (20)$$

and

$$\sum_i a_i\cos\theta_i = A-\frac{1}{2}\sum_i a_i\theta_i^2+\frac{1}{24}\sum_i a_i\theta_i^4 - \frac{1}{720}\sum_i a_i\theta_i^6+\cdots, \quad (21)$$

where all a_i's >0 and $\sum_i a_i = A$.

Suppose

$$\sum_i a_i\cos\theta_i = \text{Series } X\cdot A\exp\left[-(1/2A)\sum_i a_i\theta_i^2\right], \quad (22)$$

so that on comparing with Eq. (21), we obtain

$$\begin{aligned}\text{Series } X &= 1+\left[\frac{1}{24A}\sum_i a_i\theta_i^4-\frac{1}{8A^2}\left(\sum_i a_i\theta_i^2\right)^2\right] \\ &\quad+\left\{-\frac{1}{720A}\sum_i a_i\theta_i^6-\frac{1}{24A^3}\left(\sum_i a_i\theta_i^2\right)^3 \right. \\ &\quad\left.+\frac{1}{48A^2}\left(\sum_i a_i\theta_i^4\right)\left(\sum_j a_j\theta_j^2\right)\right\}+\cdots \\ &= 1+\frac{1}{24A^2}\sum_i(Aa_i-3a_i^2)\theta_i^4-\frac{1}{8A^2}\sum_{i\neq j}a_ia_j\theta_i^2\theta_j^2 \\ &\quad+\frac{1}{720A^3}\sum_i(15Aa_i^2-30a_i^3-A^2a_i)\theta_i^6 \\ &\quad+\frac{1}{48A^3}\sum_{i\neq j}(Aa_ia_j-6a_i^2a_j)\theta_i^4\theta_j^2 \\ &\quad-\frac{1}{24A^3}\sum_{i\neq j\neq k}a_ia_ja_k\theta_i^2\theta_j^2\theta_k^2+\cdots. \end{aligned} \quad (23)$$

Using the finite integral for large l

$$\frac{1}{\pi}\int_0^{\pi/2}\theta_i^{2p}\exp\left(-\frac{a_i\theta_i^2}{2A}l\right)d\theta_i \sim A^p\left(\frac{A}{2a_i\pi l}\right)^{1/2}\frac{1\cdot3\cdot5\cdots(2P-1)}{a_i^Pl^P}, \quad (24)$$

we can express the distribution function when $(l+\Sigma_i m_i)$ is even as

$$N^l_{m_1,m_2,\ldots,m_n}\sim 2(2A)^l\left(\frac{A}{2\pi l}\right)^{n/2}\left(\prod_i a_i\right)^{-1/2}\cdot\text{Series } Y, \quad (25)$$

where Series Y is a series having inverse powers of l. The new series can be evaluated by the product

$$(\text{Series } X)^l\cdot\prod_i\cos(m,\theta_i),$$

and the preceding integral formula (24).

The first term in Series Y is obviously equal to 1. The second term is

$$-\frac{1}{2}A\sum_i\frac{m_i^2}{a_i}\cdot\frac{1}{l}+\frac{l}{24}\sum_i\frac{3(Aa_i-3a_i^2)}{a_i^2}\frac{1}{l^2} -\frac{l}{8}\sum_{i\neq j}\frac{a_ia_j}{a_ia_j}\frac{1}{l^2},$$

which is equal to

$$-\frac{1}{8l}\left(n^2+2n+A\sum_i\frac{4m_i^2-1}{a_i}\right). \quad (26)$$

The third term is

$$\begin{aligned}&\frac{A^2}{24}\sum_i\frac{3m_i^4}{a_i^2l^2}+\frac{A^2}{8}\sum_{i\neq j}\frac{m_i^2m_j^2}{a_ia_jl^2}-\frac{Al}{48}\sum_i\frac{15m_i^2(Aa_i-3a_i^2)}{a_i^3l^3} \\ &-\frac{Al}{48}\sum_{i\neq j}\frac{3m_i^2(Aa_j-3a_j^2)}{a_ia_j^2l^3}+\frac{Al}{8}\sum_{i\neq j}\frac{3a_ia_jm_i^2}{a_i^2a_jl^3} \\ &+\frac{Al}{16}\sum_{i\neq j\neq k}\frac{m_i^2a_ja_k}{a_ia_ja_kl^3}+\frac{l^2}{2\cdot24^2}\sum_i\frac{105(Aa_i-3a_i^2)^2}{a_i^4l^4} \\ &+\frac{l^2}{2\cdot24^2}\sum_{i\neq j}\frac{9(Aa_i-3a_i^2)(Aa_j-3a_j^2)}{a_i^2a_j^2l^4}+\frac{2l^2}{2\cdot8^2}\sum_{i\neq j}\frac{9a_i^2a_j^2}{a_i^2a_j^2l} \\ &+\frac{4l^2}{2\cdot8^2}\sum_{i\neq j\neq k}\frac{3a_i^2a_ja_k}{a_i^2a_ja_kl^4}+\frac{l^2}{2\cdot8^2}\sum_{i\neq j\neq k\neq m}\frac{a_ia_ja_ka_m}{a_ia_ja_ka_ml^4} \\ &-\frac{2l^2}{2\cdot8\cdot24}\sum_{i\neq j\neq k}\frac{3(Aa_i-3a_i^2)a_ja_k}{a_i^2a_ja_kl^4} \\ &-\frac{2\cdot2l^2}{2\cdot8\cdot24}\sum_{i\neq j}\frac{15(Aa_i-3a_i^2)a_ia_j}{a_i^3a_jl^4} \\ &+\frac{l}{720}\sum_i\frac{15(15Aa_i^2-30a_i^3-A^2a_i)}{a_i^3l^3} \\ &+\frac{1}{48}\sum_{i\neq j}\frac{3(Aa_ia_j-6a_i^2a_j)}{a_i^2a_jl^3}-\frac{l}{24}\sum_{i\neq j\neq k}\frac{a_ia_ja_k}{a_ia_ja_kl^3} \\ &=\frac{A^2}{128l^2}\left\{16\left(\sum_i\frac{m_i^2}{a_i}\right)^2+\sum_i\frac{9-40m_i^2}{a_i^2}+\sum_{i\neq j}\frac{1}{a_ia_j}\right. \\ &\left.-\sum_{i\neq j}\frac{8m_i^2}{a_ia_j}\right\}+\frac{(n+2)(n+4)A}{64l^2}\sum_i\frac{4m_i^2-1}{a_i} \\ &+\frac{1}{128l^2}(25n+78C_2^n+76C_3^n+24C_4^n, \end{aligned} \quad (27)$$

where C_m^n is the combinating function.

Substituting these results into expression (25), we obtain

the asymptotic expansion of $N^{l}_{m_1,m_2,\ldots,m_n}$ as

$$
\begin{aligned}
&2(2A)^{l}\left(\frac{A}{2\pi l}\right)^{n/2}\left(\prod_i a_i\right)^{-1/2}\\
&\quad\times\left\{1-\frac{1}{8l}\left[n^2+2n+A\sum_i\frac{4m_i^2-1}{a_i}\right]\right.\\
&\quad+\frac{A^2}{128l^2}\left[16\left(\sum_i\frac{m_i^2}{a_i}\right)^2\right]+\sum_i\frac{9-40m_i^2}{a_i^2}\\
&\quad+\sum_{i\neq j}\frac{1}{a_ia_j}\\
&\quad-\sum_{i\neq j}\frac{8m_i^2}{a_ia_j}+\frac{2(n+2)(n+4)}{A}\\
&\quad\times\sum_i\frac{4m_i^2-1}{a_i}\\
&\quad+\frac{1}{A^2}(25n+78C_2^n+76C_3^n+24C_4^n)\Big]\\
&\quad\left.+O\left(\frac{1}{l^3}\right)\right\}. \qquad (28)
\end{aligned}
$$

Having obtained the asymptotic expansion of $N^{l}_{m_1,m_2,\ldots,m_n}$, we shall examine the ratio r, which is the probability of producing hadron h_1 to the probability of producing hadron h_2 given by Eq. (14). We have

$$
r=\frac{N^{l-k_t}_{(m_1-p_1),\ldots,(m_i-p_i),\ldots,(m_n-p_n)}}{N^{l-k_b}_{(m_1-q_1),\ldots,(m_i-q_i),\ldots,(m_n-q_n)}}, \qquad (29)
$$

which can be expressed as

$$
\begin{aligned}
r&=(2A)^{(k_b-k_t)}\left(\frac{l-k_b}{l-k_t}\right)^{n/2}\\
&\quad\times\frac{1+t_1/(l-k_t)+t_2/(l-k_t)^2+\cdots}{1+b_1/(l-k_b)+b_2/(l-k_b)^2+\cdots}, \qquad (30)
\end{aligned}
$$

where t_1, t_2, b_1, b_2 are the coefficients of the asymptotic expansion of the distribution function and can be obtained from Eq. (28).

The second factor of Eq. (30) is

$$
\begin{aligned}
\left(\frac{l-k_b}{l-k_t}\right)^{n/2}&=1+\frac{n}{2l}(k_t-k_b)+\frac{n}{8l^2}(k_t-k_b)\\
&\quad\times[n(k_t-k_b)+2(k_t+k_b)]+\cdots \qquad (31)
\end{aligned}
$$

and the last factor can be simplified to

$$
1+\frac{t_1-b_1}{l}+\frac{1}{l^2}[k_tt_1-k_bb_1+t_2-b_2-(t_1-b_1)b_1].
$$

Hence the ratio becomes

$$
\begin{aligned}
r=(2A)^{(k_b-k_t)}&\left\{1+\frac{1}{l}\left[\frac{n}{2}(k_t-k_b)+(t_1-b_1)\right]\right.\\
&+\frac{1}{l^2}\left[\frac{n^2}{8}(k_t-k_b)^2+\frac{n}{4}(k_t^2-k_b^2)\right.\\
&+\frac{n}{2}(k_t-k_b)(t_1-b_1)+k_tt_1-k_bb_1\\
&\left.\left.+t_2-b_2-(t_1-b_1)b_1\right]+\cdots\right\}. \qquad (32)
\end{aligned}
$$

Note that

$$
t_1-b_1=-\frac{A}{2}\sum_i\frac{(m_i-p_i)^2-(m_i-q_i)^2}{a_i} \qquad (33)
$$

and

$$
\begin{aligned}
t_2-b_2=&\frac{A^2}{8}\left\{\left[\sum_i\frac{(m_i-p_i)^2}{a_i}\right]^2-\left[\sum_i\frac{(m_i-q_i)^2}{a_i}\right]^2\right\}\\
&-\frac{5A^2}{16}\sum_i\frac{(m_i-p_i)^2-(m_i-q_i)^2}{a_i^2}\\
&-\frac{A^2}{16}\sum_{i\neq j}\frac{(m_i-p_i)^2-(m_i-q_i)^2}{a_ia_j}\\
&+\frac{A}{16}(n+2)(n+4)\\
&\times\sum_i\frac{(m_i-p_i)^2-(m_i-q_i)^2}{a_i}. \qquad (34)
\end{aligned}
$$

[1] A. Chao and C. N. Yang, Phys. Rev. D **9**, 2505 (1974).
[2] C. K. Chew, H. B. Low, S. Y. Lo, and K. K. Phua, J. Physics G **6**, 17 (1980).
[3] C. K. Chew, L. C. Chee, H. B. Low, and K. K. Phua, Phys. Rev. D **19**, 3274 (1979).
[4] C. K. Chew, D. Kiang, and K. K. Phua, Phys. Rev. D **21**, 2525 (1980).

Splines and the projection collocation method for solving integral equations in scattering theory

M. Brannigan
University of Georgia, Department of Statistics and Computer Science, Athens, Georgia 30602
D. Eyre
National Research Institute for Mathematical Sciences of the CSIR, P. O. Box 395, Pretoria, Republic of South Africa

(Received 2 September 1981; accepted for publication 7 October 1981)

This paper investigates the method of projection collocation using cubic B-spline approximants to solve singular integral equations arising in scattering theory. Theoretical error bounds are provided for the approximation which give criteria for estimating the efficiency and convergence of the method. As numerical examples we solve the two-body K-matrix equation with a separable potential and the Reid $^1S_0$ soft-core potential.

PACS numbers: 24.10. − i, 02.30.Rz, 25.10. + s, 02.60.Nm

1. INTRODUCTION

The aim of this paper is to investigate the method of projection collocation, with cubic B splines as basis functions, to obtain approximate solutions of the singular integral equations that arise in scattering theory.

It has been shown[1] that three-body scattering can be described by the solution of singular multidimensional integral equations. The numerical solution of these equations is known to be difficult and complicated.

The present paper is devoted to a discussion of the numerical solution of the simpler two-body scattering problem, but a straightforward application of the methods described here may also be used to obtain numerical solutions of the integral equations that describe three-body scattering.

The general method of projection[2–5] has been used successfully to obtain approximate solutions of the integral equations that describe few-body systems. Osborn[6] investigates the use of moment methods to obtain approximate solutions of the singular two-body Lippmann–Schwinger equation. The use of splines and the Galerkin method to solve the corresponding homogeneous equation is described in Ref. 7. More recently Fiebig[8] has advocated the use of splines to solve scattering and bound-state problems. Similar methods have been used to solve integral equations that describe the three-body bound-state[9] and scattering problem.[10,11] In Ref. 11 use is made of bicubic splines to construct an approximate kernel that is degenerate. It should be remarked that a collocation method with bicubic spline approximants has also been used to solve the three-body integral–differential equations for the bound state problem in configuration space.[12]

In this paper it will be shown how the use of splines as approximants to the solution of two-body integral equations yields an easily programmable method for solving the scattering problem. An error analysis shows that this method is numerically stable. The method is also shown to be efficient provided that the fourth derivative of the scattering solution is sufficiently small.

Section 2 gives a mathematical formulation of the method, and provides error bounds for the approximation. Section 3 shows how this method may be applied to scattering integral equations, and in Sec. 4 we give our numerical results.

2. THEORY

The method of projection for solving integral equations of the second kind[2,3] has not only proved successful but has enabled the theoretical investigators readily to provide error estimates for the solution. The particular subclass of methods of projection we employ is that of collocation with spline approximants, which in turn gives rise to the need for the evaluation of principal value integrals. The method used for evaluating these integrals is that of subtracting the singularity and computing the resulting nonsingular integrals by means of Gaussian quadrature.

The general problem, from which our physical problem is taken, is the solution of the operator equation

$$(I-\mathscr{K})f=y\,, \tag{2.1}$$

where $f, y\in C(X)$, the space of continuous linear functionals defined on the compact set X; $\mathscr{K}$, I are linear operators mapping $C(X)$ into itself with I the identity operator.

The essence of the projection method of solving this equation is first to choose a linear subspace S of $C(X)$, with which to approximate f, and an appropriate bounded projection operator P mapping $C(X)$ onto $(I-\mathscr{K})[S]$. The approximate solution to our problem, using this projection, is the $g\in S$ such that

$$P(I-\mathscr{K})g=Py\,. \tag{2.2}$$

A general error analysis[2] shows us that if $\mathscr{K}$ is bounded $(I-\mathscr{K})^{-1}$ exists, and $\|\mathscr{K}-P\mathscr{K}\|$ is bounded by $\|(I-\mathscr{K})^{-1}\|^{-1}$, then

$$\|f-g\|\leqslant\|(I-P\mathscr{K})^{-1}\|\,\|f-Pf\|\,. \tag{2.3}$$

The particular problem we address is the solution of the integral equation

$$f(s)-\int_a^b K(s,t)f(t)dt=y(s),\quad a\leqslant s\leqslant b\,. \tag{2.4}$$

We choose for our linear subspace S of $C[a,b]$ a finite-dimensional space of cubic splines with a given set of knots. To be precise, let π_n be a partition of the interval $[a,b]$ defined by the knots $a=t_1<t_2<\cdots t_n=b$ with mesh spacing $h_n=\max\{(t_{i+1}-t_i):1\leqslant i\leqslant n\}$. On this partition together with the extended knots $t_{-2}\leqslant t_{-1}\leqslant t_0\leqslant t_1<\cdots<t_n\leqslant t_{n+1}\leqslant t_{n+2}\leqslant t_{n+3}$, we can construct the cubic B splines $\{B_{ni}: i=0,\ldots,n+1\}$. Each B spline B_{ni} is a cubic spline hav-

 0022-2488/83/010177-07$02.50

ing nonzero values over the interval (t_{i-2}, t_{i+2}), and the set of B splines form a basis for the $(n+2)$-dimensional subspace of cubic splines for the partition π_n.

Given $n+2$ distinct points $s_0,...,s_{n+1}$ in $[a,b]$ and any function $f \in C[a,b]$, we can define the operator P_n mapping $C[a,b]$ into S such that

$$P_n f(s_i) = f(s_i), \quad i = 0,...,n+1.$$

From the properties of cubic splines we have that

$$P_n(\alpha f + \beta g) = \alpha P_n f + \beta P_n g, \tag{2.5}$$

$$P_n^2 f = P_n(P_n(f)) = P_n f; \tag{2.6}$$

hence P_n is a projection operator.

Using the B-spline basis we let $P_n f$ be defined by

$$P_n f = \sum_{i=0}^{n+1} \alpha_{ni} B_{ni}; \tag{2.7}$$

then

$$P_n(I - \mathscr{K}) f = \sum_{i=0}^{n+1} \alpha_{ni}(I - \mathscr{K}) B_{ni}$$

and the coefficients $\{\alpha_{ni}\}$ are found from the system of linear equations

$$\sum_{i=0}^{n+1} \alpha_{ni} [(I - \mathscr{K}) B_{ni}] \,|_{s_j} = y(s_j), \quad j = 0,...,n+1, \tag{2.8}$$

where $|_S$ denotes the value of the operator at the point s.

Owing to the identity term, the linear equations so formed are well-conditioned integral equations of the second kind.

The method described here is useful because error estimates are available.

From De Boor and Schwartz[3] we have the following inequality: if $f \in C^4[a,b]$, then

$$\|f - P_n f\| \leqslant \tfrac{5}{384} \|f''''\| h_n^4, \tag{2.9}$$

and in particular, if the collocation points of the set $\{S_i\}$ are the points t_1, $(t_1 + t_2)/2$, $t_2,...,t_{n-1}$, $(t_{n-1} + t_n)/2$, t_n, then

$$\|(I - P_n \mathscr{K})^{-1}\| \leqslant 1 + \tfrac{5}{2}(h_n/h_n^1)^2, \tag{2.10}$$

where $h_n^1 = \min\{(t_{i+1} - t_i): 1 \leqslant i \leqslant n-1\}$.

Hence using the equation given above, if g is our approximate solution then an upper bound for the error is given by

$$\|f - g\| \leqslant (1 + \tfrac{5}{2}(h_n/h_n^1)^2)(\tfrac{5}{384}\|f''''\| h_n^4). \tag{2.11}$$

and a rate of convergence of $O(h_n^4)$ results.

We note here that the collocation process described above does not prescribe any form to the kernel. However, the coefficients of the linear equations involve the integrals

$$\int_a^b K(s,t) B_{ni}(t)\,dt, \quad i = 0,...,n+1, \tag{2.12}$$

and to effect a numerically stable algorithm these integrals need to be evaluated accurately.

In scattering problems we assume the kernel K is singular and of the principal value type. The evaluation of the moment integrals is best performed using the method of subtraction of the singularity. An error estimate for the evaluation of this integral using the method of subtraction and a quadrature method of at least second order can be derived as in Ref. 13. Thus if

$$K(\cdot,t) = G(\cdot,t)/(t - u), \tag{2.13}$$

then for $a \leqslant \alpha \leqslant u \leqslant \beta \leqslant b$ we have an error estimate for the numerical solution of Eq. (2.12) of the form

$$(\delta^4/720)[F''''(\beta) - F''''(\alpha)] + O(\delta^6), \tag{2.14}$$

where δ is the step length used in the integration and

$$F'''' = 4G'B_{ni}''' + 6G''B_{ni}'' + 4G'''B_{ni}' + G''''B_{ni}. \tag{2.15}$$

As shown by Sloan[4] we can improve on the approximation g_0 of f which interpolates the points $\{(s_i, g_0(s_i)): i = 0,..., n+1\}$ by constructing the sequence of functions $\{g_i\} \in C[a,b]$ from

$$g_{i+1}(s) = y(s) + \int_a^b K(s,t) g_i(t)\,dt. \tag{2.16}$$

As we shall see later, this procedure, which we refer to as the iterative improvement, has the effect of smoothing out oscillations in the approximation g_0. Note that for all $i, g_i(s_j) = g_0(s_j), j = 0,..., n+1$; hence this smoothing operation does not change the approximation at the collocation points.

3. SCATTERING EQUATIONS

We now apply the method described in Sec. 2 to a physical problem, viz., the solution of the principal-value integral equations that describe two-body scattering. The partial-wave equation for the half-shell K matrix $M(p,k)$ has the form

$$M(p,k) = v(p,k) - \frac{2}{\pi} \int_0^\infty v(p,p') \times \frac{p'^2 dp'}{p'^2 - k^2} M(p',k), \quad p \in [0,\infty] \tag{3.1}$$

where $v(p,p')$ is the potential and k is the on-shell momentum. The integral in Eq. (3.1) is evaluated with respect to the principal value prescription. The solution $M(p,k)$ is a real-valued function and can be expressed in terms of the phase shift $\delta(k)$ by

$$M(k,k) = -[k \cot \delta(k)]^{-1}. \tag{3.2}$$

It is convenient to map the integral in Eq. (3.1) onto a finite interval. To do this we introduce the variable $x \in [-1, +1]$ by the mapping

$$p(x) = \eta\left(\frac{1+x}{1-x}\right), \tag{3.3}$$

where η is a constant scale parameter. Equation (3.1) can now be written in the form

$$M(p(x),k) = v(p(x),k) - 2\eta^3 \frac{2}{\pi} \int_{-1}^{1} v(p(x), p'(x')) \left(\frac{1+x'}{1-x'}\right)^2 \times \frac{M(p'(x'),k)\,dx'}{[\eta^2(1+x')^2 - k^2(1-x')^2]}, \quad x \in [-1, +1]. \tag{3.4}$$

Let M_n be the spline approximation $P_n M$ to M given by

$$M_n(p(x),k) = \sum_{i=0}^{n+1} \alpha_{ni}(k) B_{ni}(x) . \tag{3.5}$$

The linear system of equations formed is

$$\sum_{i=0}^{n+1} [B_{ni}(x_j) + I_{ni}(p_j,k)] d_{ni}(k) = v(p_j,k), \quad j = 0,...,n+1 , \tag{3.6}$$

where $x_j, j = 0,...,n+1$, are the $n+2$ collocation points and $p_j \equiv p(x_j)$ and $I_{ni}(x_j,k)$ are moment integrals formed from the kernel of Eq. (3.1) convoluted with the cubic B spline, viz.,

$$I_{ni}(p_j,k) = 2\eta^3 \frac{2}{\pi} \int_{-1}^{1} v(p_j,p(x)) \left(\frac{1+x}{1-x}\right)^2 \times \frac{B_{ni}(x)dx}{[\eta^2(1+x)^2 - k^2(1-x)^2]} , \quad i,j = 0,...,n+1 . \tag{3.7}$$

The choice of $\eta = k$ leads to a pole in the integrand at the on-shell value $x = 0$. Thus

$$I_{ni}(p_j,k) = \frac{k}{\pi} \int_{-1}^{1} v(p_j,p(x)) \left(\frac{1+x}{1-x}\right)^2 \times B_{ni}(x) \frac{dx}{x} , \quad i,j = 0,...,n+1 . \tag{3.8}$$

We therefore see that by applying the projection collocation method to the integral equation in Eq. (3.1) we have reduced the problem of solving an integral equation to that of setting up a solvable linear system. Moreover, by expanding the K matrix in the cubic B-spline basis the moment integrals are restricted to be a convolution of the kernel with a function no more oscillatory than a cubic polynomial.

Using one iteration of the iterative improvement scheme we obtain a new approximation $\bar{M}_n$ given by the formula

$$\bar{M}_n(p(x),k) = v(p(x),k) - \sum_{i=0}^{n+1} \alpha_{ni}(k) I_{ni}(p(x),k) , \tag{3.9}$$

where $I_{ni}(p(x),k)$ is given by Eq. (3.7) or (3.8) with $p(x)$ replacing the discrete values p_j. At $p(x) = p_j$ the approximation $\bar{M}_n$ coincides with the approximation M_n, i.e.,

$$\bar{M}_n(p_j,k) = M_n(p_j,k) . \tag{3.10}$$

It follows that $\bar{M}_n$ provides no additional information at these points. However, along $p(x) \neq p_j$ the approximation $\bar{M}_n$ differs from the cubic B-spline interpolate M_n.

4. NUMERICAL RESULTS

In the formulation of the computer procedure to test the validity of this method, particular attention was paid to numerically stable methods used in the subprocedures.

The evaluation of the B-splines were performed using the iterative technique of Cox[14] given by

$$B_{ni}^{(m)}(x) = \frac{(x - x_{i-m}) B_{ni-1}^{(m-1)}(x) + (x_i - x) B_{ni}^{(m-1)}(x)}{x_i - x_{i-m}} , \tag{4.1}$$

where m is the order of the spline. For the cubic spline, $m = 4$. This procedure is both computationally fast and accurate, as the analysis in Ref. 14 shows.

Having this stable method of calculating the B-spline basis we need to evaluate the moments I_{ni} which are of the principal value type. Different methods were compared in Ref. 13, the conclusion being that use of the method of subtraction of the singularity gave better results. In our program, we used a Gauss–Legendre quadrature formula with error analysis as described above. The actual moments were calculated over the intervals $[t_{i+1},t_i]$ and summed. It is to be noted that over each interval only four of the cubic B splines are nonzero so that each B spline needs to be evaluated only once at the quadrature points, which contributes to computational efficiency.

The linear equations formed were solved using an LU decomposition, a method that is satisfactory because the matrix of coefficients is well conditioned.

All computations were performed on a CDC Cyber with a 48-bit mantissa.

To illustrate the properties of our method we first considered an equation for which a simple analytic solution exists, viz., a one-term separable potential of the Yamaguchi type.[15] In momentum space this potential has the form

$$v(p,p') = \lambda / (p^2 + \beta^2)(p'^2 + \beta^2) . \tag{4.2}$$

We shall assume that this system can support a bound state, i.e., that the parameter λ is fixed by ensuring that the K matrix has a pole at the binding energy $k^2 = -\epsilon$ in Eq. (3.1). Over the scattering region, $k \geqslant 0$, the half-off-shell K matrix is given by

$$M(p,k) = \left(\frac{k^2+\beta^2}{p^2+\beta^2}\right) \left[-\beta + \frac{k^2+\beta^2}{2\beta} + \frac{(k^2+\beta^2)^2}{2\beta(\beta+\alpha)^2} \right]^{-1} , \tag{4.3}$$

where $\alpha = (\epsilon)^{1/2}$. Thus we have a simple analytical form for the half-off-shell K matrix against which we can test our method.

We computed our results using the constants $\beta = 1.44401\ \text{fm}^{-1}$ and $\epsilon = 0.053695\ \text{fm}^{-2}$. These parameters were chosen so that the Yamaguchi potential will approximately describe low energy neutron–proton scattering in the s-wave spin-triplet channel. For each calculation we used evenly spaced knots and collocation points as indicated in our error analysis.

TABLE I. Convergence of the projection collocation method for a separable potential at scattering threshold. Scattering length a is measured in fm.

n	$a = M(0,0)$
Exact	−5.3800
3	−5.5333
4	−5.6077
5	−5.4025
6	−5.3777
7	−5.3782
8	−5.3791
9	−5.3795
10	−5.3797
15	−5.3800

We first calculated the scattering length for this potential, and the results are given in Table I. We note the fast convergence to the correct value of $a = -5.3800$ fm.

We next considered the solution at a scattering energy of $k^2 = 1$ fm^{-2}. In Fig. 1 we graph this solution over the interval $[-1,1]$. Figure 2 illustrates the difference between the exact solution for the K matrix and its value when calculated by our procedure for various numbers of knots. As can be seen, the error decreases rapidly as the number of knots is increased. Using iterative improvement once, we get an approximation $\bar{M}_n$ given by Eq. (3.9). In practice this approximation requires little additional computation as it requires only the evaluation of the moment integrals $I_{ni}(p,k), i = 0,\ldots,n+1$, at the interpolation point p. To compute these integrals we use the recursion relation given in the Appendix. In our Fig. 2 we also graph the result of iterative improvement for four knots. We note how the error has been smoothed out, but passes through the interpolation points given by our original collocation points.

To show the convergence properties of our procedure we tabulate, see Table II, the $L_\infty[-1,1]$ or Tchebysheff error norm. For a function $f(x), x\in[-1,1]$, the Tchebysheff norm is defined by $\|f\| = \max_{-1\leqslant x\leqslant 1}|f(x)|$. The first column gives the error for the procedure for various different knot spacings, column 2, the error after one iteration of iterative improvement. For this particular potential we can also calculate the theoretical error bound given in our analysis, which for $k^2 = 1$ fm^{-2} is

$$\|M_n - M\| \leqslant 0.4333\left(\frac{2}{n-1}\right)^4. \tag{4.4}$$

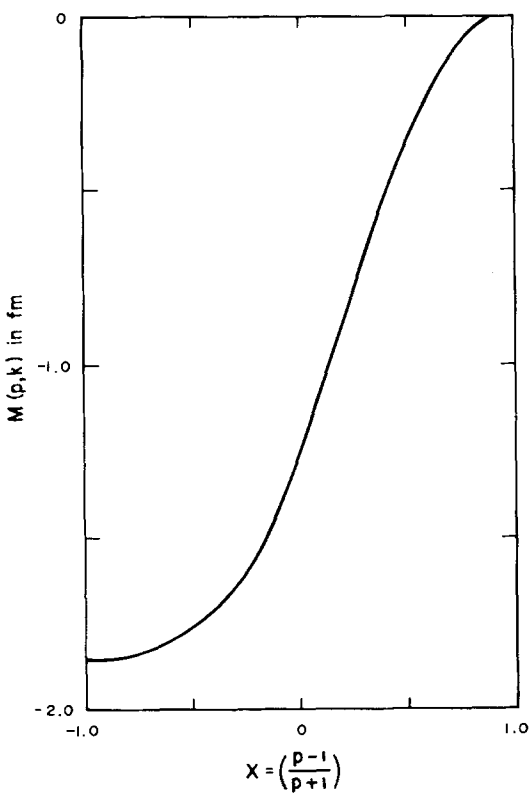


FIG. 1. Exact K matrix given by Eq. (4.3) at a scattering energy of $k^2 = 1$ fm^{-2}.

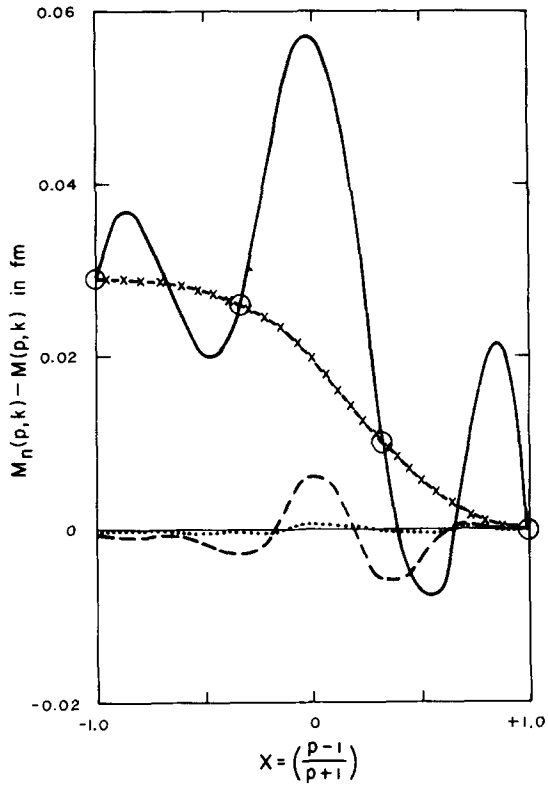


FIG. 2. Approximate K matrix obtained using the separable potential. Error function $M_n(p,k) - M(p,k)$ for the cubic B-spline interpolate with n knots. The — curve is for $n = 4$, the ---- curve for $n = 6$, and the .... curve for $n = 8$. The -x-x curve is the result of using the iterative improvement, where 0 denotes the position of the knots.

In column 3 of Table II the value for this error bound is given.

We first note that, as column 1 shows, the rate of convergence for our procedure is of order h^4. Column 2 shows this rate of convergence also and a consistently smaller error for each value of n. Finally, as predicted by the theoretical analysis, each of these errors is smaller than the theoretical error bound given in column 3.

The results we have shown so far have been used to test the method against a problem for which an exact solution is known. We next consider the solution of a more complicated system.

To this end we carried out the numerical calculation with the Reid $^1S_0$ soft-core potential.[16] In momentum space

TABLE II. Convergence of the Tchebysheff norm at a scattering energy of $k^2 = 1$ fm^{-2}.

n	$\|M_n - M\|$	$\|\bar{M}_n - M\|$	Bound
4	0.57(−1)	0.29(−1)	0.86(−1)
6	0.58(−2)	0.90(−3)	0.11(−1)
8	0.93(−3)	0.29(−3)	0.29(−2)
10	0.29(−3)	0.95(−4)	0.11(−2)
12	0.20(−3)	0.41(−4)	0.47(−3)
16	0.48(−4)	0.11(−4)	0.14(−3)
24	0.71(−5)	0.20(−5)	0.25(−4)
32	0.16(−5)	0.58(−6)	0.75(−5)

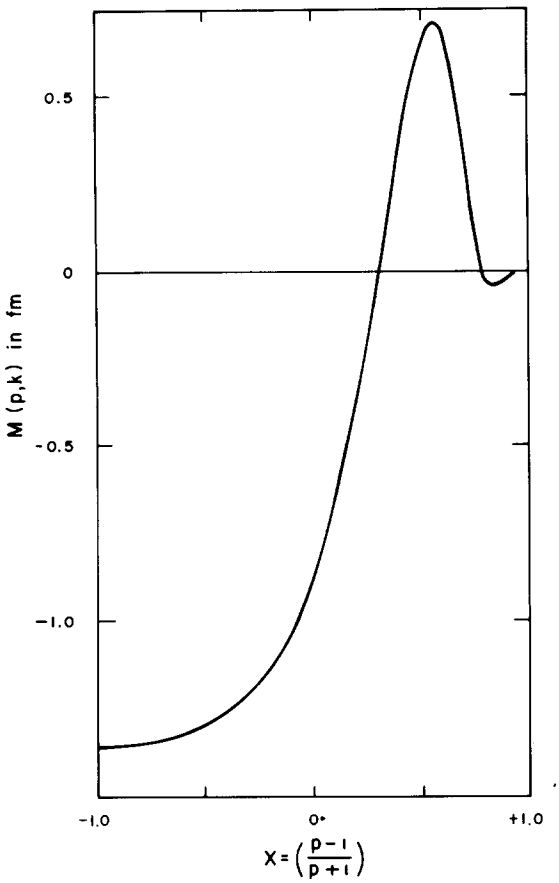


FIG. 3. K matrix for the Reid $^1S_0$ soft-core potential at a scattering energy of $E_{\rm lab} = 48$ MeV.

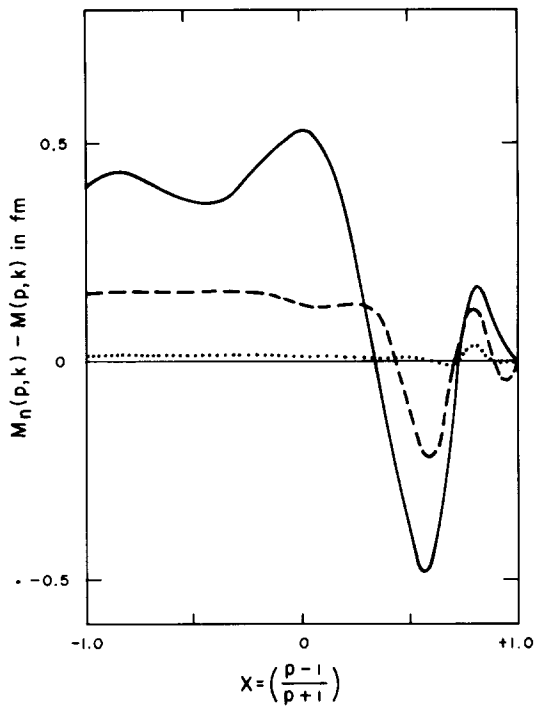


FIG. 4. Approximate K matrix obtained using the Reid $^1S_0$ soft-core potential. Error function $M_n(p,k) - M(p,k)$ for the cubic B-spline interpolate with n knots. The — curve is for $n = 4$, the ---- curve for $n = 8$, and the .... curve for $n = 16$. Knots are evenly spaced over the interval $[-1, +1]$.

this potential has the form

$$v(p,p') = \frac{1}{4\mu_1 pp'} \sum_{i=1}^{3} V_i \ln\left[\frac{(p+p')^2+\mu_i^2}{(p-p')^2+\mu_i^2}\right], \quad (4.5)$$

where $\mu_1 = 0.7$ fm^{-1}, $\mu_2 = 4\mu_1$, $\mu_3 = 7\mu_1$, $V_1 = -10.463$ MeV fm^{-3}, $V_2 = -1650.6$ MeV fm^{-3}, and $V_3 = 6484.2$ MeV fm^{-3}. To obtain a reference solution we used the method of Haftel and Tabakin[17] to solve Eq. (3.1).

Figure 3 illustrates the reference solution at a laboratory scattering energy of $E_{\rm lab} = 48$ MeV. This solution exhibits a much more complicated structure than our previous example, and therefore provides a more stringent test of our method than the system illustrated in Fig. 1.

Since we do not have an analytical form for the moment integrals, we have to evaluate these integrals by numerical quadrature. We partition the region of integration according to the mesh spacing π_n, i.e., we write

$$I_{ni}(p_j,k) = \sum_{l=1}^{n-1} I_{ni}^{(l)}(p_j,k) \quad (4.6)$$

and evaluate the integrals

$$I_{ni}^{(l)}(p_j,k) = \frac{k}{\pi}\int_{x_l}^{x_{l+1}} v(p_j,p(x))\left(\frac{1+x}{1-x}\right)^2 B_{ni}(x)\frac{dx}{x}. \quad (4.7)$$

Of course only integrals over the interval $[x_{i-2}, x_{i+2}]$ will contribute to the sum in Eq. (4.6). In the case when $0 \in [x_l, x_{l+1}]$, the integral in Eq. (4.7) is evaluated with respect to the principal value prescription with error analysis as given above. Using the method of subtracting the singularity, we write

$$\begin{aligned} I_{ni}^{(l)}(p_j,k) &= \frac{k}{\pi}\int_{x_l}^{x_{l+1}} \left[v(p_j,p(x))\left(\frac{1+x}{1-x}\right)^2 \right. \\ &\quad \left. \times B_{ni}(x) - v(p_j,k)B_{ni}(0)\right]\frac{dx}{x} \\ &\quad + v(p_j,k)B_{ni}(0)\frac{k}{\pi}\ln\left|\frac{x_{l+1}}{x_l}\right|. \end{aligned} \quad (4.8)$$

The integral is now replaced by a standard Gauss–Legendre quadrature formula.

Figure 4 illustrates the difference between the approximate spline solution M_n and the reference solution of Eq. (3.1). The knots are evenly spaced over the interval $[-1, +1]$.

An important practical consideration in applying this spline approximant is the correct positioning of the knots. We do not attempt a detailed analysis of this problem in the present paper; however, it is interesting to modify the mesh spacing and then solve the system in Fig. 3 with a different choice of knots. For this purpose we choose the knot positions as Clenshaw–Curtis points over the interval $[-1, +1]$, i.e., the knots are given by the formula

$$x_{ni} = -\cos\frac{(i-1)\pi}{n-1}, \quad i = 1,\ldots,n. \quad (4.9)$$

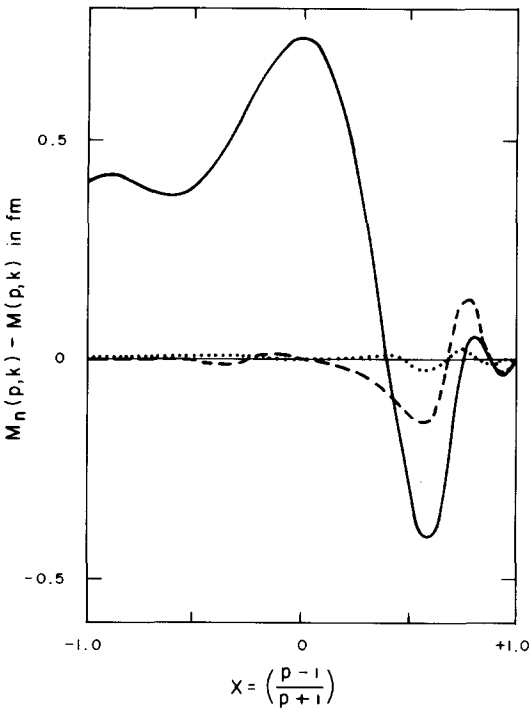


FIG. 5. Same as in Fig. 4 but with the knots spaced along Clenshaw–Curtis points.

Figure 5 illustrates the difference between the approximate spline solution M_n and the reference solution. Although the rate of convergence is found to be similar to that of the evenly spaced knots, the choice of knot positions is seen to significantly influence the approximate solution.

It should be remarked that in any practical application of this method it is advisable to concentrate the knots in a region where the scattering solution has the most structure, and in this way one may hope to optimize the accuracy of the solution for a given number of knots.

These results demonstrate that the projection collocation method with spline approximants is a practical and numerically stable procedure for solving the integral equations arising in scattering theory. As shown by the error analysis the approximate solution can be found to arbitrary accuracy. In the test problems that we have considered, the method has demonstrated that accurate solutions can be obtained with only a small number of knots and hence small linear systems.

APPENDIX

In this Appendix we evaluate the moment integrals

$$I_{ni}(p,k) = \frac{\lambda}{p^2+\beta^2}\frac{k}{\pi}\int_{-1}^{1}\frac{(1+x)^2}{k^2(1+x)^2+\beta^2(1-x)^2}B_{ni}(x)\frac{dx}{x}, \quad i=0,\ldots,n+1. \tag{A1}$$

We write

$$I_{ni}(p,k) = \frac{k}{k^2+\beta^2}\frac{\lambda}{p^2+\beta^2}h_{ni}(k), \tag{A2}$$

where

$$h_{ni}(k) = \frac{1}{\pi}\int_{-1}^{1}\left(\frac{1}{x}+2+x\right)B_{ni}(x)\frac{dx}{1+a(k)x+x^2}, \tag{A3}$$

$$a(k) = 2\left(\frac{k^2-\beta^2}{k^2+\beta^2}\right). \tag{A4}$$

To simplify matters we shall assume a partition π_n: $-1 \leqslant x_1 < x_2 < \ldots < x_n \leqslant +1$ of uniformly spaced knots. Consider the interval between two adjacent knots $[x_i, x_{i+1}]$ with spacing $h = (x_{i+1} - x_i)$. The four cubic B splines over this interval are

$$B_{nl}(x) = \frac{1}{24h^4}\sum_{r=0}^{3} b_{l,r}x^r, \quad l = i-1,\ldots,i+2 \tag{A5}$$

where coefficients $b_{l,r}$ are defined with the knots $x_{i-2}, x_{i-1},\ldots, x_{i+3}$ by

$$\begin{aligned}
&b_{i\ 1,0} = x_{i+1}^3, \quad b_{i-1,1} = -3x_{i+1}^2,\\
&b_{i-1,2} = 3x_{i+1}, \quad b_{i\ 1,3} = -1,\\
&b_{i,0} = -(x_{i\ 2}x_{i+1}^2 + x_i x_{i+2}^2 + x_{i-2}x_{i+1}x_{i+2}),\\
&b_{i,1} = x_{i+1}^2 + x_{i+2}^2 + x_{i-2}x_{i+1} + 2x_i x_{i+2}\\
&\qquad + x_{i-1}x_{i+1} + x_{i-1}x_{i+2} + x_{i+1}x_{i+2},\\
&b_{i,2} = -(x_{i-2} + x_{i-1} + x_i + 3x_{i+1} + 3x_{i+2}),\\
&b_{i,3} = 3,\\
&b_{i+1,0} = x_{i-1}x_i x_{i+2} + x_{i-1}^2 x_{i+1} + x^2{}_i x_{i+3},\\
&b_{i+1,1} = -(x_{i-1}^2 + x_{i-1}x_i + 2x_{i\ 1}x_{i+1} + x_{i-1}x_{i+2}\\
&\qquad + x_i^2 + x_i x_{i+2} + 2x_i x_{i+3}),\\
&b_{i+1,2} = 3x_{i\ 1} + 3x_i + x_{i+1} + x_{i+2} + x_{i+3},\\
&b_{i+1,3} = -3,\\
&b_{i+2,0} = -x_i^3, \quad b_{i+2,1} = 3x_i^2,\\
&b_{i+2,2} = -3x_i, \quad b_{i+2,3} = 1.
\end{aligned} \tag{A6}$$

After substituting Eq. (A5) into Eq. (A3) we obtain

$$h_{nl}(k) = \frac{1}{24h^4}\sum_{r=-1}^{4} c_{l,r}\frac{1}{\pi}\int_{x_i}^{x_{i+1}} \times x^r\frac{dx}{1+a(k)x+x^2}, \quad l = i-1,\ldots,i+2, \tag{A7}$$

where

$$\begin{aligned}
&c_{l,\ -1} = b_{l,0}, \quad c_{l,0} = 2b_{l,0} + b_{l,1},\\
&c_{l,1} = b_{l,0} + 2b_{l,1} + b_{l,2}, \quad c_{l,2} = b_{l,1} + 2b_{l,2} + b_{l,3},\\
&c_{l,3} = b_{l,2} + 2b_{l,3}, \quad c_{l,4} = b_{l,3}.
\end{aligned} \tag{A8}$$

We now write Eq. (A7) as the sum

$$h_{nl}(k) = \frac{1}{24h^4}\sum_{r=-1}^{4} c_{l,r}g_r(k), \quad l = i-1,\ldots,i+2, \tag{A9}$$

where

$$g_r(k) = \frac{1}{\pi}\int_{x_i}^{x_{i+1}} x^r\frac{dx}{1+a(k)x+x^2}. \tag{A10}$$

To determine the integrals in Eq. (A10), reference can be

made to standard tables of integrals.[18] We find that

$$g_{-1}(k) = \frac{1}{2\pi}\left[\ln\left(\frac{h}{1+a(k)h+h^2}\right) - g_0(k)\right], \tag{A11}$$

$$g_0(k) = \frac{1}{\pi}\left(\frac{k^2+\beta^2}{2k\beta}\right)\arctan\left[\left(\frac{k^2+\beta^2}{2k\beta}\right)\left(\tfrac{1}{2}a(k)+h\right)\right], \tag{A12}$$

and for $r \geqslant 1, g_r$ is given by the recurrence relation

$$g_r(k) = \frac{1}{\pi}\frac{h^r}{r} - a(k)g_{r-1}(k) - g_{r-2}(k). \tag{A13}$$

[1]L. D. Faddeev, *Mathematical Aspects of the Three-Body Problem in the Quantum Scattering Theory*, Steklov Math. Institute 69 (1963) (Davey, New York, 1965).
[2]Kendall E. Atkinson, *A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind* (SIAM, Philadelphia, 1976).
[3]Carl de Boor and Blair Swartz, SIAM J. Numer. Anal. **10**, 582(1973).
[4]Ian H. Sloan, Math. Comp. **30**, 758 (1976).
[5]Ian H. Sloan, E. Naussair, and B. J. Burn, J. Math. Anal. Appl. **69**, 85 (1979).
[6]T. A. Osborn, Nucl. Phys. A **211**, 211 (1973).
[7]J. Horáček and L. Malina, Czech. J. Phys. B **27**, 134 (1972).
[8]H. R. Fiebig, Comp. Phys. Commun. **20**, 181 (1980).
[9]Y. E. Kim, J. Math. Phys. **10**, 1491 (1969); W. Glockle and R. Offerman, Phys. Rev. C **16**, 2039 (1977).
[10]F. Sohre and H. Ziegelmann, Phys. Lett. **34B**, 579 (1971); E. Harms, *ibid.* **41B**, 26 (1973); Nancy M. Larson and J. H. Hetherington, Phys. Rev. C **9**, 699 (1974).
[11]T. A. Osborn and D. Eyre, Nucl. Phys. A **327**, 125 (1979).
[12]G. L. Payne, J. L. Friar, B. F. Gibson, and I. R. Afnan, Phys. Rev. C **22**, 823 (1980).
[13]B. Noble and S. Beighton, J. Inst. Math. Appl. **26**, 431 (1980).
[14]M. G. Cox, J. Inst. Math. Appl. **10**, 134 (1972).
[15]Y. Yamaguchi, Phys. Rev. **95**, 1628 (1954).
[16]Roderick V. Reid, Jr., Ann. Phys. (N.Y.) **50**, 411 (1968).
[17]M. I. Haftel and F. Tabakin, Nucl. Phys. A **158**, 1 (1970).
[18]I. S. Gradsteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1980).

A refinement of the Thomas–Fermi approximation for the N body problem

Joseph G. Conlon
Department of Mathematics, University of Missouri, Columbia, Missouri 65211

(Received 3 February 1981; accepted for publication 7 October 1981)

The author suggests a refinement of the Thomas–Fermi approximation for the ground state energy E_N for an N electron atom. It is known that E_N can be written asymptotically as $N\to\infty$ as $E_N \sim \alpha N^{7/3}$, where α is given by Thomas–Fermi theory. It has been further conjectured that this asymptotic formula may be refined to $E_N \sim \alpha N^{7/3} + \beta N^{6/3} + \gamma N^{5/3}$. Suggestions for the contributions to $\beta N^{6/3}$ and $\gamma N^{5/3}$ have been made by Dirac and Von Weizsäcker. Here the author uses known results on short-time asymptotics for diffusion equations to obtain a refinement of the Thomas–Fermi approximation which includes the Dirac and Von Weizsäcker corrections. He also obtains new terms. These are related to the scalar curvature of the Jacobi metric corresponding to the Thomas–Fermi potential.

PACS numbers: 31.20.Lr

1. INTRODUCTION

We are concerned with finding approximations to the ground state energy for an N-electron atom where N is a large integer.

Suppose there are k nuclei with positive charges Z_i fixed at points $R_i \in \mathbb{R}^3$, respectively, $1 \leqslant i \leqslant k$. The total potential at a point $x \in \mathbb{R}^3$ due to the nuclei is $-V(x)$, where

$$V(x) = \sum_{i=1}^{k} \frac{Z_i}{|x - R_i|}. \tag{1.1}$$

Next we introduce N electrons, each with charge -1 and mass m, moving in the field of the potential $-V(x)$. Let $x_i \in \mathbb{R}^3$ be the position of the ith electron and $\sigma_i = \pm 1$ be its spin, $1 \leqslant i \leqslant N$. Then the N electron wave function ψ may be written as $\psi \equiv \psi(x_1,...,x_N; \sigma_1,...,\sigma_N)$, where $\psi \in L^2(\mathbb{R}^{3N}; \mathbb{C}^{2N})$. Let $\mathscr{H}_N$ be the Hilbert space of all such ψ which are antisymmetric in the (x_i, σ_i), $1 \leqslant i \leqslant N$. By Pauli exclusion $\mathscr{H}_N$ is the state space for the N electron system. The corresponding Hamiltonian H_N is given by

$$H_N = -h^2(8\pi^2 m)^{-1} \sum_{i=1}^{N} \Delta_i - \sum_{i=1}^{N} V(x_i) + \sum_{i<j=1}^{N} |x_i - x_j|^{-1}, \tag{1.2}$$

with h being Planck's constant and Δ_i the Laplacian in the x_i variable, $1 \leqslant i \leqslant N$. If $(\,,\,)$ denotes the inner product on $\mathscr{H}_N$ then the ground state energy E_N for the N electron system is

$$E_N = \inf\{(\psi, H_N\psi): \psi \in \mathscr{H}_N, \|\psi\| = 1\}. \tag{1.3}$$

One method of approximating E_N is to limit the class of functions $\psi \in \mathscr{H}_N$ over which one minimizes in (1.3). Thus by restricting ψ to antisymmetric products of single-particle wave functions $\psi_1(x_1, \sigma_1),...,\psi_N(x_N, \sigma_N)$, one obtains Hartree–Fock theory.[1] If $\psi_1,...,\psi_N \in \mathscr{H}_1$ form an orthonormal set of real functions and $\psi \in \mathscr{H}_N$ is the corresponding N electron wave function then $(\psi, H_N\psi) = \epsilon_{\mathrm{HF}}(\psi_1,...,\psi_N)$, where

$$\epsilon_{\mathrm{HF}} = K + A + R + Ex. \tag{1.4}$$

The kinetic energy K is given by

$$K(\psi_1,...,\psi_N) = h^2(8\pi^2 m)^{-1} \sum_{i=1}^{N} \sum_{\sigma=\pm 1} \int_{\mathbb{R}^3} [\nabla\psi_i(x, \sigma)]^2 dx. \tag{1.5}$$

The other terms on the right in (1.4) can be expressed as integrals of the two-body density $\rho(x, \sigma; y, \sigma')$, where

$$\rho(x, \sigma;\ y, \sigma') = \sum_{i=1}^{N} \psi_i(x, \sigma)\psi_i(y, \sigma'); x, y \in \mathbb{R}^3, \quad \sigma, \sigma' = \pm 1. \tag{1.6}$$

Thus letting $\rho(x)$ be the one-body density,

$$\rho(x) = \sum_{\sigma=\pm 1} \rho(x, \sigma;\ x, \sigma), \tag{1.7}$$

the potential energy A due to nuclear attraction is

$$A(\rho, V) = -\int_{\mathbb{R}^3} \rho(x)V(x)dx. \tag{1.8}$$

The potential energy R due to electron repulsion is

$$R(\rho) = \frac{1}{2} \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x - y|}\, dx\, dy. \tag{1.9}$$

The nonclassical exchange energy Ex is

$$\mathrm{Ex}(\psi_1,...,\psi_N) = -\frac{1}{2} \sum_{\sigma,\sigma'=\pm 1} \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \frac{[\rho(x, \sigma;\ y, \sigma')]^2}{|x - y|}\, dx\, dy. \tag{1.10}$$

The Hartree–Fock theory then yields the approximate value $E_N(\mathrm{HF})$ for the ground state energy, where

$$E_N(\mathrm{HF}) = \inf\{\epsilon_{\mathrm{HF}}(\psi_1,...,\psi_N): \psi_i \in \mathscr{H}_1, (\psi_i, \psi_j) = \delta_{ij}\}. \tag{1.11}$$

Evidently $E_N \leqslant E_N(\mathrm{HF})$.

As a further approximation one tries to express the kinetic and exchange energies in terms of the one-body density $\rho(x)$ alone. Hence, since A and R are already expressed in terms of $\rho(x)$ the total energy, ϵ_{HF} is a functional of $\rho(x)$ alone. To approximate E_N one then has just to minimize ϵ_{HF} over all functions $\rho(x)$ such that

$$\rho(x) \geqslant 0, \quad x \in \mathbb{R}^3, \quad \int_{\mathbb{R}^3} \rho(x)dx = N. \tag{1.12}$$

To approximate the kinetic energy in terms of $\rho(x)$ let us think of $\rho(x)$ as being a fixed function satisfying (1.12) and regard (1.7) as a constraint on the wave functions $\psi_1,...,\psi_N$. Let $K_{\min}(\rho)$ be the minimum kinetic energy (1.5) subject to the constraint (1.7). Then if we assume following Thomas

 0022-2488/83/010184-09$02.50

and Fermi[2] that a volume h^3 in the classical phase space can accomodate exactly two electrons we find for $K_{\min}(\rho)$ the value

$$K_{\min}(\rho) = \frac{3}{5} c \int_{\mathbb{R}^3} \rho(x)^{5/3}\, dx \tag{1.13}$$

with c given by

$$c = h^2(2m)^{-1}3^{2/3}(8\pi)^{-2/3}. \tag{1.14}$$

To approximate the exchange energy in terms of $\rho(x)$ one needs to express the two-body density (1.6) approximately in terms of $\rho(x)$. Dirac[3] achieved this by using a formal analogy between classical observables and quantum observables. His value for the two-body density $\rho(x,\sigma; y,\sigma')$ is

$$\rho(x, \sigma;\ y, \sigma') = \tfrac{1}{2}\rho(\xi)g(2|\eta|(3\pi^2\rho(\xi))^{1/3})\delta_{\sigma\sigma'}, \tag{1.15}$$

where $\xi = (x+y)/2$, $\eta = (x-y)/2$, $\delta_{\sigma\sigma'} = 0, 1$ according as $\sigma \neq \sigma'$ or $\sigma = \sigma'$, respectively. The function $g(z)$ is defined by

$$g(z) = z^{-3}[\sin z - z\cos z]. \tag{1.16}$$

Substituting (1.15) into (1.10) one obtains the exchange energy in terms of $\rho(x)$ as

$$\mathrm{Ex} = -3^{4/3}\pi^{-1/3}4^{-1}\int_{\mathbb{R}^3} \rho(x)^{4/3}dx. \tag{1.17}$$

It was not until recent years that the nature of the approximations (1.13) and (1.17) for kinetic and exchange energies was understood. In Ref. 4 it is shown that $K_{\min}(\rho)$ with a different value of the constant c is a lower bound for kinetic energy. Lieb and Oxford prove in Ref. 5 that the Dirac energy (1.17), again with a different constant, is a lower bound for exchange energy. Here we are more concerned with the results of Ref. 2 where it is shown that kinetic energy converges in a certain asymptotic sense to $K_{\min}(\rho)$ with the constant c of (1.14) as the number of electrons $N \to \infty$. To explain this we consider the Thomas–Fermi energy $\epsilon_{\mathrm{TF}}(\rho, V)$ defined by

$$\epsilon_{\mathrm{TF}}(\rho, V) = K_{\min}(\rho) + A(\rho, V) + R(\rho), \tag{1.18}$$

where $A(\rho, V)$ and $R(\rho)$ are given by (1.8) and (1.9), respectively. The corresponding minimum energy for the λ electron system, $\epsilon_\lambda(V)$, is

$$\epsilon_\lambda(V) = \inf\left\{\epsilon_{\mathrm{TF}}(\rho,V):\ \rho(x)\geqslant 0, \rho\in L^{5/3}(\mathbb{R}^3),\ \int_{\mathbb{R}^3}\rho(x)dx = \lambda\right\}. \tag{1.19}$$

It is known[2] that there is a unique minimizing ρ for (1.19) provided

$$\lambda \leqslant \sum_{i=1}^{k} Z_i = Z. \tag{1.20}$$

The minimizing $\rho(x)$ satisfies the Euler equation

$$c\rho(x)^{2/3} = \max[\phi_\rho(x) - \phi_0, 0], \tag{1.21}$$

where $\phi_\rho(x)$ is given by

$$\phi_\rho(x) = V(x) - \int_{\mathbb{R}^3} \frac{\rho(y)}{|x-y|}\, dy, \tag{1.22}$$

and ϕ_0 is a non-negative constant. Thus $-\phi_\rho(x)$ is the total electrostatic potential at x due to nuclei and electrons. The constant $-\phi_0$ may be interpreted as the maximum energy of an electron in the system. Let $\rho_{\mathrm{TF}}(x)$ be the minimizing function for (1.19), and for $N \geqslant 1$ let $V_N(x)$ be

$$V_N(x) = N^{4/3}V(N^{1/3}x). \tag{1.23}$$

Then $N^2\rho_{\mathrm{TF}}(N^{1/3}x)$ minimizes $\epsilon_{\mathrm{TF}}(\rho, V_N)$ subject to

$$\int_{\mathbb{R}^3}\rho(x)dx = \lambda N. \tag{1.24}$$

It follows that

$$\epsilon_{\lambda N}(V_N) = N^{7/3}\epsilon_\lambda(V). \tag{1.25}$$

Next let $E_{\lambda N}(V_N)$ be the ground state energy (1.3) for the λN electron system with nuclear potential $-V_N$. It is known[2] that

$$E_{\lambda N}(V_N) = N^{7/3}\epsilon_\lambda(V) + o(N^{7/3}), \quad N\to\infty. \tag{1.26}$$

Hence the Thomas–Fermi energy approximates the quantum mechanical energy in a definite asymptotic sense. Note that if $k = 1, \lambda = Z = 1$, then $E_{\lambda N}(V_N)$ is the minimum energy of an N electron atom.

It has been conjectured, based on calculations for the hydrogenic atom,[6] that the asymptotic formula (1.26) may be refined to

$$E_{\lambda N}(V_N) = N^{7/3}\epsilon_\lambda(V) + \alpha N^{6/3} + \beta N^{5/3} + o(N^{5/3}),\quad N\to\infty. \tag{1.27}$$

Thus on substituting $\rho(x) = N^2\rho_{\mathrm{TF}}(N^{1/3}x)$ into (1.17) we see that the exchange energy should make a contribution to β. Von Weizsäcker[7] argued that kinetic energy also makes a contribution to β by suggesting that K can be written more accurately in terms of $\rho(x)$ as

$$K = K_{\min}(\rho) + C_W\int_{\mathbb{R}^3}(\nabla\rho^{1/2})^2dx, \tag{1.28}$$

where C_W is a positive constant. The actual value of the constant C_W seems to be in some doubt.[8] If we substitute $\rho(x) = N^2\rho_{\mathrm{TF}}(N^{1/3}x)$ into (1.28) then it is easy to see that the Von Weizsäcker term is order $N^{5/3}$ provided

$$\int_{\mathbb{R}^3}(\nabla\rho_{\mathrm{TF}}^{1/2})^2dx < \infty. \tag{1.29}$$

The $\alpha N^{6/3}$ term was first suggested by Scott,[9] who claimed that α should have the same value as for the hydrogenic atom since the $N^{6/3}$ correction is caused by the singularities of the potential (1.1). Since the density $\rho_{\mathrm{TF}}(x)$ corresponding to (1.1) satisfies $\rho_{\mathrm{TF}}(x) \sim (Z_i/c|x-R_i|)^{3/2}$ as $x\to R_i$, $1\leqslant i\leqslant k$, the integral (1.29) is not finite. Therefore we might expect the Von Weizsäcker term to make a contribution to α as well as β. This has been shown to be the case by Lieb.[10]

In this paper we intend to pursue Dirac's idea of associating a two-body density with a one-body density function $\rho(x)$. Then this two-body density will be used to calculate the total electronic energy in terms of $\rho(x)$. Our starting point is the Thomas–Fermi equation (1.21). Classically an electron in the system moves under the potential $-\phi_\rho(x)$. Hence, if g is the Euclidean metric in $\mathbb{R}^3$, an electron with maximum energy moves along a geodesic in the Jacobi metric $[\phi_\rho(x) - \phi_0]g$. From (1.21) this metric is just $\rho(x)^{2/3}g$. Our idea is to choose the eigenfunctions of the Laplace operator on $\mathbb{R}^3$ in the metric $\rho(x)^{2/3}g$ to form the two-body density associated with $\rho(x)$.

Let $\rho(x)$ be a suitably smooth function which is positive for all $x\in\mathbb{R}^3$ and such that

$$\int_{\mathbb{R}^3}\rho(x)dx = N. \tag{1.30}$$

We regard $\mathbb{R}^3$ as a Riemannian manifold M with metric $\rho(x)^{2/3}g$. Hence the Laplace operator Q on M is given by

$$-Q = \rho(x)^{-1}\nabla\cdot(\rho(x)^{1/3}\nabla). \tag{1.31}$$

We assume that Q is essentially self-adjoint on M with pure point spectrum. Then Q has real eigenvalues $\lambda_j, j = 1, 2,...,$ with $0<\lambda_1<\lambda_2\leqslant\lambda_3\leqslant...$ and corresponding real eigenfunctions $\phi_j(x), j = 1, 2,...$. If N is an even integer we define a set of functions $\psi_1,...,\psi_N\in\mathscr{H}_1$ by

$$\psi_j(x,\sigma) = \rho(x)^{1/2}\phi_j(x)\delta_{\sigma,1},$$
$$\psi_{j+N/2}(x,\sigma) = \rho(x)^{1/2}\phi_j(x)\delta_{\sigma,-1}, \quad 1\leqslant j\leqslant N/2. \tag{1.32}$$

Since $\phi_1,...,\phi_{N/2}$ form an orthonormal set in $L^2(M)$ it follows that $\psi_1,...,\psi_N$ form an orthonormal set in $\mathscr{H}_1$. We then define the two-body density associated with $\rho(x)$ by (1.6).

Our aim here is to show that when we use the functions (1.32) to compute the Hartree–Fock energy (1.4) we obtain the semiclassical approximations (1.17) and (1.28) in a certain asymptotic sense as $N\to\infty$. In Sec. 2 we show that the Hartree–Fock energy approaches the Thomas–Fermi energy to order $N^{7/3}$. This corresponds to (1.26). We also show that the Hartree–Fock exchange energy approaches the Dirac exchange energy to order $N^{5/3}$. In Sec. 3 we derive heuristically a refinement of the Thomas–Fermi approximation which includes the Dirac and Von Weizsäcker corrections. We use an attractive nuclear potential which smooths the Coulomb singularity. Hence the refinement does not contain a term $\alpha N^{6/3}$ as in (1.27).

Our method of approach in this paper is to use known results on short time asymptotics for the heat equation associated with Q. In Sec. 2 we consider smooth compact manifolds for which rigorous results are known.[11] In Sec. 3 we assume that the results for the compact case extend to the noncompact case.

This work is a direct extension of the March and Young work[12] for the one-dimensional case. In the one-dimensional situation it is possible to write down the eigenfunctions of the Laplace operator explicitly,[13] which leads to considerable simplification.

2. THOMAS–FERMI AND DIRAC EXCHANGE ENERGIES

Let $\Omega\subset\mathbb{R}^3$ be a bounded open set with smooth boundary $\partial\Omega$ and

$$\rho:\ \Omega\to\mathbb{R} \tag{2.1}$$

be a C^∞ function on the closure $\bar{\Omega}$ of Ω such that $\rho(x)>0$, $x\in\bar{\Omega}$, and

$$\int_\Omega\rho(x)dx = 1. \tag{2.2}$$

We regard $\bar{\Omega}$ as a Riemannian manifold M with the metric $\rho(x)^{2/3}g$. Hence M is a C^∞ manifold of unit volume and with smooth boundary ∂M. The Laplace operator Q on M is defined by (1.31). Let Q_D be the self-adjoint extension of Q in $L^2(M)$, which corresponds to Dirichlet boundary conditions. For $t>0$ the operator $\exp(-tQ_D)$ is a real symmetric compact integral operator on $L^2(M)$ with kernel $G(x,y,t), x,y\in\Omega$. The function $G(x,y,t)$ is in $C^\infty(\Omega\times\Omega\times(0,\infty))$ and for each $t>0$ $G(x,y,t)$ is in $C^1(\bar{\Omega}\times\bar{\Omega})$ such that $G(x,y,t) = 0$ if $x\in\partial\Omega$. The operator Q_D has pure point spectrum with real eigenvalues $\lambda_j, j = 1,2,...,$ such that $0<\lambda_1<\lambda_2\leqslant\lambda_3\leqslant\cdots$. We write the corresponding real normalized eigenfunctions as $\phi_j(x)$, $j = 1,2,\cdots$. Each function $\phi_j(x)$ is in $C^\infty(\Omega)\cap C^1(\bar{\Omega})$ and $\phi_j(x) = 0$ if $x\in\partial\Omega$.

For positive even integers $N = 2, 4,...,$ we define a set of functions ${}_N\psi_1, {}_N\psi_2,...,{}_N\psi_N$, by

$${}_N\psi_j(x,\sigma) = N^{1/2}\rho(N^{1/3}x)^{1/2}\phi_j(N^{1/3}x)\delta_{\sigma,1},$$
$${}_N\psi_{j+N/2}(x,\sigma) = N^{1/2}\rho(N^{1/3}x)^{1/2}\phi_j(N^{1/3}x)\delta_{\sigma,-1},$$
$$1\leqslant j\leqslant N/2. \tag{2.3}$$

It is easy to see that the functions ${}_N\psi_j$, $1\leqslant j\leqslant N$, are the single-particle wave functions associated to the one-body density $N^2\rho(N^{1/3}x)$ by the prescription (1.32). We extend the ${}_N\psi_j(x,\sigma)$ to $\mathbb{R}^3$ by setting ${}_N\psi_j(x,\sigma) = 0$ if $x\in\mathbb{R}^3-\Omega$. Thus the functions ${}_N\psi_j$, $1\leqslant j\leqslant N$, form an orthonormal set in the space $\mathscr{H}_1$ and we can also see that each ${}_N\psi_j(x,\sigma)$ is in the Sobolev space $H^1(\mathbb{R}^3)$. Let $\rho_N(x)$ be the one-body density associated with the ${}_N\psi_j$, $1\leqslant j\leqslant N$, so

$$\rho_N(x) = \sum_{\sigma=\pm1}\sum_{j=1}^{N}[{}_N\psi_j(x,\sigma)]^2. \tag{2.4}$$

We prove the following result

Theorem 2.1:

(a) $\lim_{N\to\infty}N^{-7/3}K({}_N\psi_1,...,{}_N\psi_N) = K_{\min}(\rho)$,

(b) $\lim_{N\to\infty}N^{-7/3}A(\rho_N, V_N) = A(\rho, V)$,

(c) $\lim N^{-7/3}R(\rho_N) = R(\rho)$.

Proof: (a) We have

$$N^{-7/3}K({}_N\psi_1,...,{}_N\psi_N) = h^2(4\pi^2m)^{-1}N^{-5/3}\sum_{j=1}^{N/2}\int_\Omega[\nabla\rho(x)^{1/2}\phi_j(x)]^2dx. \tag{2.5}$$

Let $I(n)$ be given by

$$I(n) = \sum_{i=1}^{n}\int_\Omega[\nabla\rho(x)^{1/2}\phi_i(x)]^2dx. \tag{2.6}$$

On using the identity

$$\phi Q\phi = \tfrac{1}{2}Q\phi^2 + \rho^{-2/3}(\nabla\phi)^2, \tag{2.7}$$

we see that

$$I(n) = I_1(n) + I_2(n) + I_3(n) + I_4(n), \tag{2.8}$$

where

$$I_1(n) = \sum_{i=1}^{n}\int_\Omega\lambda_i\phi_i(x)^2\rho(x)^{5/3}dx, \tag{2.9}$$

$$I_2(n) = \sum_{i=1}^{n}\frac{1}{2}\int_\Omega Q\phi_i^2(x)\rho(x)^{5/3}dx, \tag{2.10}$$

$$I_3(n) = \sum_{i=1}^{n}\int_\Omega(\nabla\rho^{1/2})^2\phi_i(x)^2dx, \tag{2.11}$$

$$I_4(n) = \sum_{i=1}^{n} \frac{1}{2} \int_\Omega \nabla\rho \cdot \nabla \phi_i^2(x) dx. \tag{2.12}$$

We now use the well-known identity

$$G(x,x,t) = \sum_{i=1}^{\infty} e^{-\lambda_i t} \phi_i(x)^2, \quad x \in \Omega, \ t > 0. \tag{2.13}$$

We may differentiate (2.13) with respect to t to obtain

$$-\frac{\partial G}{\partial t}(x,x,t) = \sum_{i=1}^{\infty} e^{-\lambda_i t} \lambda_i \phi_i(x)^2. \tag{2.14}$$

If we integrate (2.14) against $\rho(x)^{5/3}$ we have

$$\int_\Omega -\frac{\partial G}{\partial t}(x,x,t)\rho(x)^{5/3} dx = \sum_{i=1}^{\infty} e^{-\lambda_i t} \int_\Omega \lambda_i \phi_i(x)^2 \rho(x)^{5/3} dx. \tag{2.15}$$

On using the asymptotic formula,[14]

$$-\frac{\partial G}{\partial t}(x,x,t) \sim \frac{3}{2t}(4\pi t)^{-3/2}, \quad t \to 0, \tag{2.16}$$

we see that

$$\lim_{t \to 0} t^{5/2} \int_\Omega -\frac{\partial G}{\partial t}(x,x,t)\rho(x)^{5/3} dx = \frac{3}{16} \pi^{-3/2} \int_\Omega \rho(x)^{5/3} dx. \tag{2.17}$$

Let $n(\lambda)$ be given by

$$n(\lambda) = \#\{\lambda_j : \lambda_j \leqslant \lambda\}. \tag{2.18}$$

Then Weyl's theorem[15] yields

$$\lim_{\lambda \to \infty} \lambda^{-3/2} n(\lambda) = (6\pi^2)^{-1}. \tag{2.19}$$

From (2.15) and (2.17) we obtain via the Karamata Tauberian theorem[16]

$$\lim_{\lambda \to \infty} \lambda^{-5/2} I_1(n(\lambda)) = (10\pi^2)^{-1} \int_\Omega \rho(x)^{5/3} dx. \tag{2.20}$$

This asymptotic formula may be rewritten using (2.19) as

$$\lim_{n \to \infty} n^{-5/3} I_1(n) = \frac{3}{5} 6^{2/3} \pi^{4/3} \int_\Omega \rho(x)^{5/3} dx. \tag{2.21}$$

Next we define $I_5(n)$ by

$$I_5(n) = \sum_{i=1}^{n} \frac{1}{2} \int_\Omega \phi_i^2(x) |Q\rho^{2/3}(x)| \rho(x) dx. \tag{2.22}$$

It is obvious from (2.10) on integration by parts that

$$|I_2(n)| \leqslant I_5(n). \tag{2.23}$$

From (2.13) we have

$$\frac{1}{2} \int_\Omega G(x,x,t) |Q\rho^{2/3}(x)| \rho(x) dx = \sum_{i=1}^{\infty} e^{-\lambda_i t} \frac{1}{2} \int_\Omega \phi_i^2(x) |Q\rho^{2/3}(x)| \rho(x) dx. \tag{2.24}$$

If we use the asymptotic formula[14]

$$G(x,x,t) \sim (4\pi t)^{-3/2}, \quad t \to 0, \tag{2.25}$$

we have that

$$\lim_{t \to 0} t^{3/2} \int_\Omega \frac{1}{2} G(x,x,t) |Q\rho^{2/3}(x)| \rho(x) dx = \frac{\pi^{-3/2}}{16} \int_\Omega |Q\rho^{2/3}(x)| \rho(x) dx. \tag{2.26}$$

Thus from (2.23), (2.24), and (2.26) and Karamata's Tauberian theorem we conclude that

$$\lim_{n \to \infty} n^{-5/3} I_2(n) = 0. \tag{2.27}$$

Similarly we conclude that

$$\lim_{n \to \infty} n^{-5/3} I_3(n) = 0, \tag{2.28}$$

$$\lim_{n \to \infty} n^{-5/3} I_4(n) = 0. \tag{2.29}$$

The conclusion of (a) then follows from (2.5), (2.8), (2.21), and (2.27)–(2.29). (b) We observe that

$$N^{-7/3} A(\rho_N, V_N) = 2N^{-1} \sum_{j=1}^{N/2} \int_\Omega V(x)\rho(x)\phi_j^2(x) dx. \tag{2.30}$$

We put

$$J(n) = \sum_{i=1}^{n} \int_\Omega V(x) \rho(x) \phi_i^2(x) dx. \tag{2.31}$$

Suppose that Ω contains the singularities $R_1, \ldots, R_k$ of $V(x)$ in its interior. From (2.13) we have

$$\int_\Omega G(x,x,t) V(x)\rho(x) dx = \sum_{i=1}^{\infty} e^{-\lambda_i t} \int_\Omega V(x)\rho(x)\phi_i^2(x) dx. \tag{2.32}$$

From (2.25) it follows that

$$\lim_{t \to 0} t^{3/2} \int_\Omega G(x,x,t) V(x)\rho(x) dx = \frac{\pi^{-3/2}}{8} \int_\Omega V(x)\rho(x) dx. \tag{2.33}$$

By the Karamata theorem it follows from (2.32) and (2.33) that

$$\lim_{\lambda \to \infty} \lambda^{-3/2} J(n(\lambda)) = (6\pi^2)^{-1} A(\rho, V). \tag{2.34}$$

Hence from (2.19) we have

$$\lim_{n \to \infty} n^{-1} J(n) = A(\rho, V), \tag{2.35}$$

from which (b) follows. Before turning to (c) we prove a slight generalization of the Karamata theorem.

Lemma 2.2: Let m be a positive Borel measure on the quadrant $[0,\infty) \times [0,\infty)$ such that

$$\int_0^\infty \int_0^\infty e^{-(\alpha u + \beta v)} dm(u,v) < \infty \tag{2.36}$$

for $\alpha, \beta > 0$, and $b(u,v)$ be a nonnegative Borel measurable function on $[0,\infty) \times [0,\infty)$ with

$$\int_0^\infty \int_0^\infty e^{-(\alpha u + \beta v)} b(u,v) du\, dv < \infty \tag{2.37}$$

if $\alpha, \beta > 0$. Suppose there is a $\gamma > 0$ such that

$$\lim_{t\to 0} t^{\gamma}\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)t}dm(u,v)$$
$$=\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)}b(u,v)du\,dv \tag{2.38}$$

for all $\alpha,\beta>0$. Then

$$\lim_{a\to\infty} a^{-\gamma}m\{[0,a]\times[0,a]\}=\int_0^1\int_0^1 b(u,v)du\,dv. \tag{2.39}$$

Proof: Define a family of measures m_t, $t>0$, by

$$m_t(A)=t^{\gamma}m(t^{-1}A), \tag{2.40}$$

where $A\subset\mathbb{R}^3$ is a Borel set. Then (2.38) becomes

$$\lim_{t\to 0}\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)}dm_t(u,v)$$
$$=\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)}b(u,v)du\,dv \tag{2.41}$$

for all $\alpha,\beta>0$. Since the family of measures

$$e^{-(u+v)}dm_t(u,v),\quad t>0, \tag{2.42}$$

is uniformly bounded we can conclude from (2.41) and the Stone–Weierstrasse theorem that for every continuous function f: $[0,\infty)\times[0,\infty)\to\mathbb{R}^3$ which disappears at ∞,

$$\lim_{t\to 0}\int_0^\infty\int_0^\infty f(u,v)e^{-(u+v)}dm_t(u,v)$$
$$=\int_0^\infty\int_0^\infty f(u,v)e^{-(u+v)}b(u,v)du\,dv. \tag{2.43}$$

From (2.43) it is easy to see that

$$\lim_{t\to 0}\int_0^1\int_0^1 dm_t(u,v)=\int_0^1\int_0^1 b(u,v)du\,dv, \tag{2.44}$$

and (2.44) is equivalent to (2.39).

Proof of (c): We have

$$N^{-7/3}R(\rho_N)$$
$$=4N^{-2}\int_\Omega\int_\Omega\rho(x)\rho(y)\sum_{i,j=1}^{N/2}\phi_i^2(x)\phi_j^2(y)\frac{dx\,dy}{|x-y|}. \tag{2.45}$$

For $\alpha,\beta,t>0$, we consider

$$\int_\Omega\int_\Omega\rho(x)\rho(y)G(x,x,\alpha t)G(y,y,\beta t)\frac{dx\,dy}{|x-y|}$$
$$=\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)t}dm(u,v), \tag{2.46}$$

where

$$m(A)=\sum_{(\lambda_i,\lambda_j)\in A}\int_\Omega\int_\Omega\rho(x)\rho(y)\phi_i^2(x)\phi_j^2(y)\frac{dx\,dy}{|x-y|}, \tag{2.47}$$

for every Borel set $A\subset\mathbb{R}^2$. From (2.25) we conclude that

$$\lim_{t\to 0} t^3\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)t}dm(u,v)$$
$$=(\alpha\beta)^{-3/2}(4\pi)^{-3}R(\rho). \tag{2.48}$$

Using the fact that

$$\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)}(uv)^{1/2}du\,dv=\frac{\pi}{4}(\alpha\beta)^{-3/2} \tag{2.49}$$

and Lemma 2.2, we see that

$$\lim_{\lambda\to\infty}\lambda^{-3}m\{[0,\lambda]\times[0,\lambda]\}=(6\pi^2)^{-2}R(\rho). \tag{2.50}$$

Then (c) follows from (2.50) and (2.19).

As a corollary to Theorem 2.1 we may prove part of (1.26).

Corollary 2.3:

$$\overline{\lim_{N\to\infty}} N^{-7/3}E_{\lambda N}(V_N)\leqslant\epsilon_\lambda(V),$$

where N goes to ∞ through integer values of λN.

Proof: Since each function ${}_N\psi_j(x,\sigma)$ of (2.3) is in $H^1(\mathbb{R}^3)$ and the exchange energy is negative we conclude from Theorem 2.1 that

$$\overline{\lim_{N\to\infty}} N^{-7/3}E_{\lambda N}(V_N)$$
$$\leqslant\lambda^{7/3}[K_{\min}(\rho)+A(\rho,V_{1/\lambda})+R(\rho)]$$
$$=\epsilon_{\rm TF}(\rho^\lambda,V), \tag{2.51}$$

where $\rho^\lambda(x)$ is defined by

$$\rho^\lambda(x)=\lambda^2\rho(\lambda^{1/3}x). \tag{2.52}$$

One can easily see that it is possible to choose a sequence of functions $\rho_i(x)$, $i=1,2,\ldots$, satisfying (2.1) and (2.2) such that

$$\lim_{i\to\infty}\epsilon_{\rm TF}(\rho_i^\lambda,V)=\epsilon_\lambda(V). \tag{2.53}$$

This completes the proof of the corollary.

Next we consider the exchange energy.

Theorem 2.4:

$$\lim_{N\to\infty} N^{-5/3}{\rm Ex}({}_N\psi_1,\ldots,{}_N\psi_N)=-3^{4/3}\pi^{-1/3}4^{-1}X(\rho),$$

where $X(\rho)$ is given by

$$X(\rho)=\int_{\mathbb{R}^3}\rho(x)^{4/3}dx. \tag{2.54}$$

Proof: We have

$$N^{-5/3}{\rm Ex}({}_N\psi_1,\ldots,{}_N\psi_N)$$
$$=-N^{-4/3}\int_\Omega\int_\Omega\rho(x)\rho(y)\sum_{i,j=1}^{N/2}\phi_i(x)\phi_i(y)\phi_j(x)\phi_j(y)\frac{dx\,dy}{|x-y|}. \tag{2.55}$$

For $t>0$ the Green's function $G(x,y,t)$ is given by

$$G(x,y,t)=\sum_{i=1}^\infty e^{-\lambda_i t}\phi_i(x)\phi_i(y), \tag{2.56}$$

where the series on the right is absolutely convergent. Hence if $\alpha,\beta>0$, then

$$\int_\Omega\int_\Omega\rho(x)\rho(y)G(x,y,\alpha t)G(x,y,\beta t)\frac{dx\,dy}{|x-y|}$$
$$=\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)t}dm(u,v), \tag{2.57}$$

where

$$m(A)=\sum_{(\lambda_i,\lambda_j)\in A}\int_\Omega\int_\Omega\rho(x)\rho(y)\phi_i(x)\phi_i(y)\phi_j(x)\phi_j(y)\frac{dx\,dy}{|x-y|}, \tag{2.58}$$

for every Borel set $A\subset\mathbb{R}^2$. Since the function $|x|^{-1}$ is positive

definite it follows that m is a positive measure on $[0,\infty)\times[0,\infty)$. We now use the asymptotic formula[11] for $G(x,y,t)$,

$$G(x,y,t)\sim(4\pi t)^{-3/2}\exp\left[-\frac{d^2(x,y)}{4t}\right],\quad t\to 0. \tag{2.59}$$

Here $d(x,y)$ is the Riemannian distance from x to y in the metric $\rho(x)^{2/3}g$. Thus for y close to x we have

$$d^2(x,y)\sim\rho(x)^{2/3}|x-y|^2. \tag{2.60}$$

We conclude that

$$\begin{aligned}\lim_{t\to 0} t^2\int_\Omega\int_\Omega \rho(x)\rho(y)G(x,y,\alpha t)G(x,y,\beta t)\frac{dx\,dy}{|x-y|}&\\ =\lim_{t\to 0} t^{-1}(4\pi)^{-2}(\alpha\beta)^{-3/2}\int_\Omega\int_0^\infty \rho(x)^2&\\ \times\exp\left[-\frac{\rho^{2/3}(x)r^2}{4t}(\alpha^{-1}+\beta^{-1})\right]r\,dr\,dx&\\ =\frac{1}{8\pi^2}\frac{1}{(\alpha\beta)^{1/2}(\alpha+\beta)}X(\rho).&\end{aligned} \tag{2.61}$$

In order to apply Lemma 2.2 we need to find a positive function $f(u,v)$ on $[0,\infty)\times[0,\infty)$ such that

$$\int_0^\infty\int_0^\infty e^{-(\alpha u+\beta v)}f(u,v)du\,dv=(\alpha\beta)^{-1/2}(\alpha+\beta)^{-1} \tag{2.62}$$

for all $\alpha,\beta>0$. We put

$$F_\alpha(v)=\int_0^\infty e^{-\alpha u}f(u,v)du. \tag{2.63}$$

Hence (2.62) becomes

$$\mathcal{L}F_\alpha(\beta)=(\alpha\beta)^{-1/2}(\alpha+\beta)^{-1},\quad \beta>0 \tag{2.64}$$

where $\mathcal{L}$ denotes Laplace transform. On using the convolution theorem for Laplace transforms we conclude that

$$F_\alpha(v)=\int_0^v e^{-\alpha(v-w)}(\pi\alpha w)^{-1/2}dw. \tag{2.65}$$

For $0<w<v<\infty$ and $u>0$ let $h_{v,w}(u)$ be defined by

$$\begin{aligned}h_{v,w}(u)&=\pi^{-1}\left[w(u-\overline{v-w})\right]^{-1/2}\quad\text{if } u>\overline{v-w},\\ &=0\quad\text{if } u\leqslant\overline{v-w}.\end{aligned} \tag{2.66}$$

It is evident that

$$\mathcal{L}h_{v,w}(\alpha)=e^{-\alpha(v-w)}(\pi\alpha w)^{-1/2},\quad \alpha>0. \tag{2.67}$$

If we put $f_v(u)=f(u,v)$ then (2.63) may be written as

$$F_\alpha(v)=\mathcal{L}f_v(\alpha),\quad \alpha>0. \tag{2.68}$$

From (2.65) and (2.67) we deduce that

$$f_v(u)=\int_0^v h_{v,w}(u)dw. \tag{2.69}$$

Evaluating the integral in (2.69) we obtain

$$f(u,v)=\frac{2}{\pi}\ln\left[\frac{u^{1/2}+v^{1/2}}{|u-v|^{1/2}}\right], \tag{2.70}$$

and an elementary calculation yields

$$\int_0^1\int_0^1 f(u,v)du\,dv=\frac{2}{\pi}. \tag{2.71}$$

Now from (2.57), (2.58), (2.61), and Lemma 2.2 we conclude that

$$\lim_{\lambda\to\infty}\lambda^{-2}m\{[0,\lambda]\times[0,\lambda]\}=(4\pi^3)^{-1}X(\rho). \tag{2.72}$$

Hence from (2.19) and (2.55) we obtain the result of the theorem.

3. REFINEMENT OF THOMAS–FERMI THEORY

We have seen in Sec. 2 that by choosing the functions (1.32) to form the two body density we obtain the semiclassical approximations (1.13) and (1.17). Here we again use these functions to obtain the refinement (1.27) of the Thomas–Fermi approximation.

We consider a modified atomic potential $V(x)$ defined by

$$V(x)=\int_{\mathbb{R}^3}\frac{h(y)}{|x-y|}dy, \tag{3.1}$$

where h: $\mathbb{R}^3\to\mathbb{R}$ is a non-negative C^∞ function which is spherically symmetric and has support in a region $|x|\leqslant\epsilon$ such that

$$\int_{|x|<\epsilon}h(x)dx=1. \tag{3.2}$$

Thus $V(x)$ is a C^∞ spherically symmetric function such that $V(x)=|x|^{-1}$ if $|x|\geqslant\epsilon$. Let $\rho_{\mathrm{TF}}(x)$ be the neutral Thomas–Fermi density associated with $V(x)$. Hence $\rho_{\mathrm{TF}}(x)$ satisfies (1.21) and (1.22) with $\phi_0=0$. If ϵ is sufficiently small then $\rho_{\mathrm{TF}}(x)$ is a strictly positive C^∞ function in $\mathbb{R}^3$ and is spherically symmetric. We regard $\mathbb{R}^3$ as a Riemannian manifold M with the metric $\rho_{\mathrm{TF}}(x)^{2/3}g$. Let Q be the Laplace operator on M and Q_D be the Friederichs' extension of Q to $L^2(M)$.

Theorem 3.1: Q_D has pure point spectrum.

Proof: Since $\rho_{\mathrm{TF}}(x)$ is spherically symmetric we may by introducing spherical harmonics reduce the problem to a one-dimensional one. For $l=0,1,2,\ldots$, let Q^l be the operator on functions with domain $0<r<\infty$ defined by

$$Q^l=\frac{-1}{r^2\rho_{\mathrm{TF}}(r)}\frac{1}{dr}\left[\rho_{\mathrm{TF}}(r)^{1/3}r^2\frac{d}{dr}\right]+\frac{l(l+1)}{\rho_{\mathrm{TF}}(r)^{2/3}r^2}. \tag{3.3}$$

Evidently Q^l is formally self-adjoint and positive on the space $L^2[(0,\infty),\rho_{\mathrm{TF}}(r)r^2dr]$. Let Q^l_D be the Friederichs' extension of Q^l. We need to show that Q^l_D has pure point spectrum for $l=0,1,2,\ldots$.

We proceed in the standard manner.[17] We make a change of variable $r\leftrightarrow s$ given by

$$\frac{ds}{dr}=r^2\rho_{\mathrm{TF}}(r). \tag{3.4}$$

In the s variable the operator Q^l becomes

$$A^l=\frac{-d}{ds}\left[p(s)\frac{d}{ds}\right]+q_l(s), \tag{3.5}$$

where

$$p(s)=r^4\rho_{\mathrm{TF}}(r)^{4/3}, \tag{3.6}$$

$$q_l(s)=l(l+1)\rho_{\mathrm{TF}}(r)^{-2/3}r^{-2}. \tag{3.7}$$

Defining s_∞ by

$$s_\infty = \int_0^\infty r^2 \rho_{\mathrm{TF}}(r) dr, \tag{3.8}$$

we see that A^l is formally self-adjoint on $L^2(0, s_\infty)$ and its Friederichs' extension A^l_D is unitarily equivalent to Q^l_D.

We make a further change of variable $s \leftrightarrow t$ by

$$\frac{ds}{dt} = p(s)^{1/2}. \tag{3.9}$$

Let t_∞ be given by

$$t_\infty = \int_0^{s_\infty} p(s)^{-1/2} ds. \tag{3.10}$$

We define a unitary transformation $\mathscr{U}$ from $L^2(0, s_\infty)$ onto $L^2(0, t_\infty)$ by

$$\mathscr{U} g(t) = g(s(t)) s'(t)^{1/2}, \quad g \in L^2(0, s_\infty). \tag{3.11}$$

In the t variable the operator A^l becomes

$$B^l = -\frac{d^2}{dt^2} + v_l(t), \tag{3.12}$$

where

$$v_l(t) = q_l(s) - \tfrac{1}{16}[p'(s)^2/p(s)] + \tfrac{1}{4} p''(s). \tag{3.13}$$

Thus B^l is formally self-adjoint on $L^2(0, t_\infty)$ and its Friederichs' extension B^l_D is unitarily equivalent to Q^l_D.

We show that B^l is essentially self-adjoint with pure point spectrum if $l \geqslant 1$. To do this we need to use the fact[18] that

$$\lim_{r \to \infty} \left(\frac{d}{dr} \right)^m \rho_{\mathrm{TF}}(r) \sim 27\pi^{-3} \left(\frac{d}{dr} \right)^m r^{-6}, \quad m = 0,1,2, \tag{3.14}$$

where we have taken the constant c in (1.21) to be 1. From (3.14) we deduce that $t_\infty < \infty$ and that

$$v_l(t) \sim l(l+1)(t_\infty - t)^{-2}, \quad t \to t_\infty. \tag{3.15}$$

Using the fact that $\rho_{\mathrm{TF}}(r)$ is C^∞ at $r = 0$ we also see that

$$v_l(t) \sim [l(l+1) + 1] t^{-2}, \quad t \to 0. \tag{3.16}$$

From (3.15) and (3.16) we see by an application of Theorem 6.23 in Chap. 13 of Ref. 17 that B^l is essentially self-adjoint if $l > 0$. By Theorem 7.17 in Chap. 13 of Ref. 17 it follows that B^l_D has pure point spectrum for $l > 0$.

We must deal with the case of B^l for $l = 0$ separately. In that case we apply Theorem 5 of Ref. 18 to deduce that

$$\lim_{t \to t_\infty} \sup |(t_\infty - t)^2 v_0(t)| = 0. \tag{3.17}$$

It follows then from Theorem 6.23 of Ref. 17 that B^0 has two boundary values at t_∞. Hence by Theorem 6.12 in Chap. 13 of Ref. 17 the endpoint t_∞ does not contribute to the essential spectrum. Just as for $l \geqslant 1$, neither does the endpoint 0 and so we conclude that B^0_D has pure point spectrum. This proves the theorem.

Corollary 3.2: Let $\phi(x)$ be an eigenfunction of Q_D and put $\psi(x) = [\rho_{\mathrm{TF}}(x)]^{1/2} \phi(x)$. Then $\psi(x)$ is in the Sobolev space $H^1(\mathbb{R}^3)$.

Proof: Since $\phi(x)$ is an eigenfunction of Q_D it follows that

$$\int_{\mathbb{R}^3} \rho_{\mathrm{TF}}(x) \phi(x)^2 dx < \infty, \tag{3.18}$$

$$\int_{\mathbb{R}^3} \rho_{\mathrm{TF}}(x)^{1/3} [\nabla \phi(x)]^2 dx < \infty. \tag{3.19}$$

The result now follows from (3.14).

We wish to investigate the behavior of the eigenfunctions $\phi(x)$ of Q_D as $x \to \infty$. To do this we need some lemmas.

Lemma 3.3: Let q: $(0,1] \to \mathbb{R}$ be a continuous function such that

$$\lim_{t \to 0} t^2 q(t) = \alpha, \tag{3.20}$$

where $\alpha > \frac{3}{4}$, and $u(t)$, $0 < t < 1$, be a solution of the equation

$$u''(t) = q(t) u(t) \tag{3.21}$$

such that

$$\int_0^1 u(t)^2 dt < \infty. \tag{3.22}$$

Then there is a constant $A > 0$ such that

$$|u'(t)| \leqslant A, \quad |u(t)| \leqslant At, \quad 0 < t < 1. \tag{3.23}$$

Proof: Choose ϵ with $0 < \epsilon < 1$ such that

$$t^2 q(t) \geqslant \tfrac{3}{4}, \quad 0 < t \leqslant \epsilon. \tag{3.24}$$

Suppose $u(\epsilon) > 0$, $u'(\epsilon) < 0$. From (3.21) we see that $u(t)$ is convex for $0 < t \leqslant \epsilon$ and so $u(t) > 0$, $0 < t \leqslant \epsilon$. Let $v(t)$, $0 < t \leqslant \epsilon$, satisfy the equation

$$v''(t) = \tfrac{3}{4} t^{-2} v(t), \tag{3.25}$$

with the initial condition

$$v(\epsilon) = u(\epsilon), \quad v'(\epsilon) = u'(\epsilon). \tag{3.26}$$

It is easy to see that

$$0 < v(t) \leqslant u(t), \quad 0 < t < \epsilon. \tag{3.27}$$

We may solve (3.25) explicitly to obtain

$$v(t) = c_1 t^{3/2} + c_2 t^{-1/2}, \tag{3.28}$$

where c_1 and c_2 are constants. Since $v(\epsilon)$ and $v'(\epsilon)$ have opposite signs we must have $c_2 \neq 0$. Consequently

$$\int_0^\epsilon v(t)^2 dt = \infty, \tag{3.29}$$

and from (3.27) it follows that (3.22) does not hold.

We may therefore assume that $u(t) > 0$, $u'(t) > 0$, $0 < t \leqslant \epsilon$. In view of (3.21) $u'(t)$ is increasing for $0 < t < \epsilon$. Hence $\lim_{t \to 0} u'(t)$ exists. From this fact we may easily deduce the inequalities (3.23).

Lemma 3.4: Let q: $(0,1] \to \mathbb{R}$ be a continuous function such that

$$\lim_{t \to 0} t^2 q(t) = 0 \tag{3.30}$$

and $u(t)$, $0 < t < 1$, be a solution of

$$u''(t) = q(t) u(t). \tag{3.31}$$

Then for any $\delta > 0$ the following inequalities hold:

$$\lim_{t \to 0} \sup |t^\delta u(t)| < \infty, \tag{3.32}$$

$$\lim_{t \to 0} \sup |t^{1+\delta} u'(t)| < \infty. \tag{3.33}$$

Proof: Similar to Lemma 3.3.

Theorem 3.5: Let $\phi(x)$ be an eigenfunction of Q_D. Then $\phi(x)$ ia a C^∞ function such that for any $\delta > 0$ the following inequalities hold:

$$\lim_{x\to\infty} \sup |x|^{-1-\delta}|\phi(x)| < \infty, \tag{3.34}$$

$$\lim_{x\to\infty} \sup |x|^{-\delta}|\nabla\phi(x)| < \infty. \tag{3.35}$$

Proof: Follows from Lemmas 3.3 and 3.4.

In order to obtain the refinement of Thomas–Fermi theory we must make several important assumptions. For $t > 0$ the operator $\exp(-tQ_D)$ is a real symmetric compact operator on $L^2(M)$. We shall suppose it is an integral operator with kernel $G(x,y,t), x,y \in \mathbb{R}^3$, where $G(x,y,t)$ is in $C^\infty(\mathbb{R}^3\times\mathbb{R}^3\times(0,\infty))$ and is a fundamental solution for the heat equation associated with Q. It follows[14] that $G(x,x,t)$ behaves asymptotically, as $t\to 0$, like

$$G(x,x,t) \sim (4\pi t)^{-3/2}[1 + (t/3)\kappa(x) + o(t)], \tag{3.36}$$

for any fixed $x \in \mathbb{R}^3$. Here $\kappa(x)$ is the scalar curvature of M, which turns out to be

$$\kappa(x) = 4\rho_{\mathrm{TF}}(x)^{-1}[\rho_{\mathrm{TF}}(x)^{1/6}\Delta\rho_{\mathrm{TF}}(x)^{1/6}]. \tag{3.37}$$

Similarly for fixed $x \in \mathbb{R}^3$, $\partial G/\partial t(x,x,t)$ is given asymptotically as $t\to 0$ by

$$\frac{\partial G}{\partial t}(x,x,t) \sim (4\pi t)^{-3/2}\left[\frac{-3}{2t} - \tfrac{1}{6}\kappa(x) + o(1)\right]. \tag{3.38}$$

Next we assume that the asymptotic formulas (3.36) and (3.38) are uniform in x to the extent that we may integrate (3.36) and (3.38) against a continuous function $f(x)$ over $\mathbb{R}^3$, where $f(x)$ decays like $|x|^{-6}$ as $x\to\infty$. In particular we have from (3.36) the formula

$$\int_{\mathbb{R}^3} G(x,x,t)\rho_{\mathrm{TF}}(x)dx \sim (4\pi t)^{-3/2}[1 - (4t/3)J + o(t)], \tag{3.39}$$

where J is given by

$$J = \int_{\mathbb{R}^3} [\nabla\rho_{\mathrm{TF}}(x)^{1/6}]^2 dx. \tag{3.40}$$

As in Sec. 2 let Q_D have real eigenvalues $\lambda_j, j = 1,2,\ldots$, with corresponding real eigenfunctions $\phi_j(x)$. Then (3.39) is equivalent to

$$\sum_{j=1}^{\infty} e^{-\lambda_j t} \sim (4\pi t)^{-3/2}[1 - (4t/3)J + o(t)]. \tag{3.41}$$

If $n(\lambda)$ is defined as in (2.18) then (3.41) and the Karamata Tauberian theorem yields (2.19). A conjecture of Weyl[15] suggests that we may separate out the second term in the asymptotic expansion of $n(\lambda)$. In that case $n(\lambda)$ must be given by

$$n(\lambda) \sim (6\pi^2)^{-1}[\lambda^{3/2} - 2J\lambda^{1/2} + o(\lambda^{1/2})], \quad \lambda\to\infty. \tag{3.42}$$

We shall assume that (3.42) holds.

Now with the notation of Sec. 2 and taking $\rho = \rho_{\mathrm{TF}}$, $\Omega = \mathbb{R}^3$, we consider

$$K(_N\psi_1,\ldots,_N\psi_N) = h^2(4\pi^2 m)^{-1}N^{2/3}I(N/2). \tag{3.43}$$

The function $I(n)$ may be written as the sum (2.8) with $I_1(n)$ and $I_3(n)$ given by (2.9) and (2.11), respectively. From Theorem 3.5 we may integrate by parts in (2.10) and (2.12) to obtain

$$I_2(n) = \sum_{i=1}^{n} \frac{1}{2}\int_{\mathbb{R}^3} \phi_i^2(Q\rho_{\mathrm{TF}}{}^{2/3})\rho_{\mathrm{TF}}(x)dx, \tag{3.44}$$

$$I_4(n) = -\sum_{i=1}^{n} \frac{1}{2}\int_{\mathbb{R}^3} \phi_i^2\Delta\rho_{\mathrm{TF}}(x)dx. \tag{3.45}$$

We define Borel measures m, M on $\mathbb{R}^+$ by

$$m[0,\lambda] = I_2(n(\lambda)), \tag{3.46}$$

$$M[0,\lambda] = I_5(n(\lambda)), \tag{3.47}$$

with I_5 as in (2.22). It is evident that M is a positive measure and that the signed measure m satisfies

$$|m| \leqslant M. \tag{3.48}$$

Further, by our assumption on the uniformity of the asymptotic formula (3.36) we have

$$\lim_{t\to 0} t^{3/2}\int_0^\infty e^{-\lambda t} dm(\lambda) = (4\pi)^{-3/2}\frac{1}{2}\int_{\mathbb{R}^3} (Q\rho_{\mathrm{TF}}{}^{2/3})\rho_{\mathrm{TF}}(x)dx = 0, \tag{3.49}$$

$$\lim_{t\to 0} t^{3/2}\int_0^\infty e^{-\lambda t} dM(\lambda) = (4\pi)^{-3/2}\frac{1}{2}\int_{\mathbb{R}^3} |Q\rho_{\mathrm{TF}}^{2/3}|\rho_{\mathrm{TF}}(x)dx. \tag{3.50}$$

We would like to conclude from (3.49) and the Karamata theorem that

$$\lim_{\lambda\to\infty} \lambda^{-3/2} m[0,\lambda] = 0. \tag{3.51}$$

Since m is not a positive measure we cannot apply the Karamata theorem directly. However, from (3.48), (3.49), (3.50), and the proof of the Karamata theorem we see that (3.51) holds. Thus

$$\lim_{\lambda\to\infty} \lambda^{-3/2} I_2(n(\lambda)) = 0. \tag{3.52}$$

Similarly we see that

$$\lim_{\lambda\to\infty} \lambda^{-3/2} I_4(n(\lambda)) = 0, \tag{3.53}$$

and by direct application of the Karamata theorem we conclude that

$$\lim_{\lambda\to\infty} \lambda^{-3/2} I_3(n(\lambda)) = (6\pi^2)^{-1}\int_{\mathbb{R}^3} (\nabla\rho_{\mathrm{TF}}^{1/2})^2 dx. \tag{3.54}$$

We wish to estimate $I_1(n(\lambda))$ as $\lambda\to\infty$ to order $\lambda^{3/2}$. To do this we use the identity

$$\int_0^\infty e^{-\lambda t} dI_1(n(\lambda)) = \int_{\mathbb{R}^3} \frac{-\partial G}{\partial t}(x,x,t)\rho_{\mathrm{TF}}(x)^{5/3}dx. \tag{3.55}$$

Making the assumption that the asymptotic formula (3.38) is uniform in x we have from (3.55)

$$\int_0^\infty e^{-\lambda t} dI_1(n(\lambda))$$
$$\sim \tfrac{3}{16}\pi^{-3/2} t^{-5/2} \int_{\mathbf{R}^3} \rho_{\mathrm{TF}}(x)^{5/3} dx$$
$$- \tfrac{5}{108}\pi^{-3/2} t^{-3/2} L + o(t^{-3/2}), \tag{3.56}$$

where L is given by

$$L = \int_{\mathbf{R}^3} (\nabla \rho_{\mathrm{TF}}{}^{1/2})^2 dx. \tag{3.57}$$

In analogy with the Weyl conjecture which led to (3.42) we shall assume by virtue of (3.56) that $I_1(n(\lambda))$ is given to order $\lambda^{3/2}$ by

$$I_1(n(\lambda)) \sim (10\pi^2)^{-1} \lambda^{5/2} \int_{\mathbf{R}^3} \rho_{\mathrm{TF}}(x)^{5/3} dx$$
$$- \tfrac{5}{81}\pi^{-2} L \lambda^{3/2} + o(\lambda^{3/2}). \tag{3.58}$$

We may now estimate $I(n)$ as $n \to \infty$ to order n by using (3.42), (3.52), (3.54), and (3.58). We obtain

$$I(n) \sim \left[\tfrac{3}{5}(6\pi^2)^{2/3} n^{5/3} + 2Jn\right] \int_{\mathbf{R}^3} \rho_{\mathrm{TF}}(x)^{5/3} dx$$
$$+ \tfrac{17}{27} Ln + o(n). \tag{3.59}$$

Consequently from (3.43) we have

$$K({}_N\psi_1, \ldots, {}_N\psi_N)$$
$$\sim [N^{7/3} + 10J\, 3^{-5/3}\pi^{-4/3} N^{5/3}] K_{\min}(\rho_{\mathrm{TF}})$$
$$+ N^{5/3} C_W L + o(N^{5/3}), \tag{3.60}$$

where the Von Weizsäcker constant C_W is given by

$$C_W = h^2 (4\pi^2 m)^{-1} \tfrac{17}{54}. \tag{3.61}$$

This value of C_W differs from the value proposed by Von Weizsäcker.[7] In fact his constant is $h^2(8\pi^2 m)^{-1}$.

Let V_N be defined by (1.23), where the potential $V(x)$ is given by (3.1), and $\rho_N(x)$ by (2.4), where the functions ψ_j are associated with $\rho_{\mathrm{TF}}(x)$. By making assumptions on the uniformity of the asymptotic formula (3.36) and assuming a Weyl-type conjecture we can estimate $A(\rho_N, V_N)$ to order $N^{5/3}$ as $N \to \infty$. We obtain

$$A(\rho_N, V_N) \sim [N^{7/3} + 2J(3\pi^2)^{-2/3} N^{5/3}] A(\rho_{\mathrm{TF}}, V)$$
$$+ 2(3\pi^2)^{-2/3} N^{5/3} \int_{\mathbf{R}^3} V \rho_{\mathrm{TF}}{}^{1/6} \Delta \rho_{\mathrm{TF}}{}^{1/6} dx + o(N^{5/3}). \tag{3.62}$$

In a similar fashion we have

$$R(\rho_N) \sim [N^{7/3} + 4J(3\pi^2)^{-2/3} N^{5/3}] R(\rho_{\mathrm{TF}})$$
$$+ 4(3\pi^2)^{-2/3} N^{5/3} \int_{\mathbf{R}^3} \int_{\mathbf{R}^3} \frac{\rho_{\mathrm{TF}}(x) \rho_{\mathrm{TF}}^{1/6}(y) \Delta \rho_{\mathrm{TF}}^{1/6}(y)}{|x-y|} dx\, dy$$
$$+ o(N^{5/3}). \tag{3.63}$$

We conjecture that the asymptotic formulas (3.60), (3.62), and (3.63) are correct and that Theorem 2.4 may be extended to the case of $\rho = \rho_{\mathrm{TF}}$. It then follows from Corollary 3.2 that we may obtain a refinement of Corollary 2.3, namely,

$$E_N(V_N) \leqslant N^{7/3} \epsilon_1(V) + \beta N^{5/3} + o(N^{5/3}), \tag{3.64}$$

where $\epsilon_1(V)$ is the Thomas–Fermi energy for the atom with nuclear attractive potential (3.1), and β is the sum of the terms in (3.60), (3.62), and (3.63) of order $N^{5/3}$ plus the exchange energy.

Remark 1: The main obstruction to proving the asymptotic formulas (3.60), (3.62), and (3.63) is undoubtedly the Weyl-type conjectures which we have assumed. Even for the Euclidean Laplacian in a bounded domain the Weyl conjecture has been verified in only a few cases.[19]

Remark 2: If instead of the potential (3.1) we had taken the Coulomb potential $V(x) = |x|^{-1}$ then Theorem 3.1 would still hold but Corollary 3.2 would not. We might expect this to be the case to account for the term $\alpha N^{6/3}$ of (1.27).

ACKNOWLEDGMENTS

I wish to thank Percy Deift for his encouragement of this research. This research was supported by a University of Missouri Alumni Award and NSF Grant No. MCS 8100761.

[1] E. Lieb and B. Simon, Commun. Math. Phys. **53**, 185 (1977).
[2] E. Lieb and B. Simon, Adv. Math. **23**, 22 (1977).
[3] P. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).
[4] E. Lieb and W. Thirring, Phys. Rev. Lett. **35**, 687 (1975); Erratum: Phys. Rev. Lett. **35**, 1116 (1975).
[5] E. Lieb and S. Oxford, Int. J. Quantum Chem. **19**, 427 (1981).
[6] N. March, Adv. Phys. **6**, 1 (1957).
[7] C. Von Weizsäcker, Z. Phys. **96**, 431 (1935).
[8] A. Kompaneets and E. Pavlovskii, Sov. Phys. JETP **4**, 328 (1957).
[9] J. Scott, Philos. Mag. **43**, 859 (1952).
[10] E. Lieb, "Thomas–Fermi and related theories of Atoms and Molecules," Rev. Mod. Phys. **53**, 603–641 (1981).
[11] Y. Kannai, Commun. Partial Differential Equations **2**, 781 (1977).
[12] N. March and W. Young, Proc. Phys. Soc. **72**, 182 (1958).
[13] E. Lieb, Springer Lect. Notes in Phys. **116**, 91 (1980).
[14] H. McKean, J. Diff. Geom. **1**, 43 (1967).
[15] H. Weyl, Math. Ann. **71**, 441 (1912).
[16] B. Simon, *Functional Integration and Quantum Physics* (Academic, New York, 1979).
[17] N. Dunford and J. Schwartz, *Linear Operators* (Interscience, New York, 1963).
[18] E. Hille, J. Anal. Math. **23**, 147 (1970).
[19] V. Mikhailets, Russ. Math. Surveys **33**, 259 (1978).

Functional integrals in Navier–Stokes incompressible fluid turbulence

Gerald Rosen[a)]
Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104

(Received 27 May 1982; accepted for publication 27 August 1982)

A variational principle is formulated for the dynamical evolution of the Hopf characteristic functional $\Phi = \Phi[\mathbf{y}(\mathbf{x}),t]$ by employing an appropriate functional integral over all parameter fields $\mathbf{y}(\mathbf{x})$. It follows that the ratio of functional integrals $\Gamma \equiv \int \Phi^* \dot{\Phi} D(\mathbf{y}) / \int |\Phi|^2 D(\mathbf{y})$ is an *exact constant of the motion* during the decay of boundary-free Navier–Stokes incompressible fluid turbulence. Bearing the physical dimensions of inverse time, the constant of the motion Γ is a scalar function of the multipoint velocity correlation tensors embodied in Φ. For statistical situations such that the probability measure over the velocity-field ensemble is semi-Gaussian (i.e., the real part of $\ln \Phi$ is a quadratic functional of $\mathbf{y}$), Γ is evaluated explicitly in terms of the two-point velocity correlation tensor.

PACS numbers: 47.25. − c, 47.10. + g

I. INTRODUCTION

Although the essentially nonlinear, dissipative Navier–Stokes equation

$$\partial \mathbf{u}/\partial t = \nu \nabla^2 \mathbf{u} - \mathbf{u}\cdot\nabla\mathbf{u} - \rho^{-1}\nabla p, \quad \nabla\cdot\mathbf{u} \equiv 0, \tag{1}$$

$\mathbf{u} = \mathbf{u}(\mathbf{x},t)$ defined for all $\mathbf{x} \in R_3$, $\nu, \rho \equiv$ positive constants,

does not admit a conventional variational principle, it has been shown recently that a physical *minimum principle* can be formulated for solutions to (1).[1] However, if one considers the multipoint velocity correlation tensors of turbulent flows, there remains an open question regarding the existence of a minimum or a variational principle for the dynamical evolution of the latter statistical quantities.

In the statistical theory for turbulent incompressible fluid flows, all equal-time multipoint velocity correlation tensors are contained in the Hopf characteristic functional[2]

$$\begin{aligned}\Phi = \Phi(\mathbf{y},t) &\equiv \left\langle \exp i \int u_j(\mathbf{x}',t) y_j(\mathbf{x}') d^3x' \right\rangle \\ &= 1 + i \int \langle u_j(\mathbf{x}',t)\rangle y_j(\mathbf{x}')\, d^3x' \\ &\quad - \frac{1}{2}\int\int \langle u_j(\mathbf{x}',t) u_k(\mathbf{x}'',t)\rangle y_j(\mathbf{x}') y_k(\mathbf{x}'') d^3x' d^3x'' \\ &\quad - \frac{i}{6}\int\int\int \langle u_j(\mathbf{x}',t) u_k(\mathbf{x}'',t) u_l(\mathbf{x}''',t)\rangle \\ &\quad \times y_j(\mathbf{x}') y_k(\mathbf{x}'') y_l(\mathbf{x}''') d^3x' d^3x'' d^3x''' + \cdots,\end{aligned} \tag{2}$$

the complex-valued Fourier transform of the probability measure over $\mathbf{u}$. It follows from the definition (2) that Φ satisfies the conditions

$$\begin{aligned}&\Phi(\mathbf{0},t) \equiv 1, && |\Phi(\mathbf{y},t)| \leqslant 1, \\ &\Phi(\mathbf{y},t)^* \equiv \Phi(-\mathbf{y},t), && \Phi(\mathbf{y},t) \equiv \Phi(\mathbf{y}^{\mathrm{tr}},t),\end{aligned} \tag{3}$$

where $\mathbf{y}^{\mathrm{tr}} \equiv \mathbf{y} - \nabla^{-2}\nabla(\nabla\cdot\mathbf{y})$ is the transverse part of the real parameter vector-field $\mathbf{y} = \mathbf{y}(\mathbf{x})$. The characteristic functional (2) changes with time according to the linear equation first derived by Hopf[2,3]

$$\dot{\Phi} \equiv \partial\Phi/\partial t = (F - iS)\Phi, \tag{4}$$

in which there appear the first-order and second-order time-independent functional differential operators[4]

$$F \equiv \nu \int y_j(\mathbf{x}) \nabla^2 (\delta/\delta y_j(\mathbf{x})) d^3x = F^*, \tag{5}$$

$$S \equiv \int \frac{\partial y_j^{\mathrm{tr}}(\mathbf{x})}{\partial x_k} \frac{\delta^2}{\delta y_j(\mathbf{x}) \delta y_k(\mathbf{x})} d^3x = S^*. \tag{6}$$

Equation (4) is a consequence of the fact that every velocity field in the ensemble evolves according to (1), and the real operators (5) and (6) manifest the statistical dynamical effects of the viscous and inertial forces evident in (1). Intertwining of viscous and inertial effects in Navier–Stokes incompressible fluid turbulence shows up in the nonzero commutator of (5) and (6),

$$[F,S] = 2\nu \int \frac{\partial y_j^{\mathrm{tr}}(\mathbf{x})}{\partial x_k} \left(\frac{\partial}{\partial x_l} \frac{\delta}{\delta y_j(\mathbf{x})}\right)\left(\frac{\partial}{\partial x_l} \frac{\delta}{\delta y_k(\mathbf{x})}\right) d^3x. \tag{7}$$

In Sec. II a variational principle is formulated for the Hopf dynamical equation (4) by employing an appropriate functional integral over all parameter fields $\mathbf{y}(\mathbf{x})$. This result is particularly significant in view of the current interest attached to self-organizing variational principles in lower-dimensional and related forms of turbulence.[5] From this variational principle displayed below in (20), one obtains the time dependence of the quantity

$$(\Phi,\Phi) \equiv \int |\Phi(\mathbf{y},t)|^2 D(\mathbf{y}) \tag{8}$$

by way of Noether's theorem. It follows immediately that the ratio of functional integrals of the same type

$$\Gamma \equiv (\Phi,\dot{\Phi})/(\Phi,\Phi) \tag{9}$$

is *an exact constant of the motion* during the decay of boundary-free Navier–Stokes incompressible fluid turbulence,

$$d\Gamma/dt = 0, \tag{10}$$

as shown by the calculation in Sec. III. Bearing the physical dimensions of inverse time, the quantity (9) is a scalar function of the multipoint velocity correlation tensors embodied

[a)] The work reported here was supported by NASA grant NAG1-110.

 0022-2488/83/010193-03$02.50

in Φ [see (2) above]. In Sec. IV Γ is shown to be given exclusively in terms of the two-point velocity correlation tensor (31) by the integral (34) for statistical situations such that the probability measure is semi-Gaussian, i.e., the characteristic funtional takes the form (29) during a certain time-interval of the decay. Amenable to practical evaluation, the functional integral concomitants of Φ in (9) and (20) provide valuable new insights into the theory for Navier–Stokes incompressible fluid turbulence.[6]

II. VARIATIONAL PRINCIPLE FOR THE Φ EQUATION

Let an inner-product for complex-valued functionals of $\mathbf{y}$ be defined by

$$(\Phi_{(1)},\Phi_{(2)}) \equiv \int \Phi^*_{(1)} \Phi_{(2)} D(\mathbf{y}). \tag{11}$$

The infinitesimal volume element or measure in (11) is expressed symbolically as[7]

$$D(\mathbf{y}) = \begin{pmatrix}\text{normalization}\\ \text{constant}\end{pmatrix} \prod_{\mathbf{x}\in R_3} [d^3 y(\mathbf{x})] \tag{12}$$

and the integration in (11) is understood to run from $-\infty$ to $+\infty$ for each component of $\mathbf{y}(\mathbf{x})$ at all $\mathbf{x}$. Observe that (12) has the property of displacement-invariance,

$$D(\mathbf{y}+\mathbf{a}) \equiv D(\mathbf{y}), \tag{13}$$

where $\mathbf{a} = \mathbf{a}(\mathbf{x})$ is an arbitrary real vector field independent of $\mathbf{y}$; in view of (13), finiteness of the inner-product (11) for certain $\Phi_{(1)}$ and $\Phi_{(2)}$ implies that[8]

$$\int \left(\frac{\delta \Phi^*_{(1)}}{\delta y_j(\mathbf{x})} \Phi_{(2)} + \Phi^*_{(1)} \frac{\delta \Phi_{(2)}}{\delta y_j(\mathbf{x})} \right) D(\mathbf{y}) = 0. \tag{14}$$

The adjoints of operators (5) and (6) are defined implicitly by

$$(\Phi_{(1)}, F\Phi_{(2)}) \equiv (F^\dagger \Phi_{(1)}, \Phi_{(2)}), \tag{15}$$

$$(\Phi_{(1)}, S\Phi_{(2)}) \equiv (S^\dagger \Phi_{(1)}, \Phi_{(2)}), \tag{16}$$

and determined by making use of (14),

$$F^\dagger = -F + c, \tag{17}$$

$$S^\dagger = S, \tag{18}$$

where

$$c \equiv -3\nu \int \nabla^2 \delta^{(3)}(\mathbf{x})|_{\mathbf{x}=0}\, d^3x \tag{19}$$

is a positive real constant.[9]

In terms of (11) the variational principle for (4) takes the form

$$\delta \int_{t_0}^{t_1} e^{-ct} (\Phi, \dot\Phi - F\Phi + iS\Phi)\, dt = 0 \tag{20}$$

with $\delta\Phi$ and $\delta\Phi^*$ equal to zero at the terminal times t_0, t_1 but arbitrary (and treated as mutually independent) for $t_0 < t < t_1$. Clearly, the $\delta\Phi^*$ term that arises in (20) vanishes if and only if (4) is satisfied, while the $\delta\Phi$ term (obtained by parts integration with respect to t) vanishes if and only if

$$-\dot\Phi^* + c\Phi^* - F^\dagger \Phi^* + iS^\dagger \Phi^* = 0. \tag{21}$$

In view of (17) and (18), (21) becomes

$$-\dot\Phi^* + F\Phi^* + iS\Phi^* = 0, \tag{22}$$

which is just the complex-conjugate of (4). Hence (20) is equivalent to (4).

As a consequence of the variational principle (20), Noether's theorem applies and takes the form

$$e^{-ct_1}(\Phi, \delta\Phi)|_{t_1} = e^{-ct_0}(\Phi, \delta\Phi)|_{t_0}, \tag{23}$$

where $\delta\Phi$ is the change in the characteristic functional associated with an infinitesimal transformation that leaves the integral in (20) unchanged. The obvious invariance transformation

$$\Phi \to e^{i\alpha}\Phi, \quad \Phi^* \to e^{-i\alpha}\Phi^*, \quad \alpha = (\text{real constant}), \tag{24}$$

gives $\delta\Phi = i\Phi\delta\alpha$ and thus yields

$$(\Phi,\Phi)|_{t_1} = e^{c(t_1 - t_0)} (\Phi,\Phi)|_{t_0}. \tag{25}$$

Shown in (25), the time-dependence of the quantity (8) can also be derived by multiplying (4) by Φ^*, integrating the real part of the resulting equation over $\mathbf{y}$ with the measure (12), and finally using (17) and (18) to get

$$\frac{d}{dt}(\Phi,\Phi) = c(\Phi,\Phi). \tag{26}$$

III. CONSTANCY OF Γ

It follows from (4) that

$$\begin{aligned} &(\Phi,\ddot\Phi) = (\Phi, F\dot\Phi - iS\dot\Phi) = (F^\dagger\Phi, \dot\Phi), \\ &-i(S^\dagger\Phi,\dot\Phi) = -(F\Phi,\dot\Phi) + c(\Phi,\dot\Phi), \\ &-i(S\Phi,\dot\Phi) = -(\dot\Phi,\dot\Phi) + c(\Phi,\dot\Phi), \end{aligned} \tag{27}$$

where (15)–(18) and the complex-conjugate of (4) have been employed. Thus the time-derivative of (9) is

$$\frac{d\Gamma}{dt} = \frac{(\Phi,\ddot\Phi) + (\dot\Phi,\dot\Phi)}{(\Phi,\Phi)} - \frac{(\Phi,\dot\Phi)}{(\Phi,\Phi)^2} \frac{d(\Phi,\Phi)}{dt} = 0 \tag{28}$$

by virtue of (27) and (26).

Observe that the constancy of Γ defined by (9) is analogous to the constancy of the Hamiltonian expectation value $\langle H \rangle$ in quantum field theory. The slight complication here comes from the non-skew-adjointness of $F(\neq -F^\dagger)$ shown in (17), but the dynamical effects of c cancel out in (9).

IV. EVALUATION OF Γ FOR A SEMI-GAUSSIAN STATISTICAL ENSEMBLE

Suppose that

$$\Phi = \exp\left(-\frac{1}{2} \int\int R_{jk}(\mathbf{x}',\mathbf{x}'',t) y_j(\mathbf{x}') y_k(\mathbf{x}'') d^3x' d^3x'' + iA[\mathbf{y},t] \right) \tag{29}$$

is a suitable approximation for the characteristic functional during a certain time-interval of the decay, where

$$A[\mathbf{y},t] \equiv A[\mathbf{y},t]^* \equiv -A[-\mathbf{y},t] \tag{30}$$

is an arbitrary real odd functional of $\mathbf{y}$. The semi-Gaussian form (29) is consistent with (2) and (3) provided that

$$R_{jk}(\mathbf{x}',\mathbf{x}'',t) = \langle u_j(\mathbf{x}',t) u_k(\mathbf{x}'',t) \rangle \tag{31}$$

is the positive-definite symmetric solenoidal two-point velocity correlation tensor. For the numerator in (9) one obtains

$$(\Phi,\dot{\Phi}) = \int \Big(-\frac{1}{2}\iint \dot{R}_{jk}(\mathbf{x}',\mathbf{x}'',t)y_j(\mathbf{x}')y_k(\mathbf{x}'')d^3x'd^3x'' + i\dot{A}[\mathbf{y},t]\Big)|\Phi|^2 D(\mathbf{y}) \tag{32}$$

$$= -\frac{1}{2}\iint \dot{R}_{jk}(\mathbf{x}',\mathbf{x}'',t)\int y_j^{\mathrm{tr}}(\mathbf{x}')y_k^{\mathrm{tr}}(\mathbf{x}'')|\Phi|^2 D(\mathbf{y}) \times d^3x'd^3x''$$

because

$$|\Phi|^2 = \exp\Big(-\iint R_{jk}(\mathbf{x}',\mathbf{x}'',t)y_j(\mathbf{x}')y_k(\mathbf{x}'')d^3x'd^3x''\Big) \tag{33}$$

is even in $\mathbf{y}$ while $\dot{A}[\mathbf{y},t]$ is odd as a consequence of (30). The remaining Gaussian functional integral in the final member of (32) is well known,[7,8] and thus (9) is evaluated as

$$\Gamma = -\frac{1}{4}\iint \dot{R}_{jk}(\mathbf{x}',\mathbf{x}'',t)R^{-1}_{jk}(\mathbf{x}',\mathbf{x}'',t)\,d^3x'\,d^3x'', \tag{34}$$

where R^{-1} is the matrix-kernel inverse to R on the space of solenoidal vector-fields:

$$\int R^{-1}_{jk}(\mathbf{x}',\mathbf{x},t)R_{kl}(\mathbf{x},\mathbf{x}'',t)\,d^3x = \delta^{\mathrm{tr}}_{jl}(\mathbf{x}'-\mathbf{x}'')$$

$$\equiv \Big(\delta_{jl} - \nabla^{-2}_{\mathbf{x}'}\frac{\partial^2}{\partial x'_j\partial x'_l}\Big)\delta^{(3)}(\mathbf{x}'-\mathbf{x}''). \tag{35}$$

Clearly, if the time-dependence in (31) resides entirely in a scalar prefactor,[10]

$$R_{jk}(\mathbf{x}',\mathbf{x}'',t) = u^2(t)C_{jk}(\mathbf{x}',\mathbf{x}''), \tag{36}$$

then (35), (34), and (10) imply that

$$d^2[\ln u^2(t)]/dt^2 = 0. \tag{37}$$

Equation (37) shows that the decay of u^2 is exponential during the time-interval for which (29) remains valid as an approximation.

V. PROBABILISTIC SIGNIFICANCE AND VALUE OF Γ

As established by Liouville's theorem, the flow of probability is incompressible in phase space for a conservative Hamiltonian dynamical system. The Navier–Stokes equation (1) describes a nonconservative dynamical system with dissipation produced by the viscosity term. As a consequence, probability flows like a uniformly contracting compressible fluid in infinite-dimensional $\mathbf{u}(\mathbf{x})$-space. A simplifying feature of the probability flow in $\mathbf{u}(\mathbf{x})$-space is that the divergence of the flow-lines equals $-c$ for all $\mathbf{u}(\mathbf{x})$ and t, where the positive constant c is defined by (19). It is this constancy of the probability-flow divergence which produces constancy of Γ defined in (9). By introducing the probability density[3] $P[\mathbf{u},t]$ as the functional Fourier transform of (2), the quantity (9) is expressible as

$$\Gamma = \frac{1}{2}\frac{d}{dt}\ln\Big(\int P[\mathbf{u},t]^2 D(\mathbf{u})\Big) \tag{38}$$

while (26) becomes

$$\frac{d}{dt}\int P[\mathbf{u},t]^2 D(\mathbf{u}) = c\int P[\mathbf{u},t]D(\mathbf{u}). \tag{39}$$

Thus, constancy of the probability flow divergence [implicit in the Hopf equation (4)] engenders the value which follows from (38) and (39),

$$\Gamma = \tfrac{1}{2}c. \tag{40}$$

[1] G. Rosen, J. Math. Phys. **23**, 676 (1982).
[2] G. Rosen, J. Math. Phys. **22**, 1819 (1981).
[3] E. Hopf, J. Ratl. Mech. Anal. **1**, 87 (1952); E. Hopf and E. W. Titt, J. Ratl. Mech. Anal. **2**, 587 (1953).
[4] It should be noted that $S\Phi = \int(\partial y_j^{\mathrm{tr}}(\mathbf{x})/\partial x_k)(\delta^2\Phi/\delta y_j(\mathbf{x})\delta y_k(\mathbf{x}))\,d^3x = -\int y_j^{\mathrm{tr}}(\mathbf{x})(\delta/\delta y_k(\mathbf{x}))\nabla_k(\delta\Phi/\delta y_j(\mathbf{x}))\,d^3x$ by virtue of the fourth condition in (3), which implies that $\nabla_k(\delta\Phi/\delta y_k(\mathbf{x})) = 0$.
[5] A. Hasegawa, Y. Kodama, and K. Watanabe, Phys. Rev. Lett. **47**, 1525 (1981).
[6] A space-time path integral representation of the general solution to (4) has been known for over 20 years [G. Rosen, Phys. Fluids **3**, 519, 525 (1960); I. Hosokawa, J. Math. Phys. **8**, 221 (1967); G. T. Papadopoulis, in *Path Integrals and Their Applications in Quantum, Statistical, and Solid State Physics* (Plenum, New York, 1978), pp. 85–162], but the problem of extracting experimentally relevant information from such a path integral has yet to be solved.
[7] This type of functional integration measure was defined rigorously by: K. O. Friedrichs, H. N. Shapiro, *et al.*, *Integration of Functionals* (New York University, Institute of Mathematical Sciences, 1957).
[8] G. Rosen, in *Path Integrals and Their Applications in Quantum, Statistical, and Solid State Physics* (Plenum, New York, 1978), pp. 201–235.
[9] Formally infinite as defined in (19), the constant c equals $3\nu K^5 V/10\pi^2$ if one imposes a wavenumber cutoff $K \geqslant |\mathbf{k}|$ and finite spatial volume V, with $K\to\infty, V\to\infty$ understood to be taken as the final step in all calculations.
[10] See, for example: G. Rosen, Phys. Fluids **24**, 558 (1981).

Oscillatory magnetohydrodynamic flow and heat transfer between conducting plates

V. M. Soundalgekar
Department of Mathematics, Indian Institute of Technology, Powai, Bombay 400076, India
J. P. Bhat
Department of Mathematics, Goa Engineering College, Farmagudi (Goa), India

(Received 27 May 1981; accepted for publication 18 September 1981)

An approximate solution of the oscillatory flow of an electrically conducting fluid, between two parallel and electrically conducting plates and under transversely applied magnetic field, is given for the transient velocity, the transient magnetic field, amplitude and the phase of the skin friction and the rate of heat transfer. It is observed that the transient flow, amplitude, and the phase of the skin-friction and the rate of heat transfer are affected by the individual electrical conductance ratios of the plates, which is not so in the case of steady magnetohydrodynamic (MHD) channel flow between conducting plates.

PACS numbers: 47.65. + a, 47.60. + i, 47.25.Qv

1. INTRODUCTION

Steady MHD channel flows have been studied during last 20 years by a number of researchers because of its applications in MHD generators, MHD flow meters, nuclear engineering etc. These are discussed in books like Cowling,[1] Pai,[2] Sutton and Sherman,[3] and Hughes and Young[4] under different physical conditions. In all these studies, the channel is bounded by electrically nonconduction plates. However, in a number of cases, the plates of the channel become electrically conducting. This leads to a change of boundary conditions on the induced magnetic field and it is derived by Shercliff.[5] Taking into account electrically conducting plates, the steady MHD channel flow was studied by Chang and Yen,[6] whereas the heat transfer aspect of this flow was studied by Soundalgekar.[7]

In all these studies, the pressure gradient is assumed to be constant. If the pressure gradient is assumed to be oscillatory of the form

$$-\frac{1}{\rho}\frac{\partial p'}{\partial x'} = A(1+\epsilon e^{i\omega' t'}), \tag{1}$$

the MHD channel flow becomes oscillatory and such a study for a nonconducting plate MHD channel was made recently by Soundalgekar and Bhat.[8] But how the conducting plates affect the oscillatory flow, whose pressure gradient is represented by (1), has not been studied in the literature. Also, the heat transfer of such a flow has also not been studied in the literature. Hence it is now proposed to study the effects of electrically conducting plates of the channel on the oscillatory flow and heat transfer. In Refs. 6 and 7 it was observed that the flow and heat transfer are affected by the sum of the electrical conductance ratios, $\Phi_1 + \Phi_2$ of the two plates, where Φ_1 and Φ_2 are the electrical conductance ratios of the two plates. However, in the present case, the transient velocity, the transient magnetic field, the amplitude, and the phase of the skin friction and the rate of heat transfer are found to be affected by the individual electrical conductance ratios. This is the most significant change observed due to the oscillatory flow character in the MHD channel. In Sec. 2, the mathematical analysis is presented and in Sec. 3, the conclusions are set out.

2. MATHEMATICAL ANALYSIS

Consider the unsteady flow of an electrically conducting, viscous, incompressible fluid between two infinite parallel plates, separated by a distance of $2L$. The x' axis is taken along the center line of the channel and the y' axis taken normal to it. A magnetic field of uniform strength is assumed to be applied parallel to the y' axis. The flow is now governed by the following equations:

$$\frac{\partial u'}{\partial t'} = -\frac{1}{\rho}\frac{\partial p'}{\partial x'} + \nu\frac{\partial^2 u'}{\partial y'^2} + \frac{\mu_c H_0}{4\pi\rho}\frac{\partial H'_x}{\partial y'}, \tag{2}$$

$$\frac{\partial H'_x}{\partial t'} = H_0\frac{\partial u'}{\partial y'} + \eta\frac{\partial^2 H'_x}{\partial y'^2}, \tag{3}$$

and the boundary conditions are:

$$u' = 0, \quad \sigma_1 d_1\frac{dH'_x}{dy'} + \sigma H'_x = 0 \quad \text{at } y' = +L$$

$$u' = 0, \quad \sigma_2 d_2\frac{dH'_x}{dy'} - \sigma H'_x = 0 \quad \text{at } y' = -L. \tag{4}$$

Here u' is the velocity in the x' direction, ν the kinematic viscosity, μ_c the magnetic permeability, H_0 the constant applied magnetic field, ρ the density, H'_x the induced magnetic field, t' the time, $\eta = 1/4\pi\mu_c\sigma$ the magnetic diffusivity, σ the scalar electrical conductivity of the fluid, and σ_1, σ_2 the scalar electrical conductivities of the upper and lower plates with thicknesses d_1 and d_2, respectively.

To find the solutions, we now assume

$$u' = u'_0 + \epsilon e^{i\omega' t'}u'_1, \quad H'_x = h'_0 + \epsilon e^{i\omega' t'}h'_1 \tag{5}$$

and substitute (5) and (1) in Eqs. (2)–(4), equate harmonic and nonharmonic terms and get

$$\nu\frac{d^2 u'_0}{dy'^2} + \frac{\mu_c H_0}{4\pi\rho}\frac{dh'_0}{dy'} + A = 0, \tag{6}$$

 0022-2488/83/010196-06$02.50

$$\nu \frac{d^2 u_1'}{dy'^2} + \frac{\mu_c H_0}{4\pi\rho} \frac{dh_1'}{dy'} + A = i\omega' u_1' , \tag{7}$$

$$H_0 \frac{du_0'}{dy'} + \eta \frac{d^2 h_0'}{dy'^2} = 0 , \tag{8}$$

$$H_0 \frac{du_1'}{dy'} + \eta \frac{d^2 h_1'}{dy'^2} = i\omega' h_1' , \tag{9}$$

and the boundary conditions are

$$\begin{aligned} &u_0' = u_1' = 0 \quad \text{at } y' = \pm L , \\ &\sigma_1 d_1 \frac{dh_0'}{dy'} + \sigma h_0' = 0, \\ &\sigma_1 d_1 \frac{dh_1'}{dy'} + \sigma h_1' = 0 \quad \text{at } y' = L , \\ &\sigma_2 d_2 \frac{dh_0'}{dy'} - \sigma h_0' = 0, \\ &\sigma_2 d_2 \frac{dh_1'}{dy'} - \sigma h_1' = 0 \quad \text{at } y' = -L . \end{aligned} \tag{10}$$

Introducing the following nondimensional quantities:

$$\begin{aligned} &y = y'/L, \quad h_0 = h_0'/H_0 R_m, \quad h_1 = h_1'/H_0 R_m , \\ &R_m = 4\pi\mu_c \sigma L A^*, \quad A^* = AL^2/\nu, \quad u_0 = u_0'/A^* , \\ &u_1 = u_1'/A^*, \quad M^2 = \mu_c^2 H_0^2 L^2 \sigma/\mu, \quad \omega = \omega' L^2/\nu, \\ &\Phi_1 = \frac{\sigma_1 d_1}{\sigma L}, \quad \Phi_2 = \frac{\sigma_2 d_2}{\sigma L}, \quad R_l = \frac{LA^*}{\nu} \end{aligned} \tag{11}$$

in Eqs. (6)–(10), we have

$$\frac{d^2 u_0}{dy^2} + M^2 \frac{dh_0}{dy} = -1 , \tag{12}$$

$$\frac{d^2 h_0}{dy^2} + \frac{du_0}{dy} = 0 , \tag{13}$$

$$\frac{d^2 u_1}{dy^2} + M^2 \frac{dh_1}{dy} + 1 = i\omega u_1 , \tag{14}$$

$$\frac{d^2 h_1}{dy^2} + \frac{du_1}{dy} = i\omega N h_1 , \tag{15}$$

where $N = R_m/Re$, and the boundary conditions are

$$\begin{aligned} &u_0 = u_1 = 0 \qquad \text{at } y = \pm 1 , \\ &\frac{dh_0}{dy} + \frac{1}{\Phi_1} h_0 = 0 \quad \text{at } y = +1 , \\ &\frac{dh_0}{dy} - \frac{1}{\Phi_2} h_0 = 0 \quad \text{at } y = -1 , \\ &\frac{dh_1}{dy} + \frac{1}{\Phi_1} h_1 = 0 \quad \text{at } y = +1 , \\ &\frac{dh_1}{dy} - \frac{1}{\Phi_2} h_1 = 0 \quad \text{at } y = -1 . \end{aligned} \tag{16}$$

Here R_m, and M are, respectively, magnetic Reynolds and Hartmann numbers.

The solutions of Eqs. (12)–(15) satisfying the boundary conditions (16) are derived as follows:

$$u_0 = C_2(\cosh My - \cosh M) , \tag{17}$$

$$h_0 = -C_2 \frac{\sinh My}{M} - \frac{y}{M^2} + C_3 , \tag{18}$$

$$u_1 = \{X_1(C_4 \sinh b_1 y + C_5 \cosh b_1 y) + X_2(C_6 \sinh b_2 y + C_7 \cosh b_2 y) + 1\}/i\omega , \tag{19}$$

$$h_1 = C_4 \cosh b_1 y + C_5 \sinh b_1 y + C_6 \cosh b_2 y + C_7 \sinh b_2 y , \tag{20}$$

where the constants b_1, b_2, C_2,..., C_7 are defined in the Appendix.

The steady velocity profiles were already studied in Ref. 5. The velocity and induced magnetic field are given by

$$\begin{aligned} &u = u_0 + \epsilon e^{i\omega t} u_1(y) \\ &h = h_0 + \epsilon e^{i\omega t} h_1(y) . \end{aligned} \tag{21}$$

We can write the expressions for u and h in terms of their fluctuating parts for $\omega t = \pi/2$ as

$$\begin{aligned} &u = u_0 - \epsilon M_i , \\ &h = h_0 - \epsilon h_{1i} , \end{aligned} \tag{22}$$

where

$$M_r + iM_i = u_1, \quad h_{1r} + ih_{1i} = h_1 . \tag{23}$$

In order to get physical insight into the problem, we have calculated u and h from Eqs. (22) and these are shown in Figs. 1 and 2 respectively.

It has been observed in Ref. 5 that the steady velocity is affected by the sum of the electrical conductance ratios of the two plates. But in the present case, the transient velocity has been found to be affected by the individual electrical conductance ratios of the plate. We observe from Fig. 1 that when Φ_1, Φ_2, M are constant, an increase in ω has different effects. At small values of ω, the transient velocity increases but at large values of ω (~ 100), the transient velocity decreases. The effect of increasing M is the same as in steady-state case

	ϕ_1	ϕ_2	M	ω
I	1.0	0.2	2.0	5.0
II	1.0	0.2	2.0	100.0
III	1.0	0.2	6.0	5.0
IV	1.0	0.2	6.0	15.0
V	5.0	0.6	2.0	5.0
VI	0.2	1.0	2.0	5.0
VII	0.2	0.2	2.0	5.0
VIII	5.0	0.2	2.0	5.0
IX	1.0	0.6	2.0	5.0
X	1.0	0.2	6.0	100.0

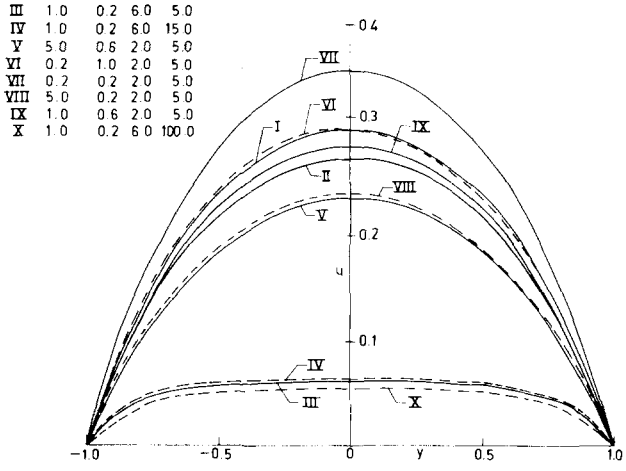


FIG. 1. Transient velocity profiles $Rm/Re = 0.02$, $\epsilon = 0.2$, $\omega t = \pi/2$.

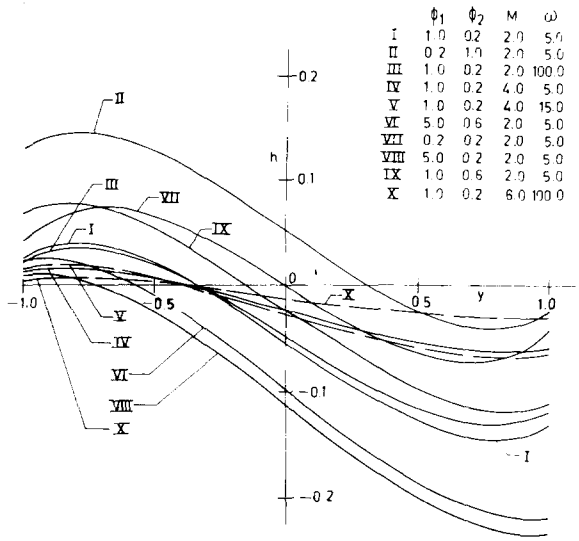


FIG. 2. Transient magnetic field $Rm/Re = 0.02$, $\epsilon = 0.2$, $\omega t = \pi/2$.

and we observe that the transient velocity decreases and gets flattened. Also an increase in Φ_1 or Φ_2 leads to a decrease in the transient velocity. The transient induced magnetic field is shown in Fig. 2. We observe from Fig. 2 that the transient induced magnetic field decreases with increasing ω or M. But an increase in Φ_2 leads to an increase in h, whereas an increase in Φ_1 leads to a decrease in h.

From the velocity field, we now study the skin friction. It is given by

$$\tau' = -\mu \left.\frac{du'}{dy'}\right|_{y' = \pm L}, \tag{24}$$

which in view of (11) reduces to

$$\tau = \tau'/(\mu A^*/L) = \left.\frac{du}{dy}\right|_{y = \pm 1} = \left.\frac{du_0}{dy}\right|_{y = \pm 1} + \epsilon e^{i\omega t} \left.\frac{du_1}{dy}\right|_{y = \pm 1}. \tag{25}$$

The numerical values of $\tau_m = du_0/dy|_{y=-1}$ are calculated and they are shown in Table I. We observe from this table that the mean skin friction τ_m decreases with increasing $\Phi_1 + \Phi_2$.

We can express τ in terms of the amplitude and phase as

$$\tau = \tau_m + \epsilon |B| \cos(\omega t + \alpha),$$

where

$$B = \left.\frac{du_1}{dy}\right|_{y = \pm 1} = B_r + iB_i \tag{26}$$

and

$$\tan\alpha = B_i/B_r.$$

The numerical values of $|B|$ at the two plates are entered in Table I. They are affected by the individual conductance ratios of the plates. Let us denote the amplitude of the skin friction at the lower plate as $|B_1|$ and that at the upper plate by $|B_2|$. We observe that an increase in Φ_2 or Φ_1 leads to a decrease in both $|B_1|$ and $|B_2|$. But an increase in ω leads to a decrease in $|B_1|$ and $|B_2|$. $|B_1|$, $|B_2|$ also decrease with increasing M.

The values of $\tan\alpha$, the phase of the skin friction, are also entered in Table I. We conclude from this table that, both being negative, there is always a phase lag.

A. Energy equation

The unsteady energy equation for the present case is given by

$$\rho c_p \frac{\partial T'}{\partial t'} = k \frac{\partial^2 T'}{\partial y'^2} + \mu \left(\frac{\partial u'}{\partial y'}\right)^2 + \frac{j_z^2}{\sigma}, \tag{27}$$

which takes account of both viscous and Joule dissipation effects.

We assume the solution in the form

$$\theta = \theta_0 + \frac{\epsilon}{2}(e^{i\omega' t'}\theta_1 + e^{-i\omega' t'}\bar{\theta}_1) + \frac{\epsilon^2}{2}(e^{2i\omega' t'}\theta_2 + e^{-2i\omega' t'}\bar{\theta}_2), \tag{28}$$

where $\theta = T' - T_1/(T_2 - T_1)$ and $\bar{}$ denotes the complex conjugate. Here T_1 and T_2 are the temperatures of the lower and upper plates, respectively. In addition to (28), we write (5) as

$$u' = u_0' + \frac{\epsilon}{2}(e^{i\omega' t'}u_1' + e^{-i\omega' t'}\bar{u}_1')$$

and (29)

TABLE I. Values of τ_m, $|B|$, $\tan\alpha$.

M	Rm/Re	ω	ϕ_1	ϕ_2	Mean skin-friction at $Y = 1$	$\lvert B_1\rvert$	$\lvert B_2\rvert$	$\tan\alpha$ at $y = -1$	$y = +1$
2.0	0.02	5.0	0.2	1.0	0.7128	0.4959	0.4991	−0.7230	−0.7056
2.0	0.02	5.0	1.0	0.2	0.7128	0.4991	0.4959	−0.7056	−0.7230
2.0	0.02	5.0	0.6	5.0	0.5581	0.4413	0.4492	−0.5482	−0.5168
2.0	0.02	5.0	0.2	0.2	0.8481	0.5316	0.5316	−0.8543	−0.8543
2.0	0.02	10.0	0.2	1.0	0.7128	0.3364	0.3421	−0.9685	−0.9438
2.0	0.02	15.0	0.2	1.0	0.7128	0.2650	0.2713	−1.013	−0.9954
2.0	0.01	5.0	0.2	1.0	0.7128	0.4954	0.4969	−0.7198	−0.7109
2.0	0.01	15.0	0.2	1.0	0.7128	0.2661	0.2691	−1.006	−0.9933
4.0	0.02	5.0	0.2	1.0	0.4704	0.4279	0.4313	−0.3712	−0.3414
4.0	0.02	5.0	0.6	5.0	0.3113	0.3012	0.3086	−0.2070	−0.1470

$$j_z = j_0 + \frac{\epsilon}{2}(e^{i\omega' t'} j_1 + e^{-i\omega' t'} \bar{j}_1).$$

Substituting (28) and (29) in Eqs. (27), equating harmonic and nonharmonic terms, and neglecting the coefficients of ϵ^3, we get, in view of (11), the following equations in nondimensional form:

$$\frac{d^2\theta_0}{dy^2} + PE\left[\left(\frac{du_0}{dy}\right)^2 + \frac{\epsilon^2}{2}\left(\frac{du_1}{dy}\right)\left(\frac{d\bar{u}_1}{dy}\right)\right] + M^2 PE\left(J_0^2 + \frac{\epsilon^2}{2} J_1 \bar{J}_1\right) = 0, \tag{30}$$

$$\frac{d^2\theta_1}{dy^2} - i\omega P\theta_1 = -2PE\left(\frac{du_0}{dy}\right)\left(\frac{du_1}{dy}\right) - 2M^2 PEJ_0 J_1, \tag{31}$$

$$\frac{d^2\theta_2}{dy^2} - 2i\omega P\theta_2 = -\frac{1}{2} PE\left[\left(\frac{du_1}{dy}\right)^2 + M^2 J_1^2\right], \tag{32}$$

and two more equations for $\bar{\theta}_1$, $\bar{\theta}_2$ similar to (31) and (32), respectively. Here $P = \mu c_p/k$ is the Prandtl number, $E = A^{*2}/c_p(T_2 - T_1)$, the Eckert number, and $J = j_z/\sigma\mu_e H_0 A^*$.

The boundary conditions are:

$$\theta_0(-1) = 0, \quad \theta_0(1) = 1, \quad \theta_1(\pm 1) = 0, \quad \theta_2(\pm 1) = 0. \tag{33}$$

Remembering that $J = -dh/dy$, and substituting for u_0, u_1, h_0, and h_1 from (17)–(20), in Eqs. (30)–(32), and solving these under the boundary conditions (33), we have

$$\theta_0 = C_8 + C_9 y + P_1(y), \tag{34}$$

where

$$P_1(y) = -PE\left\{P_{11}(y) + \frac{\epsilon^2}{2} P_{12}(y)\right\},$$

$$P_{11}(y) = X_9 \cosh 2My + X_{10} \cosh My + y^2/2M^2,$$

$$\begin{aligned} P_{12}(y) = {} & X_{15} \cosh b_5 y + X_{16} \sinh b_5 y + X_{17} \cosh b_6 y \\ & + X_{18} \sinh b_6 y + X_{19} \cosh b_7 y + X_{20} \sinh b_7 y \\ & + X_{21} \cosh b_8 y + X_{22} \sinh b_8 y + X_{23} \cosh b_9 y \\ & + X_{24} \sinh b_9 y + X_{25} \cosh b_{10} y + X_{26} \sinh b_{10} y \\ & + X_{27} \cosh b_{11} y + X_{28} \sinh b_{11} y \\ & + X_{29} \cosh b_{12} y + X_{30} \sinh b_{12} y, \end{aligned}$$

where the constants $C_8, C_9, \ldots, X_{30}, b_5, \ldots, b_{12}$ are given in the Appendix.

$$\theta_1 = C_{10} \cosh b_3 y + C_{11} \sinh b_3 y - 2PEP_2(y), \tag{35}$$

where

$$\begin{aligned} P_2(y) = {} & X_{31} \sinh b_{13} y + X_{32} \cosh b_{13} y + X_{33} \sinh b_{14} y \\ & + X_{34} \cosh b_{14} y + X_{35} \sinh b_{15} y + X_{36} \cosh b_{15} y \\ & + X_{37} \sinh b_{16} y + X_{38} \cosh b_{16} y + X_{39} \sinh b_1 y \\ & + X_{40} \cosh b_1 y + X_{41} \sinh b_2 y + X_{42} \cosh b_2 y. \end{aligned}$$

The constants are defined in the Appendix.

$$\theta_2 = C_{12} \cosh b_4 y + C_{13} \sinh b_4 y - \tfrac{1}{2} PEP_3(y), \tag{36}$$

where

$$\begin{aligned} P_3(y) = {} & X_{43} + X_{44} \cosh 2b_1 y + X_{45} \sinh 2b_1 y \\ & + X_{46} \cosh 2b_2 y + X_{47} \sinh 2b_2 y + X_{48} \cosh b_{17} y \\ & + X_{49} \sinh b_{17} y + X_{50} \cosh b_{18} y \\ & + X_{51} \sinh b_{18} y. \end{aligned}$$

The constants $X_{43}, \ldots, X_{51}$ are defined in the Appendix.

Substituting for θ_0, θ_1, and θ_2 in the expression for θ, we can get the expression for the temperature field. But we are interested in the rate of heat transfer. It is given by

$$q' = -k \left.\frac{\partial T'}{\partial y'}\right|_{y' = \pm L} \tag{37}$$

and in view of (11), Eq. (37) reduces to

$$\begin{aligned} q = \left.\frac{d\theta}{dy}\right|_{y=\pm 1} = {} & \left.\frac{d\theta_0}{dy}\right|_{y=\pm 1} + \epsilon e^{i\omega t} \left.\frac{d\theta_1}{dy}\right|_{y=\pm 1} \\ & + \epsilon^2 e^{2i\omega t} \left.\frac{d\theta_2}{dy}\right|_{y=\pm 1}, \end{aligned} \tag{38}$$

where $q = -q'L/k(T_2 - T_1)$.

The mean rate of heat transfer is given by

$$q_m = \left.\frac{d\theta_0}{dy}\right|_{y=\pm 1}. \tag{39}$$

From (34) and (39), we have calculated the expression for q_m and the numerical values of q_m are entered in Table II. We observe from this table that q_m at the lower plate decreases with increasing Φ_2 whereas q_m at the upper plate increases with increasing Φ_1. q_m is not significantly affected by ω at both the plates. An increase in E leads to an increase in q_m at the lower plate and a decrease in q_m at the upper plate. But an increase in M leads to an increase in q_m at the upper plate and a decrease in q_m at the lower plate.

We can now express the expression for the rate of heat transfer in terms of the amplitude and phase as

$$q = q_m + \epsilon|Q_1| \cos(\omega t + \alpha_1) + \epsilon^2|Q_2| \cos(\omega t + \alpha_2), \tag{40}$$

where

$$\begin{aligned} Q_1 &= \left.\frac{d\theta_1}{dy}\right|_{y=\pm 1} = Q_{1r} + iQ_{1i}, \\ Q_2 &= \left.\frac{d\theta_2}{dy}\right|_{y=\pm 1} = Q_{2r} + iQ_{2i}, \\ \tan\alpha_1 &= Q_{1i}/Q_{1r}, \quad \tan\alpha_2 = Q_{2i}/Q_{2r}. \end{aligned} \tag{41}$$

We have calculated the numerical values of $|Q_1|$, $|Q_2|$, $\tan\alpha_1$, and $\tan\alpha_2$ and their numerical values are entered in Table II. We observe from this table that an increase in Φ_2 or Φ_1 leads to a decrease in the amplitude of the first harmonic of the rate of heat transfer. $|Q_1|$ decreases with increasing ω or M. The effects of Φ_1, Φ_2, M, or ω on $|Q_2|$, the amplitude of the second harmonic, are the same as in $|Q_1|$.

The values of $\tan\alpha_1$, the phase of the first harmonic, being negative, we conclude that there is a phase lag, whereas the values of $\tan\alpha_2$ being positive, there is always a phase lead.

3. CONCLUSIONS

1) The transient velocity increases with increasing ω at small values of ω, whereas at large values of ω it decreases.

2) The transient velocity decreases with increasing M, Φ_1, or Φ_2.

3) The amplitudes of the skin friction at both the plates decrease with increasing Φ_1, Φ_2, M, or ω.

TABLE II. Values of q_m, $|Q_1|$, $|Q_2|$, $\tan\alpha_1$, $\tan\alpha_2$ ($P = 0.71$).

						At the wall $y =$	q_m	$\lvert Q_1\rvert$	$\tan\alpha_1$	$\lvert Q_2\rvert$	$\tan\alpha_2$
0.2	1.0	2.0	0.02	5	0.01	− 1	0.5011	0.9745×10^{-3}	− 2.903	0.1200×10^{-3}	2.443
						+ 1	0.4989	0.9743×10^{-3}	− 2.830	0.1213×10^{-3}	2.682
1.0	0.2	2.0	0.02	5	0.02	− 1	0.5011	0.9743×10^{-3}	− 2.830	0.1213×10^{-3}	2.682
						+ 1	0.4989	0.9745×10^{-3}	− 2.903	0.1200×10^{-3}	2.443
0.6	5.0	2.0	0.02	5	0.01	− 1	0.5010	0.8533×10^{-3}	− 3.385	0.1032×10^{-3}	2.800
						+ 1	0.4990	0.8546×10^{-3}	− 3.356	0.1044×10^{-3}	3.130
0.2	0.2	2.0	0.02	5	0.01	− 1	0.5014	0.1189×10^{-2}	− 2.831	0.1426×10^{-3}	1.971
						+ 1	0.4986	0.1189×10^{-2}	− 2.831	0.1426×10^{-3}	1.971
0.2	1.0	2.0	0.02	15	0.01	− 1	0.5011	0.3068×10^{-3}	− 8.006	0.2004×10^{-4}	1.033
						+ 1	0.4989	0.3104×10^{-3}	− 6.929	0.2132×10^{-4}	1.108
0.2	1.0	2.0	0.02	5	0.02	− 1	0.5023	0.1949×10^{-2}	− 2.903	0.2401×10^{-3}	2.443
						+ 1	0.4977	0.1949×10^{-2}	− 2.830	0.2426×10^{-3}	2.682
0.2	1.0	2.0	0.01	5	0.01	− 1	0.5011	0.9732×10^{-3}	− 2.881	0.1200×10^{-3}	2.485
						+ 1	0.4989	0.9729×10^{-3}	− 2.844	0.1206×10^{-3}	2.602
0.2	1.0	4.0	0.02	5	0.01	− 1	0.5004	0.4760×10^{-3}	− 1.442	0.7595×10^{-4}	− 4.779
						+ 1	0.4996	0.4743×10^{-3}	− 1.397	0.7534×10^{-4}	− 3.740

4) The mean rate of heat transfer at the lower plate decreases with increasing Φ_2 but increases with increasing Φ_1.

5) The mean rate of heat transfer is not significantly affected by ω at both the plates.

6) An increase in E leads to an increase in the mean rate of heat transfer at the lower plate and a decrease at the upper plate.

7) An increase in M leads to an increase in the mean rate of heat transfer at the upper plate and a decrease at the lower plate.

8) The amplitudes of the first and second harmonics of the rate of heat transfer at both the plates decrease with increasing Φ_1, Φ_2, ω, or M.

9) There is always a phase lag in the first harmonic of the heat transfer at both the plates whereas in the case of the second harmonic there is a phase lead.

APPENDIX

$$b_{1,2}^2 = \{M^2 + i\omega(N+1) \pm \sqrt{[M^2 + i\omega(N+1)]^2 + 4\omega^2 N}\}/2, \quad b_3 = \sqrt{i\omega P}, \quad b_4 = \sqrt{2i\omega P}, \quad b_5 = b_1 + \bar{b}_1,$$
$$b_6 = b_1 - \bar{b}_1, \quad b_7 = b_1 + \bar{b}_2, \quad b_8 = b_1 - \bar{b}_2, \quad b_9 = b_2 + \bar{b}_1, \quad b_{10} = b_2 - \bar{b}_1, \quad b_{11} = b_2 + \bar{b}_2, \quad b_{12} = b_2 - \bar{b}_2,$$
$$b_{13} = M + b_1, \quad b_{14} = M - b_1, \quad b_{15} = M + b_2, \quad b_{16} = M - b_2, \quad b_{17} = b_1 + b_2, \quad b_{18} = b_1 - b_2,$$
$$a_1 = M^2 b_1 \bar{b}_1, \quad a_2 = M^2 b_1 \bar{b}_2, \quad a_3 = M^2 b_2 \bar{b}_1, \quad a_4 = M^2 b_2 \bar{b}_2,$$
$$C_1 = -(1/M^2 + C_2 \cosh M), \quad C_2 = -(2 + \Phi_1 + \Phi_2)/M\{(\Phi_1 + \Phi_2)M \cosh M + 2\sinh M\},$$
$$C_3 = (\Phi_2 - \Phi_1)C_1/2, \quad C_4 = (X_5X_7 - X_4X_8)/(X_3X_7 - X_4X_6), \quad C_5 = (X_3X_8 - X_5X_6)/(X_3X_7 - X_4X_6),$$
$$C_6 = -C_4X_1 \sinh b_1/X_2 \sinh b_2, \quad C_7 = -(1 + X_1C_5 \cosh b_1)/X_2 \cosh b_2, \quad C_8 = \tfrac{1}{2}[1 - P_1(1) - P_1(-1)],$$
$$C_9 = \tfrac{1}{2}[1 - P_1(1) + P_1(-1), \quad C_{10} = PE[P_2(1) + P_2(-1)]/\cosh b_3, \quad C_{11} = PE[P_2(1) - P_2(-1)]/\sinh b_3,$$
$$C_{12} = PE[P_3(1) + P_3(-1)]/4\cosh b_4, \quad C_{13} = PE[P_3(1) - P_3(-1)]/4\sinh b_4,$$
$$X_1 = b_1(M^2 + i\omega N - b_1^2), \quad X_2 = b_2(M^2 + i\omega N - b_2^2), \quad X_3 = (\Phi_1 - \Phi_2)(b_1X_2 - X_1b_2)\sinh b_1/X_2,$$
$$X_4 = b_1(\Phi_1 + \Phi_2)\cosh b_1 + 2\sinh b_1 - \{b_2(\Phi_1 + \Phi_2)\cosh b_2 + 2\sinh b_2\}X_1 \cosh b_1/X_2 \cosh b_2,$$
$$X_5 = \{b_2(\Phi_1 + \Phi_2)\cosh b_2 + 2\sinh b_2\}/X_2 \cosh b_2,$$

$X_6 = b_1(\Phi_1 + \Phi_2)\sinh b_1 + 2\cosh b_1 - \{b_2(\Phi_1 + \Phi_2)\sinh b_2 + 2\cosh b_2\}X_1 \sinh b_1/X_2 \sinh b_2,$

$X_7 = X_3 \cosh b_1/\sinh b_1, \quad X_8 = b_2(\Phi_1 - \Phi_2)/X_2, \quad X_9 = C_2^2/4, \quad X_{10} = 2C_2^2/M^2, \quad X_{11} = X_1C_4b_1/i\omega,$

$X_{12} = X_1C_5b_1/i\omega, \quad X_{13} = X_2C_6b_2/i\omega, \quad X_{14} = X_2C_7b_2/i\omega, \quad X_{15} = [X_{11}\overline{X}_{11} + X_{12}\overline{X}_{12} + a_1(C_4\overline{C}_4 + C_5\overline{C}_5)]/2b_5^2,$

$X_{16} = [X_{11}\overline{X}_{12} + X_{12}\overline{X}_{11} + a_1(C_4\overline{C}_5 + C_5\overline{C}_4)]/2b_5^2, \quad X_{17} = [X_{11}\overline{X}_{11} - X_{12}\overline{X}_{12} + a_1(-C_4\overline{C}_4 + C_5\overline{C}_5)]/2b_6^2,$

$X_{18} = [-X_{11}\overline{X}_{12} + X_{12}\overline{X}_{11} + a_1(C_4\overline{C}_5 - C_5\overline{C}_4)]/2b_6^2, \quad X_{19} = [X_{11}\overline{X}_{13} + X_{12}\overline{X}_{14} + a_2(C_4\overline{C}_6 + C_5\overline{C}_7)]/2b_7^2,$

$X_{20} = [X_{11}\overline{X}_{14} + X_{12}\overline{X}_{13} + a_2(C_4\overline{C}_7 + C_5\overline{C}_6)]/2b_7^2, \quad X_{21} = [X_{11}\overline{X}_{13} - X_{12}\overline{X}_{14} + a_2(-C_4\overline{C}_6 + C_5\overline{C}_7)]/2b_8^2,$

$X_{22} = [-X_{11}\overline{X}_{14} + X_{12}\overline{X}_{13} + a_2(C_4\overline{C}_7 - C_5\overline{C}_6)]/2b_8^2, \quad X_{23} = [X_{13}\overline{X}_{11} + X_{14}\overline{X}_{12} + a_3(C_6\overline{C}_4 + C_7\overline{C}_5)]/2b_9^2,$

$X_{24} = [X_{13}\overline{X}_{12} + X_{14}\overline{X}_{11} + a_3(C_6\overline{C}_5 + C_7\overline{C}_4)]/2b_9^2, \quad X_{25} = [X_{13}\overline{X}_{11} - X_{14}\overline{X}_{12} + a_3(-C_6\overline{C}_4 + C_7\overline{C}_5)]/2b_{10}^2,$

$X_{26} = [-X_{13}\overline{X}_{12} + X_{14}\overline{X}_{11} + a_3(C_6\overline{C}_5 - C_7\overline{C}_4)]/2b_{10}^2, \quad X_{27} = [X_{13}\overline{X}_{13} + X_{14}\overline{X}_{14} + a_4(C_6\overline{C}_6 + C_7\overline{C}_7)]/2b_{11}^2,$

$X_{28} = [X_{13}\overline{X}_{14} + X_{14}\overline{X}_{13} + a_4(C_6\overline{C}_7 + C_7\overline{C}_6)]/2b_{11}^2, \quad X_{29} = [X_{13}\overline{X}_{13} - X_{14}\overline{X}_{14} + a_4(-C_6\overline{C}_6 + C_7\overline{C}_7)]/2b_{12}^2,$

$X_{30} = [-X_{13}\overline{X}_{14} + X_{14}\overline{X}_{13} + a_4(C_6\overline{C}_7 - C_7\overline{C}_6)]/2b_{12}^2.$

In the following, $f(D) = D^2 - i\omega P$, $g(D) = D^2 - 2i\omega P$.

$X_{31} = C_2M(X_{11} - MC_4b_1)/2f(b_{13}), \quad X_{32} = C_2M(X_{12} - MC_5b_1)/2f(b_{13}), \quad X_{33} = C_2M(X_{11} + MC_4b_1)/2f(b_{14}),$

$X_{34} = -C_2M(X_{12} + MC_5b_1)/2f(b_{14}), X_{35} = C_2M(X_{13} - MC_6b_2)/2f(b_{15}), \quad X_{36} = C_2M(X_{14} - MC_7b_2)/2f(b_{15}),$

$X_{37} = C_2M(X_{13} + MC_6b_2)/2f(b_{16}), \quad X_{38} = -C_2M(X_{14} + MC_7b_2)/2f(b_{16}), \quad X_{39} = -C_4b_1/f(b_1),$

$X_{40} = -C_5b_1/f(b_1), \quad X_{41} = -C_6b_2/f(b_2), \quad X_{42} = -C_7b_2/f(b_2),$

$X_{43} = [X_{11}^2 - X_{12}^2 + X_{13}^2 - X_{14}^2 + M^2(-C_4^2b_1^2 + C_5^2b_1^2 - C_6^2b_2^2 + C_7^2b_2^2)]/2g(0),$

$X_{44} = [X_{11}^2 + X_{12}^2 + M^2b_1^2(C_4^2 + C_5^2)]/2g(b_1), \quad X_{45} = (X_{11}X_{12} + M^2C_4C_5b_1^2)/g(2b_1),$

$X_{46} = (X_{13}^2 + X_{14}^2 + M^2C_6^2b_2^2)/2g(2b_2), \quad X_{47} = (X_{13}X_{14} + M^2C_6C_7b_2^2)/g(2b_2),$

$X_{48} = [X_{11}X_{13} + X_{12}X_{14} + M^2(C_4C_6b_1b_2 + C_5C_7b_1b_2)]/g(b_{17}),$

$X_{49} = [X_{11}X_{14} + X_{12}X_{13} + M^2(C_4C_7b_1b_2 + C_5C_6b_1b_2)]/g(b_{17}),$

$X_{50} = [X_{11}X_{13} - X_{12}X_{14} + M^2b_1b_2(C_5C_7 - C_4C_6)]/g(b_{18}),$

$X_{51} = [-X_{11}X_{14} + X_{12}X_{13} + M^2b_1b_2(C_4C_7 - C_5C_6)]/g(b_{18}).$

[1]T. G. Cowling, *Magnetohydrodynamics* (Interscience, New York, 1957).

[2]S. I. Pai, *Magnetogasdynamics and Plasma Dynamics* (Springer-Verlag, New York, 1962).

[3]G. W. Sutton and A. Sherman, *Engineering Magnetohydrodynamics* (McGraw-Hill, New York, 1965).

[4]W. F. Hughes and Y. J. Young, *The Electromagnetodynamics of Fluids* (Wiley, New York, 1966).

J. A. Shercliff, J. Fluid Mech. **1**, 644 (1965).

[6]C. C. Chang and J. T. Yen, Z. Angew. Math. Phys. **13**, 266 (1962).

[7]V. M. Soundalgekar, Proc. Natl. Inst. Sci., India, Part A **35**, 329 (1969).

[8]V. M. Soundalgekar and J. P. Bhat (to be published).

Analysis of the effect of surfaces on the tricritical behavior of systems[a)]

G. Gumbs
Division of Chemistry, National Research Council of Canada, Ottawa K1A 0R6, Canada

(Received 10 February 1982; accepted for publication 20 August 1982)

The order parameter $\phi(z)$ of the ϕ^6-dominated tricritical free energy functional is calculated for film and half-space geometries. Extrapolation length (Λ) boundary conditions are used to simulate the effect of the surface. Closed-form expressions for $\phi(z)$ of a film are given in terms of Weierstrass elliptic functions, or, alternatively, Jacobi elliptic functions. For a half-space, $\phi(z)$ is expressed in terms of hyperbolic functions. In the absence of an external field, it is shown that the phase transitions which the system can undergo may be classified as ordinary ($\Lambda > 0$), surface ($\Lambda < 0$), and special ($\Lambda = \infty$), like the ϕ^4 theory for second-order phase transitions. The critical exponents for the order parameter at the surface are determined for each type of phase transition. A discussion of the free energy for the surface phase is also presented.

PACS numbers: 64.60.Kw, 64.60.Fr, 68.60. + q, 02.30. + q

I. INTRODUCTION

Progress in the theory of critical phenomena in *bulk* systems has been achieved with the use of the renormalization group (RG) method. The effects of surfaces on phase transitions have also received considerable attention recently, but the lack of translational invariance has made the application of the RG method much more difficult. Mean-field theories have also been applied to both bulk systems and systems with surfaces with considerable success. However, for a film of finite thickness even a mean-field theory (MFT) calculation of the correlation function above the critical temperature or of the order parameter in the ordered phase can be very involved.[1]

The phenomenological theory for systems with a tricritical point, such as occurs in He^3–He^4 mixtures, has been the subject of much discussion recently. However, with the exception of Binder and Landau,[2] all these studies have been confined to bulk systems.[3–9] In particular, a scaling theory for tricritical behavior has been developed by Riedel and Wegner.[3] These authors[3] as well as Stephen *et al.*[8] have noted that MFT for tricritical points is correct in three dimensions, apart from logarithmic corrections. (Similar logarithmic corrections are also needed in four dimensions for the MFT of the Ising spin model.[10]) Tricritical behavior in a metamagnetic single crystal, in the presence of an ordering field, has also been discussed recently. However, the Hamiltonian needed[11] is considerably more complicated than that of Riedel and Wegner.

In this paper, we consider the symmetrical tricritical point of Riedel and Wegner[3] in the presence of a surface. The system is described by an energy functional which contains extra terms arising from the surfaces at $z = 0$ and $z = L$:

$$F = \int d\mathbf{x}\,\{r_0\phi^2(\mathbf{x}) + \xi_0^2[\nabla\phi(\mathbf{x})]^2 + g_4\phi^4(\mathbf{x}) + \tfrac{1}{3}g_6\phi^6(\mathbf{x}) + (\xi_0^2/\Lambda)[\delta(z) + \delta(z-L)]\phi^2(\mathbf{x})\}. \tag{1.1}$$

Here $\phi(\mathbf{x})$ is a scalar order parameter and $r_0 \equiv \tau - 1$ where $\tau \equiv T/T_c^{\mathrm{MF}}(\infty)$, with $T_c^{\mathrm{MF}}(\infty)$ equal to the mean-field transition temperature for the bulk. $\mathbf{x} \equiv (\mathbf{x}_\parallel, z)$ is a spatial vector within the film, with $\mathbf{x}_\parallel$ parallel to the surface. The integration over z in (1.1) is from 0 to L, and ξ_0 is a temperature-independent length scale for the system. Λ is an extrapolation length whose significance is such that if the value of ϕ were extrapolated a distance Λ beyond each surface, ϕ would vanish there. This type of boundary condition has been used in many papers (see, for example, the references given by Cordery and Griffin[12]) which have studied the effects of a surface on the phase transition of spin systems using the continuous spin Ginzburg–Landau–Wilson (GLW) Hamiltonian. The coefficient g_4 in (1.1) depends on temperature as well as the interactions causing tricriticality.[3] At the tricritical point, g_4 vanishes while g_6 is finite and positive. Below the critical temperature T_c, in the mean-field treatment of the problem, g_4 is set equal to zero while the ϕ^6 term is retained in the energy functional (1.1).[13] Above T_c, when fluctuations are ignored, g_4 and g_6 are set equal to zero, and the two-point correlation function for a bounded system with a tricritical point is equal to that derived previously in MFT.[14] In the present paper, we derive expressions for the order parameter of the ϕ^6-dominated, tricritical free energy functional for film and half-space geometries. Owing to translational invariance parallel to the surface, ϕ depends only on the variable z.

With the free energy (1.1), we show that the system has an *ordinary*, *surface*, and *special* transition, depending on the value of the extrapolation length Λ. This classification follows that of Bray and Moore[15] and Lubensky and Rubin[14] for *usual* second-order phase transitions. For the *ordinary* transition, $\Lambda > 0$ and the system orders at the bulk transition temperature. For the *surface* transition, the surface orders spontaneously at a higher temperature than the bulk. For the *special* transition, $\Lambda = \infty$ and the system orders at the bulk transition temperature.

II. THE ORDER PARAMETER FOR A FILM BELOW T_c

Within MFT, the order parameter for a film is obtained by minimizing the free energy (1.1). Doing so and setting $g_4 = 0$, we obtain

$$\xi_0^2 \frac{d^2\phi(z)}{dz^2} = r_0\phi(z) + g_6\phi^5(z), \tag{2.1}$$

[a)] National Research Council of Canada No. 20489.

 0022-2488/83/010202-07$02.50

with the boundary conditions

$$\frac{d\phi(z)}{dz} = \frac{1}{\Lambda}\phi(z), \quad z = 0, \tag{2.2a}$$

$$\frac{d\phi(z)}{dz} = -\frac{1}{\Lambda}\phi(z), \quad z = L. \tag{2.2b}$$

With the use of these equations, one may show that in thermodynamic equilibrium the free energy is given by

$$F[\phi] = -\tfrac{2}{3}g_6\int_0^L dz\,\phi^6(z). \tag{2.3}$$

Multiplying the differential equation (2.1) by $d\phi(z)/dz$ and then integrating over z, we obtain

$$\xi_0^2\left(\frac{d\phi(z)}{dz}\right)^2 = r_0\phi^2(z) + \tfrac{1}{3}g_6\phi^6(z) + R, \tag{2.4}$$

where R is independent of the z coordinate. In general, the order parameter is either symmetric (S) or antisymmetric (A) about the midplane $z = L/2$ of the film. From symmetry considerations, $\phi(z)$ satisfies $d\phi(z)/dz = 0$ at $z = L/2$ for the S solution, whereas $\phi(z)$ satisfies $\phi(z = L/2) = 0$ for the A solution. Define a function $\psi(z)$ in terms of $\phi(z)$ by the equation

$$\phi^2(z) = \phi_0^2/[\psi(z) - c], \tag{2.5}$$

where c is independent of z. Substituting (2.5) into (2.4), we obtain after a little algebra

$$\xi_0^2\left(\frac{d\psi(z)}{dz}\right)^2 = \left(\frac{4R}{\phi_0^2}\right)\psi^3(z) + 4\left(r_0 - \frac{3cR}{\phi_0^2}\right)\psi^2(z) - 4c\left(2r_0 - \frac{3cR}{\phi_0^2}\right)\psi(z) + 4\left(r_0c^2 + \frac{1}{3}g_6\phi_0^4 - \frac{R}{\phi_0^2}c^3\right). \tag{2.6}$$

Choose c so that the ψ^2 term in (2.6) vanishes. This gives

$$c = r_0\phi_0^2/3R. \tag{2.7}$$

Also, choose R to be

$$R = -\tfrac{1}{3}g_6\phi_0^6/(3c + 1). \tag{2.8}$$

Therefore, $\phi_0^2 = \phi^2(z = L/2)$ for a *symmetric* solution, and we must have ϕ_0^2 *positive* for the S case. For an *antisymmetric* solution, however, ϕ_0^2 might be positive or negative. Substituting (2.7) and (2.8) into (2.6), we obtain

$$\frac{1}{4}\left(\frac{d\psi(v)}{dv}\right)^2 = \psi^3(v) - 3c^2\psi(v) - (1 + 3c - 2c^3), \tag{2.9}$$

where we have changed variables from z to v, with

$$v = (R^{1/2}/\phi_0\xi_0)(z - L/2). \tag{2.10}$$

In our notation, $\phi_0 \equiv |\phi_0^2|^{1/2}$.

The differential equation for ψ in (2.9) is the same as that satisfied by the Weierstrass elliptic function.[16] The roots e_1, e_2, and e_3 of the cubic equation

$$z^3 - 3c^2z - (1 + 3c - 2c^3) = 0 \tag{2.11}$$

are

$$1 + c, \quad -\left(\frac{1+c}{2}\right) + \frac{\Delta^{1/2}}{2} \quad \text{and} \quad -\left(\frac{1+c}{2}\right) - \frac{\Delta^{1/2}}{2}, \tag{2.12}$$

where Δ is the discriminant for the Weierstrass function and is given by

$$\Delta = 3(3c + 1)(c - 1). \tag{2.13}$$

For definiteness, we separate the two cases corresponding to Δ positive and negative, and adopt the following notation:

Case (i): If $\Delta < 0$, we choose $e_2 = 1 + c$ and denote the remaining (complex) roots by e_1 and e_3.

Case (ii): If $\Delta > 0$, we arrange the roots so that $e_1 > e_2 > e_3$. Therefore: (a) for $c \geqslant -\frac{1}{2}$, i.e., $c \in [-\frac{1}{2}, -\frac{1}{3})$ or $(1, \infty)$, we have

$$e_1 = 1 + c, \quad e_2 = -\left(\frac{1+c}{2}\right) + \frac{\Delta^{1/2}}{2},$$
$$e_3 = -\left(\frac{1+c}{2}\right) - \frac{\Delta^{1/2}}{2}; \tag{2.14}$$

(b) for $c < -\frac{1}{2}$

$$e_1 = -\left(\frac{1+c}{2}\right) + \frac{\Delta^{1/2}}{2}, \quad e_2 = 1 + c,$$
$$e_3 = -\left(\frac{1+c}{2}\right) - \frac{\Delta^{1/2}}{2}. \tag{2.15}$$

The solutions of (2.9) depend on the values of ϕ_0^2 and c. We now turn to calculating these solutions which are conveniently expressed in terms of Jacobi elliptic functions.

Region 1: $\phi_0^2 > 0$, $c < -\frac{1}{3}$

For $c < -\frac{1}{3}$, $\Delta > 0$ and there are four possible solutions. Introducing the variable u which is defined by

$$u = (1/\phi_0\xi_0)\sqrt{(e_1 - e_3)R}\;(z - L/2) \tag{2.16}$$

and defining the *modulus* k by

$$k \equiv \sqrt{(e_2 - e_3)/(e_1 - e_3)}, \tag{2.17}$$

the solutions are[17]

$$\psi_1\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_3 + \frac{e_1 - e_3}{\mathrm{sn}^2(u,k)}, \tag{2.18a}$$

$$\psi_2\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_2 + \frac{e_1 - e_2}{\mathrm{cn}^2(u,k)}, \tag{2.18b}$$

$$\psi_3\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_3 + (e_2 - e_3)\mathrm{sn}^2(u,k), \tag{2.18c}$$

$$\psi_4\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_1 - \frac{e_1 - e_2}{\mathrm{dn}^2(u,k)}. \tag{2.18d}$$

Here sn, cn, and dn are the sine, cosine, and delta amplitude Jacobi elliptic functions, respectively. sn(u) is antisymmetric whereas cn(u) and dn(u) are symmetric in the argument $u \to -u$. Therefore, since u is related to z by (2.16), the solution (2.18a) yields a solution for $\phi(z)$ in (2.5) which is antisymmetric about the midplane of the film. On the other hand, (2.18b)–(2.18d) yield symmetric solutions.

With the use of (2.5) and $\phi(z = L/2) = \phi_0$ for the symmetric solution, it is straightforward to show that for the *symmetric* case

$$\psi(0) = 1 + c. \tag{2.19}$$

However, setting $u = 0$ in (2.18b)–(2.18d), we find that

$$\psi_2(0) = e_1, \quad \psi_3(0) = e_3, \quad \text{and} \quad \psi_4(0) = e_2. \tag{2.20}$$

Therefore, referring to (2.14) and (2.15), we find that for the symmetric case, ψ_2 is the solution for $c \in (-\frac{1}{2}, -\frac{1}{3})$, ψ_4 is the solution for $c < -\frac{1}{2}$, and ψ_3 must be *discarded*.

Substituting (2.18b) into (2.5), we obtain the *symmetric* solution for the order parameter, where $c \in (-\frac{1}{2}, -\frac{1}{3})$:

$$\phi_S(z) = \phi_0 \mathrm{cn}(u,k)/[(e_1 - e_2) - (c - e_2)\mathrm{cn}^2(u,k)]^{1/2}. \tag{2.21a}$$

Since (2.21a) satisfies the boundary conditions (2.2), we have[18]

$$(e_1 - e_3)^{1/2}(e_1 - e_2)R^{1/2}\mathrm{sc}(u_0,k)\mathrm{dn}(u_0,k) = (\phi_0\xi_0/\Lambda)[(e_1 - e_2) - (c - e_2)\mathrm{cn}^2(u_0,k)], \tag{2.21b}$$

where

$$u_0 \equiv (L/2\phi_0\xi_0)\sqrt{(e_1 - e_3)R}\,. \tag{2.21c}$$

For $c < -\frac{1}{2}$, the symmetric solution is obtained from (2.5) and (2.18d):

$$\phi_S(z) = \phi_0 \mathrm{dn}(u,k)/[(e_1 - c)\mathrm{dn}^2(u,k) - (e_1 - e_2)]^{1/2}. \tag{2.21d}$$

Imposing the boundary conditions (2.2) on (2.21d), we have

$$(e_1 - e_3)^{1/2}(e_1 - e_2)k^2R^{1/2}\mathrm{sn}(u_0,k)\mathrm{cd}(u_0,k) = (\phi_0\xi_0/\Lambda)[(e_1 - e_2) - (e_1 - c)\mathrm{dn}^2(u_0,k)]. \tag{2.21e}$$

Substituting (2.18a) into (2.5), we obtain the antisymmetric solution for ϕ when $c < -\frac{1}{3}$:

$$\phi_A(z) = \phi_0 \mathrm{sn}(u,k)/[(e_1 - e_3) - (c - e_3)\mathrm{sn}^2(u,k)]^{1/2}. \tag{2.22a}$$

We must have from (2.22a) and (2.2)

$$(e_1 - e_3)^{3/2}R^{1/2}\mathrm{cs}(u_0,k)\mathrm{dn}(u_0,k) = -(\phi_0\xi_0/\Lambda)[(e_1 - e_3) - (c - e_3)\mathrm{sn}^2(u_0,k)]. \tag{2.22b}$$

In this region, the values of c, R, and ϕ_0 are determined by (2.7), (2.8), and (2.21b) or (2.21e) for the symmetric solutions and from (2.7), (2.8), and (2.22b) for the antisymmetric solution. Substituting these values into (2.21a), (2.21d), and (2.22a), we obtain the solutions for $\phi(z)$ in terms of the original parameters in (2.1) and (2.2).

For $c < -\frac{1}{3}$, we deduce from (2.8) that $R > 0$ since $g_6 > 0$ and ϕ_0 is real ($\phi_0^2 > 0$) as shown in Fig. 1. Therefore, in region 1, $r_0 < 0$.

Regions 2 and 3: $\phi_0^2 < 0$, $-\frac{1}{3} < c < 1$

For this range of values for c, the discriminant $\Delta < 0$, and there are two possible solutions for ψ which satisfy (2.9). These are

$$\psi_1\left(\frac{u'}{H_2^{1/2}}\right) = e_2 - H_2\left(\frac{\mathrm{cn}(u',\kappa)}{\mathrm{sn}(u',\kappa)\mathrm{dn}(u',\kappa)}\right)^2, \tag{2.23}$$

$$\psi_2\left(\frac{u'}{H_2^{1/2}}\right) = e_2 - H_2\left(\frac{\mathrm{sn}(u',\kappa)\mathrm{dn}(u',\kappa)}{\mathrm{cn}(u',\kappa)}\right)^2, \tag{2.24}$$

where (2.24) is obtained from (2.23) by replacing u' by *either* $u' + K$ *or* $u' + iK'$, where K and K' are complete elliptic integrals.[16] We have now introduced

$$H_2^2 \equiv |e_1 - e_2|^2 = \tfrac{1}{4}[9(1 + c)^2 + |\Delta|], \tag{2.25a}$$

$$u' = (1/\phi_0\xi_0)\sqrt{H_2R}\,(z - L/2), \tag{2.25b}$$

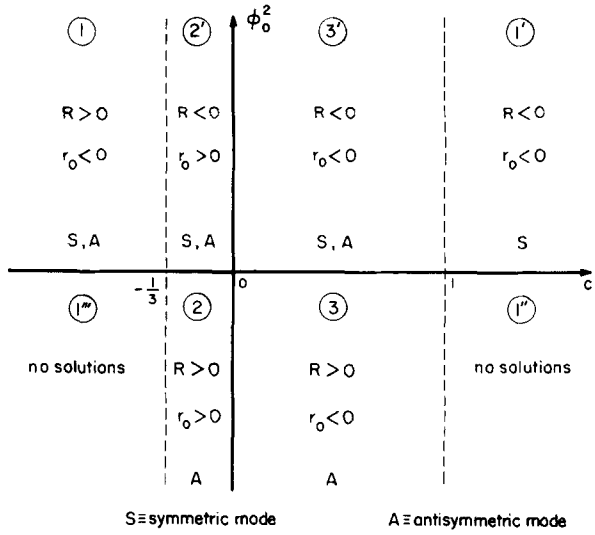


FIG. 1. Summary of the results for the order parameter, within mean-field theory, of a film. The solutions are presented in Sec. II.

$$\kappa \equiv \sqrt{\tfrac{1}{2} + 3e_2/2H_2}. \tag{2.25c}$$

Referring to Fig. 1, we see that in regions 2 and 3, $\phi_0^2 < 0$. But, for a symmetric solution ϕ_0 has to be real. Therefore, for this range of values of with $\phi_0^2 < 0$, ψ_2 must be discarded and the solution is antisymmetric. Therefore, we have

$$\phi_A(z) = \phi_0 \left[H_2\left(\frac{\mathrm{cn}(u',\kappa)}{\mathrm{sn}(u',\kappa)\mathrm{dn}(u',\kappa)}\right)^2 - 1\right]^{-1/2}, \tag{2.26a}$$

where, upon substituting (2.26a) into (2.2),

$$H_2^{3/2}R^{1/2}[1 - 2\kappa^2\mathrm{sn}^2(u_0',\kappa) + \kappa^2\mathrm{sn}^4(u_0',\kappa)] = -(\phi_0\xi_0/\Lambda)\mathrm{sc}(u_0',\kappa)\mathrm{dn}(u_0',\kappa)[H_2\mathrm{cn}^2(u_0',\kappa) - \mathrm{sn}^2(u_0',\kappa)\mathrm{dn}^2(u_0',\kappa)]. \tag{2.26b}$$

Here

$$u_0' \equiv (L/2\phi_0\xi_0)(H_2R)^{1/2}, \tag{2.26c}$$

and we stress that $\phi_0 \equiv |\phi_0^2|^{1/2}$. The values of c, R, and ϕ_0 are now given by (2.7), (2.8), and (2.26b). Equation (2.26a) thus gives the solution for ϕ in terms of the parameters in (2.1) and (2.2).

From (2.7), we find that in region 2, $r_0 > 0$, but $r_0 < 0$ in region 3. From (2.8), we deduce that $R > 0$.

Regions 2' and 3': $\phi_0^2 > 0$, $-\frac{1}{3} < c < 1$

Here the discriminant Δ of the Weierstrass function is negative and the solutions may be obtained from Eqs. (2.23) and (2.24) by making the replacement $u' \to iu'$. That is (see 8.153 of Gradshteyn and Ryzhik[16]), the solutions are

$$\psi_1\left(\frac{u'}{H_2^{1/2}}\right) = e_2 + H_2\left(\frac{\mathrm{cn}(u',\kappa')}{\mathrm{sn}(u',\kappa')\mathrm{dn}(u',\kappa')}\right)^2, \tag{2.27a}$$

$$\psi_2\left(\frac{u'}{H_2^{1/2}}\right) = e_2 + H_2\left(\frac{\mathrm{sn}(u',\kappa')\mathrm{dn}(u',\kappa')}{\mathrm{cn}(u',\kappa')}\right)^2, \tag{2.27b}$$

where the argument and modulus are, respectively,

$$u' = (1/\phi_0\xi_0)\sqrt{H_2|R|}\,(z - L/2), \tag{2.27c}$$

$$\kappa' \equiv \sqrt{\tfrac{1}{2} - 3e_2/2H_2}. \tag{2.27d}$$

Since u' is related to the z coordinate by (2.27c), and $\phi_0^2 > 0$ in this region (see Fig. 1), ψ_1 (ψ_2) gives an S (A) solution for ϕ. Note that ψ_2 satisfies the condition (2.19) for a symmetric solution. Substituting (2.27b) into (2.5), we obtain

$$\phi_S(z) = \phi_0 \left[1 + H_2\left(\frac{\mathrm{sn}(u',\kappa')\mathrm{dn}(u',\kappa')}{\mathrm{cn}(u',\kappa')}\right)^2\right]^{-1/2}. \quad (2.28a)$$

Substituting (2.28a) into the boundary conditions defined in (2.2), we obtain

$$H_2^{3/2}|R|^{1/2}\mathrm{sc}(u_0',\kappa')\mathrm{dn}(u_0',\kappa') \times [1 - 2\kappa'^2\mathrm{sn}^2(u_0',\kappa') + \kappa'^2\mathrm{sn}^4(u_0',\kappa')] = (\phi_0\xi_0/\Lambda)[H_2\mathrm{sn}^2(u_0',\kappa')\mathrm{dn}^2(u_0',\kappa') + \mathrm{cn}^2(u_0',\kappa')], \quad (2.28b)$$

where

$$u_0' \equiv (L/2\phi_0\xi_0)(H_2|R|)^{1/2}. \quad (2.28c)$$

To get the antisymmetric solution, we substitute (2.27a) into (2.5) and obtain

$$\phi_A(z) = \phi_0 \left[1 + H_2\left(\frac{\mathrm{cn}(u',\kappa')}{\mathrm{sn}(u',\kappa')\mathrm{dn}(u',\kappa')}\right)^2\right]^{-1/2}. \quad (2.29a)$$

Imposing the boundary condition (2.2) on (2.29a), we have

$$H_2^{3/2}|R|^{1/2}[1 - 2\kappa^2\mathrm{sn}^2(u_0',\kappa') + \kappa'^2\mathrm{sn}^4(u_0',\kappa')] = -(\phi_0\xi_0/\Lambda)[\mathrm{sc}(u_0',\kappa')\mathrm{dn}(u_0',\kappa')][H_2\mathrm{cn}^2(u_0',\kappa') + \mathrm{sn}^2(u_0',\kappa')\mathrm{dn}^2(u_0',\kappa')]. \quad (2.29b)$$

Equations (2.7), (2.8), and (2.28b) or (2.29b) determine the values of c, R, and ϕ_0 in this region. As shown in Fig. 1, $r_0 > 0$ in region 2', but $r_0 < 0$ in region 3'. We have $R < 0$ for this case. These results agree with (2.7) and (2.8).

Region 1': $\phi_0^2 > 0$, $c > 1$

For $c > 1$, the discriminant Δ of the Weierstrass function is positive. Also, in this region, $\phi_0^2 > 0$ and, therefore, from (2.8), $R < 0$. The solutions may be obtained from those in region 1 with the replacement $u \to iu$. Making use of the results in 8.153 of Gradshteyn and Ryzhik,[16] the solutions are

$$\psi_1\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_1 - \frac{(e_1 - e_3)}{\mathrm{sn}^2(u,k')}, \quad (2.30a)$$

$$\psi_2\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_2 + (e_1 - e_2)\mathrm{cn}^2(u,k'), \quad (2.30b)$$

$$\psi_3\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_2 - \frac{(e_2 - e_3)}{\mathrm{cn}^2(u,k')}, \quad (2.30c)$$

$$\psi_4\left(\frac{u}{\sqrt{e_1 - e_3}}\right) = e_3 + \frac{(e_2 - e_3)}{\mathrm{dn}^2(u,k')}, \quad (2.30d)$$

where

$$u = (1/\phi_0\xi_0)\sqrt{(e_1 - e_3)|R|}\,(z - L/2), \quad (2.31a)$$

$$k' \equiv \sqrt{(e_1 - e_2)/(e_1 - e_3)}. \quad (2.31b)$$

For $c > 1$, e_1, e_2, and e_3 are given by (2.14). Therefore, $(e_1 - c)\mathrm{sn}^2(u,k') - (e_1 - e_3) = \mathrm{sn}^2(u,k') - (e_1 - e_3) < 0$. That is, ψ_1 in (2.30a) does not give a real value for $\phi(z)$ in (2.5) since $\phi_0^2 > 0$. The solution for ϕ, therefore, cannot be antisymmetric. From Eqs. (2.30b)–(2.30d), we have $\psi_2 \leqslant e_1$, $\psi_3 \leqslant e_3$, and $\psi_4 \geqslant e_2$. Setting $u = 0$ in these equations, we obtain

$$\psi_2(0) = e_1 = 1 + c, \quad \psi_3(0) = e_3, \quad \psi_4(0) = e_2. \quad (2.32)$$

Since a symmetric solution must satisfy (2.5), only ψ_2 is acceptable and the solution for $c > 1$ and $\phi_0^2 > 0$ is therefore

$$\phi_S(z) = \phi_0/[(e_1 - e_2)\mathrm{cn}^2(u,k') - (c - e_2)]^{1/2}. \quad (2.33a)$$

The boundary condition (2.2) and (2.33a) give

$$(e_1 - e_3)^{1/2}(e_1 - e_2)|R|^{1/2}\mathrm{sn}(u_0,k')\mathrm{cn}(u_0,k')\mathrm{dn}(u_0,k') = -(\phi_0\xi_0/\Lambda)[(e_1 - e_2)\mathrm{cn}^2(u_0,k') - (c - e_2)], \quad (2.33b)$$

where, in (2.33b)

$$u_0 \equiv (L/2\phi_0\xi_0)\sqrt{(e_1 - e_3)|R|}. \quad (2.33c)$$

Equations (2.7), (2.8), and (2.33b) together determine the values of c, R, and ϕ_0. Moreover, from (2.7) we find that $r_0 < 0$.

In region 1'' of Fig. 1, $\phi_0^2 < 0$ and $c > 1$; however, for $c > 1$, the discriminant in (2.13) satisfies $\Delta > 0$ and the solutions for ψ in this region are given by Eqs. (2.18). One may deduce from (2.18a) that $\psi_1 \geqslant e_1$, where $e_1 = 1 + c$ from (2.14). That is, $\psi_1 > c$. Therefore, since $\phi^2 = |\phi_0^2|/(c - \psi)$, ψ_1 does not give a *real* value for the order parameter, i.e., there is *no* antisymmetric solution. ϕ_0 has to be real for there to be a symmetric solution. Therefore, we conclude that, in region 1'', there are no real solutions for $\phi(z)$.

In region 1''' of Fig. 1, where $\phi_0^2 < 0$ and $c < -\frac{1}{3}$, the discriminant $\Delta > 0$ and the solutions for ψ are given by (2.18). However, the solution for ϕ cannot be symmetric, since $\phi_0^2 < 0$ and thus (2.18b)–(2.18d) must be ruled out. For the same reason given in our considerations of region 1'', ψ_1 does not give a real value for ϕ in region 1'''. We conclude, therefore, that there are no real solutions for $\phi(z)$ in region 1'''.

In our discussion above, we did not consider the case for which $c = -\frac{1}{3}$. This value for c needs special consideration. Eliminating R from (2.7) and (2.8), we obtain $\phi_0^4 = r_0(3c + 1)/6g_6$. That is ϕ_0 *vanishes* when $c = -\frac{1}{3}$. For this value of c, it follows from (2.13) that $\Delta = 0$ and from (2.12) that $e_1 = \frac{2}{3}$, $e_2 = e_3 = -\frac{1}{3}$. Using the result in 8.1698 of Gradshteyn and Ryzhik,[16] we find that for $c = -\frac{1}{3}$ ($\Delta = 0$)

$$\psi_S(z) = -\tfrac{1}{3} + \frac{1}{\cos^2((z - L/2)(-r_{0c})^{1/2}/\xi_0)} \quad (2.34a)$$

is the *symmetric* solution for (2.9) and

$$\psi_A(z) = -\tfrac{1}{3} + \frac{1}{\sin^2((z - L/2)(-r_{0c})^{1/2}/\xi_0)} \quad (2.34b)$$

is the *antisymmetric* solution. Here $r_{0c} \equiv T_c(L)/T_c^{MF}(\infty) - 1$, where $T_c(L)$ is the transition temperature for a film.

Making use of (2.34) in (2.5), assuming that ϕ_0 is infinitesimal, and then substituting the result into the boundary conditions (2.2), we obtain (after cancelling ϕ_0)

$$(-r_{0c})^{1/2}\tan\left(\frac{L}{2\xi_0}(-r_{0c})^{1/2}\right) = \frac{\xi_0}{\Lambda} \quad (2.35a)$$

from (2.34a) and

$$(-r_{0c})^{1/2}\cot\left(\frac{L}{2\xi_0}(-r_{0c})^{1/2}\right) = \frac{\xi_0}{\Lambda} \quad (2.35b)$$

from (2.34b). There are two cases to consider, depending on

the sign of r_{0c}. If $r_{0c} < 0$, the argument of the tangent and cotangent in (2.35) is real, and there are an infinite number of solutions for these equations. However, the physical T_c corresponds to the largest value. With $\Lambda > 0$, this comes from Eq. (2.35a). That is, for $\Lambda > 0$ the reduced transition temperature is

$$T_c(L)/T_c^{\mathrm{MF}}(\infty) = 1 - x_S^2, \tag{2.36a}$$

where x_S is the smallest solution of

$$x_S \tan(Lx_S/2\xi_0) = \xi_0/\Lambda. \tag{2.36b}$$

For $\Lambda < 0$, it is (2.35a) with $r_{0c} > 0$ which gives the largest possible value for T_c. That is, for $\Lambda < 0$, the reduced transition temperature is

$$T_c(L)/T_c^{\mathrm{MF}}(\infty) = 1 + x_S^2, \tag{2.37a}$$

where x_S is the solution of

$$x_S \tanh(Lx_S/2\xi_0) = \xi_0/|\Lambda|. \tag{2.37b}$$

These results agree with the well-known results for the transition temperature of a film, within MFT, for the usual second-order phase transitions.

It is a simple matter to verify that (2.36) and (2.37) may be obtained by taking the limit $c \to -\frac{1}{3}$ from within either region 1, 2, or 2'. It also becomes apparent, when taking this limit, that for the system to have an order-disorder transition, region 1 *favors* a positive value of the extrapolation length Λ, whereas regions 2 and 2' *favor* a negative value of Λ.

For $c = 1$, the discriminant $\Delta = 0$. The solutions for $\phi(z)$ may be constructed from the degenerate results for the Weierstrass function, as we did for $c = -\frac{1}{3}$. This completes our discussion of the solution of Eq. (2.1) subject to the boundary condition (2.2).

III. THE ORDER PARAMETER FOR THE HALF-SPACE BELOW T_c

The calculation for the order parameter for a film of finite thickness is complicated by the fact that the value of R in (2.4) is given implicitly by a set of three coupled equations involving the boundary conditions (2.2). Furthermore, it does not help to rewrite (2.4) as an integral equation since, for arbitrary values of R, we cannot do the integral analytically [we had to make use of the transformation (2.5)]. However, for the semi-infinite geometry ($z \geqslant 0$) the calculation is considerably simplified. In this case, we take the limit $L \to \infty$ and satisfy the boundary condition (2.2a) at $z = 0$ only. There is no need to separate the solutions into symmetric and antisymmetric cases for this geometry. In addition, the value of R is obtained from the condition that $d\phi(z)/dz = 0 = d^2\phi(z)/dz^2$ at $z = \infty$. With this value of R, we now show that (2.4) is easily integrated to give $\phi(z)$ explicitly. The analysis proceeds in much the same way as the ϕ^4 theory.[19]

Setting $d^2\phi(z)/dz^2 = 0$ at $z = \infty$ into Eq. (2.1), we find that $\phi(z = \infty) = 0$ or $(-r_0/g_6)^{1/4}$. These two values for $\phi(\infty)$ lead to the following considerations.

Case (i): $\phi(\infty) = 0$

Here the extrapolation length $\Lambda < 0$ and the temperature must satisfy $1 < \tau < \tau_s$, where $\tau_s \equiv 1 + \xi_0^2/\Lambda^2$. In this case $R = 0$ and (2.4) may be rewritten as an integral:

$$\frac{z}{\xi_0} = -\int_{\phi(0)}^{\phi(z)} d\phi \, \frac{1}{\phi(r_0 + \frac{1}{3}g_6\phi^4)^{1/2}}, \tag{3.1}$$

where the minus sign on the right-hand side of (3.1) is chosen since, for $\Lambda < 0$, the value of ϕ at the surface is larger than in the *bulk*. We now change variables in (3.1) from ϕ to θ, with $\phi^2 = (r_0/\frac{1}{3}g_6)^{1/2}\sinh\theta$; θ is real since $r_0 > 0$. After a straightforward calculation, we obtain

$$\phi(z) = \frac{|\Lambda|^{1/2}\phi(0)}{[\xi_+ \sinh(2z/\xi_+) + |\Lambda|\cosh(2z/\xi_+)]^{1/2}}, \tag{3.2}$$

where

$$\phi(0) \equiv [(3/g_6)(\tau_s - \tau)]^{1/4} \tag{3.3a}$$

and

$$\xi_+(T) \equiv \xi_0/(\tau - 1)^{1/2}. \tag{3.3b}$$

From (3.2), we find that for $z \gg \xi_+$

$$\phi(z) \approx \left(\frac{12(\tau_s - \tau)(\tau - 1)}{g_6[\xi_0/|\Lambda| + (\tau - 1)^{1/2}]^2}\right)^{1/4} e^{-z/\xi_+}. \tag{3.4}$$

Therefore, the surface orders at τ_s [see (3.3a)] while the bulk orders at the mean-field bulk transition temperature. That is, for $\Lambda < 0$, the surface orders spontaneously at a higher temperature than the bulk. This behavior has also been shown for *usual* second-order phase transitions and is classified as a *surface* transition.[14,15]

Case (ii): $\phi(\infty) = (-r_0/g_6)^{1/4}$

For this case, $\Lambda \gtrless 0$ and $\tau < 1$ (i.e., $r_0 < 0$). We can express R of (2.4) in terms of $\phi(\infty)$. Substituting this value for R into (2.4), we rewrite this equation as an integral over ϕ:

$$\frac{z}{\xi_0} = \frac{1}{(g_6/3)^{1/2}} \int_{\phi(0)}^{\phi(z)} d\phi \, \frac{1}{[\phi^2(\infty) - \phi^2][\phi^2 + 2\phi^2(\infty)]^{1/2}}. \tag{3.5}$$

Changing the variable of integration from ϕ to θ where $\phi = \sqrt{2}\,\phi(\infty)\sinh\theta$, one may easily do the integral in (3.5), using 2.441(3) of Gradshteyn and Ryzhik.[16] The result is

$$\phi(z) = \sqrt{2}\,\phi(\infty)\left\{\left[\frac{\sqrt{3} + 3\tanh(z/\xi_-)\tanh\theta_0}{\tanh(z/\xi_-) + \sqrt{3}\tanh\theta_0}\right]^2 - 1\right\}^{-1/2}, \tag{3.6a}$$

where

$$\xi_-(T) \equiv \xi_0/(1 - \tau)^{1/2} \tag{3.6b}$$

and θ_0 is given by

$$\phi(0) = \sqrt{2}\,\phi(\infty)\sinh\theta_0. \tag{3.7}$$

With the use of the boundary condition (2.2a) at $z = 0$, together with the value of $\phi(\infty)$, one finds that $x \equiv \sinh^2\theta_0$ is given by

$$x^3 - \tfrac{3}{4}Qx + \tfrac{1}{4} = 0, \tag{3.8}$$

where

$$Q \equiv \left(\frac{1}{-r_0}\right)\left(\frac{\xi_0}{\Lambda}\right)^2 + 1. \tag{3.9}$$

The three roots of the cubic equation (3.8) are real: Two are positive in value and the other negative. The two positive

roots which are physically meaningful are

$$x_1 = Q^{1/2}\cos(C/3), \tag{3.10a}$$

$$x_2 = \tfrac{1}{2}Q^{1/2}[\sqrt{3}\sin(C/3) - \cos(C/3)], \tag{3.10b}$$

where C satisfies $\pi/2 \leqslant C < \pi$ and is given by

$$\cos C = -Q^{-3/2}. \tag{3.11}$$

Thus, for $r_0 < 0$, $\phi(z)$ has two solutions: One corresponds to $\Lambda > 0$ and the other to $\Lambda < 0$. We next identify these two solutions.

In the limit $r_0 \to 0^-$, C is given by

$$C \approx \pi/2 + (-r_0)^{3/2}(|\Lambda|/\xi_0)^3. \tag{3.12}$$

Making use of (3.12) in (3.10), we obtain in this limit

$$x_1 \approx \frac{\sqrt{3}}{2}\left(\frac{1}{-r_0}\right)^{1/2}\left(\frac{\xi_0}{|\Lambda|}\right), \tag{3.13a}$$

$$x_2 \approx \tfrac{1}{3}(-r_0)\left(\frac{|\Lambda|}{\xi_0}\right)^2. \tag{3.13b}$$

Substituting (3.13) into (3.7), we obtain for $r_0 \to 0^-$

$$\phi(0) \approx \left(\frac{3}{g_6}\right)^{1/4}\left(\frac{\xi_0}{|\Lambda|}\right)^{1/2}, \quad \text{for } x_1, \tag{3.14a}$$

$$\phi(0) \approx \left(\frac{4}{9g_6}\right)^{1/4}\left(\frac{|\Lambda|}{\xi_0}\right)(-r_0)^{3/4}, \quad \text{for } x_2. \tag{3.14b}$$

That is, in the limit $\tau \to 1^-$, $\phi(0)$ remains finite [see (3.14a)] or tends to zero [see (3.14b)]. However, we showed above that, for negative extrapolation length, the surface orders at a higher temperature than the bulk. Therefore, when $\tau < 1$, the x_1 solution in (3.10a) corresponds to $\Lambda < 0$, whereas the x_2 solution in (3.10b) corresponds to $\Lambda > 0$. From (3.14b) and the definition for $\phi(\infty)$, we find that, for $\Lambda > 0$, the surface layer and the bulk order at the mean-field transition temperature $T_c^{\mathrm{MF}}(\infty)$. This corresponds to the *ordinary* transition, discussed for usual second-order phase transitions.[14,15]

If we now let $\Lambda \to \infty$ in Eq. (3.9), we find that $Q = 1$. In this case, Eq. (3.8) has a root at $x = -1$ and a *double* root at $x = \frac{1}{2}$. The negative root must be ruled out on physical grounds. Substituting the value for the double root into (3.7), we obtain

$$\phi(0) = \phi(\infty). \tag{3.15a}$$

For $x = \frac{1}{2}$, $\tanh\theta_0 = \pm 1/\sqrt{3}$. Substituting into (3.6a), we obtain the result

$$\phi(z) = \phi(0), \tag{3.15b}$$

for $\tau < 1$ and $\Lambda = \infty$. That is the order parameter is *flat* right up to the surface and the system orders at the bulk mean-field transition temperature. Following Bray and Moore,[15] we refer to this as the *special* transition.

Figure 2 is a plot of $\phi(z)/\phi(0)$ for $\tau > 1$ and $\Lambda < 0$, using the result in (3.2). Figure 3 is a plot of $\phi(z)/\phi(0)$ in (3.6) for $\tau < 1$: Curve I corresponds to the $\Lambda < 0$ and curve II to $\Lambda > 0$.

The surface order parameter $\phi(0)$ varies as $(T_c - T)^{\beta_1}$ as $T \to T_c^-$. For the *surface* transition, it follows from (3.3a) that

$$\beta_1 = \tfrac{1}{4}. \tag{3.16a}$$

Equation (3.14) gives

$$\beta_1 = \tfrac{3}{4} \tag{3.16b}$$

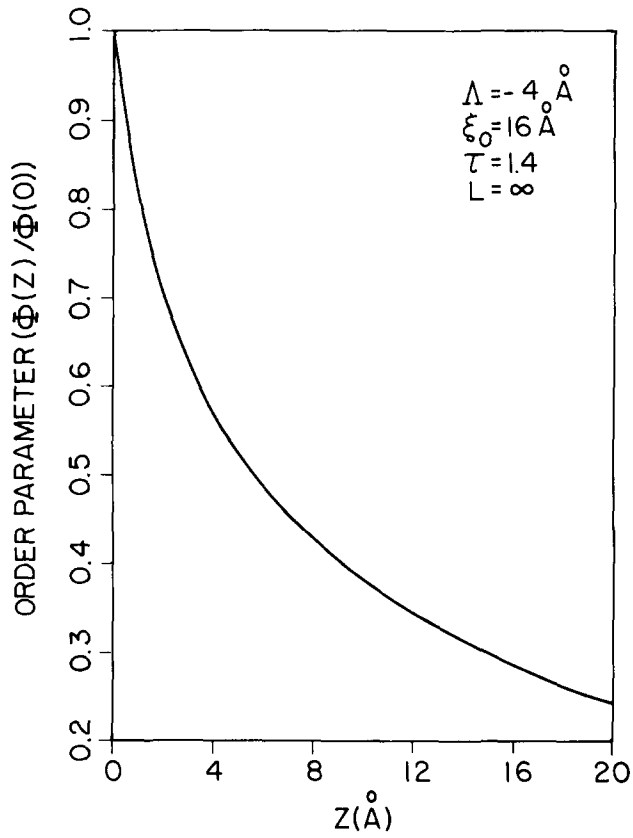


FIG. 2. The order parameter for a surface transition ($\Lambda < 0$), for a half-space, as a function of distance from the surface at $z = 0$. The plot is based on the result in (3.2).

for the *ordinary* transition and (3.15a) gives

$$\beta_1 = \tfrac{1}{4} \tag{3.16c}$$

for the *special* transition. For the surface transition, the exponent β is equal to its bulk value. This result is expected

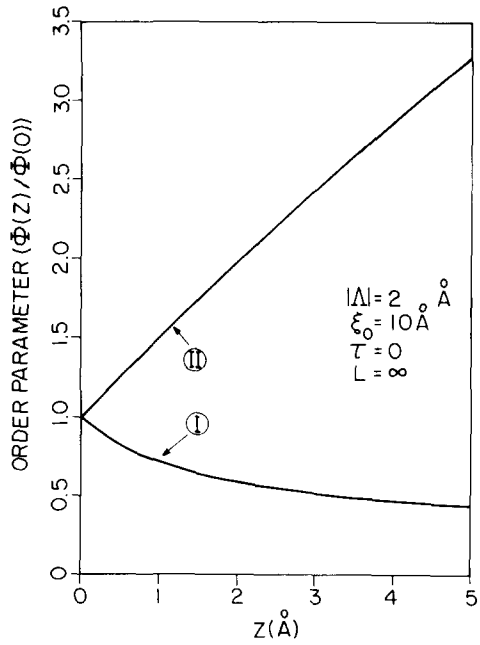


FIG. 3. The order parameter for a half-space, as a function of the distance z from the surface at $z = 0$. The plots are based on the results in (3.6)–(3.11). Curve I corresponds to a surface transition ($\Lambda < 0$) and curve II to an ordinary transition ($\Lambda > 0$).

since the exponent for the surface of a d-dimensional system is equal to the exponent for a $(d-1)$-dimensional system; in MFT the exponents are independent of dimensionality.[15] Although the system orders at the bulk transition temperature for both the *ordinary* and *special* transitions, the thermodynamic exponent β is not the same for these two types of phase transition. The difference is due to the relative value of the mean field in the surface layer compared to that in the bulk.

IV. THE FREE ENERGY FOR THE SURFACE ($\Lambda < 0$) PHASE WHEN $L = \infty$

When we substitute (3.2) for $\phi(z)$ into (2.3), we obtain the free energy for a half-space with $\Lambda < 0$ and $1 < \tau < \tau_s \equiv 1 + \xi_0^2/\Lambda^2$:

$$F_s(\tau) = -\tfrac{1}{3}g_6|\Lambda|^3\xi_+ \frac{\phi^6(0)}{(\xi_+^2 - \Lambda^2)^{3/2}} \int_0^\infty dz \frac{1}{\sinh^3(z + x_0)}, \tag{4.1}$$

where

$$x_0 \equiv \operatorname{arctanh}(|\Lambda|/\xi_+). \tag{4.2}$$

Doing the integration in (4.1), we obtain

$$F_s(\tau) = \tfrac{1}{8}g_6|\Lambda|^3\xi_+ \frac{\phi^6(0)}{(\xi_+^2 - \Lambda^2)^{3/2}} \left[\frac{\cosh x_0}{\sinh^2 x_0} - \arctan\left(\frac{1}{\cosh x_0}\right)\right]. \tag{4.3}$$

From (4.3), one finds that near τ_s, the *surface* free energy behaves asymptotically as

$$F_s(\tau) \sim (\tau_s - \tau)^{3/2}. \tag{4.4}$$

Since the specific heat $c_v = -T\partial^2 F/\partial T^2$, we find from (4.4) that the critical exponent α_s for the surface specific heat is, within MFT,

$$\alpha_s = \tfrac{1}{2}. \tag{4.5}$$

Josephson's law

$$\nu d = 2 - \alpha, \tag{4.6}$$

which involves the space dimension d, relates the exponent ν for the correlation function and the exponent α for the specific heat. Substituting $\nu = \frac{1}{2}$ from (3.4) and $\alpha = \frac{1}{2}$ from (4.5), one finds that Josephson's law is violated for the surface phase when the exponents have their classical value, except for $d = 3$. When the contribution from fluctuations is included, the corrected critical exponents only satisfy Josephson's law for $d < 3$. In fact, the width of the region in which the nonclassical exponents are observed shrinks to zero as d approaches 3.

V. CONCLUSIONS

We have presented closed-form solutions, within mean-field theory, for the order parameter $\phi(z)$ of the ϕ^6-dominated tricritical free energy functional for film and half-space geometries. We have discussed the solutions whose properties depend on the value of the extrapolation length.

Our results may be applicable in calculating the order parameter profile for a half-space, using the renormalization group ϵ expansion.[20] Bray and Moore[21] have also used the ϵ expansion to determine the shift exponent λ for *usual* second-order transitions of a thick film. Their technique may be used to calculate λ for a system having a tricritical point. These calculations are presently being done, and the results will be reported elsewhere.

ACKNOWLEDGMENTS

The author wishes to thank Professor Allan Griffin for suggesting this problem and for his helpful suggestions. Dr. K. M. Hong has made some invaluable comments for which the author is very grateful.

[1] G. Gumbs and A. Griffin, Can. J. Phys. **57**, 1686 (1979).
[2] K. Binder and D. P. Landau, Surface Sci. **61**, 577 (1976).
[3] E. K. Riedel and F. J. Wegner, Phys. Rev. Lett. **29**, 349 (1972).
[4] F. J. Wegner and E. K. Riedel, Phys. Rev. B **7**, 248 (1973).
[5] M. J. Stephen and J. L. McCauley, Phys. Lett. A **44**, 89 (1973).
[6] D. J. Amit and C. T. De Dominicis, Phys. Lett. A **45**, 193 (1973).
[7] R. B. Griffiths, Phys. Rev. B **7**, 545 (1973).
[8] M. J. Stephen, E. Abrahams, and J. P. Straley, Phys. Rev. B **12**, 256 (1975).
[9] G. Forgács and A. Zawadowski, Acta Phys. Acad. Sci. Hung. **42**, 353 (1977).
[10] A. I. Larkin and D. E. Khmel'nitskiĭ, Zh. Eksp. Theor. Fiz. **56**, 2087 (1969) [Sov. Phys. JETP **29**, 1123 (1969)].
[11] D. R. Nelson and M. E. Fisher, Phys. Rev. B **11**, 1030 (1975).
[12] R. Cordery and A. Griffin, Ann. Phys. (N.Y.) **134**, 411 (1981).
[13] P. Pfeuty and G. Toulouse, *Introduction to the Renormalization Group and to Critical Phenomena* (Wiley, New York, 1977).
[14] T. C. Lubensky and M. H. Rubin, Phys. Rev. B **12**, 3885 (1975).
[15] A. J. Bray and M. A. Moore, J. Phys. A: Math. Gen. **10**, 1927 (1977).
[16] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1980), p. 917; M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1964), p. 629.
[17] The solutions in (2.18b)–(2.18d) are obtained from (2.18a) by making the replacements $u \to u + K$, $u \to u + iK'$, $U \to u + K + iK'$, respectively in (2.18a). Here K, K' are complete elliptic integrals (Chap. 8 of Gradshteyn and Ryzhik in Ref. 16).
[18] The elliptic functions sc, cd, and cs appearing in (2.21b), (2.21e), and (2.22b) are defined on p. 570 of Abramowitz and Stegun, Ref. 16.
[19] P. Kumar, Phys. Rev. B **10**, 2928 (1974).
[20] C. A. Wilson, J. Phys. C: Solid State **13**, 925 (1980).
[21] A. J. Bray and M. A. Moore, J. Phys. A: Math. Gen. **11**, 715 (1978).

Electrostatic structural transitions in a Yukawa–Wigner solid

George L. Hall
Department of Physics, North Carolina State University, Raleigh, North Carolina 27650

(Received 22 July 1981; accepted for publication 16 September 1981)

A derivation is supplied for a functional relation between the Fuchs energy ϵ and the Madelung energy S for a Yukawa–Wigner solid (YWS) in which the usual uniform background of a Wigner solid (WS) is replaced by a periodic array of Yukawa charge distributions with variable "ripple" parameter λ allowing the WS and the empty lattice in the limits of small λ and large λ, respectively. It is the zeros of $\Delta\epsilon$, and not of ΔS, that are relevant for structural transitions between two lattices. It is known that $2\epsilon^{\mathrm{WS}} = S^{\mathrm{WS}}$, and Medeiros and Mokross incorrectly assumed $2\epsilon = S$ for the YWS. Here it is first shown by elementary means that the relation between ϵ and S varies with λ, and then the functional relation is supplied for all λ. When applied to the bcc-fcc system, it is found that $\Delta\epsilon$ has two zeros whereas ΔS has one not equal to either of those of $\Delta\epsilon$. Starting with small λ, the sequence of lowest energy structures is bcc, fcc, and bcc if these are the only two allowed to compete. The equations for the sc case have not been evaluated, but it is expected that the full sequence for the cubics will be found to be bcc, fcc, and sc, as this author reported for the Gaussian–Wigner solid.

PACS numbers: 64.70.Kb, 61.50.Lt, 71.45.Nt

1. INTRODUCTION

This is the second of a two-part report on generalized Wigner[1] solids (WS's) in which the usual uniform background charge of a WS is replaced with a background possessing variable "ripple" in the charge. The WS has point charges Q (of either sign) located at the lattice sites of a Bravais lattice with neutrality being maintained with a uniform background. The first paper[2] treats the Gaussian WS (GWS) in which the background is formed by centering about each lattice site the charge distribution $-Q(p/\pi)^{3/2}\exp(-pr^2)$. In this paper the charge distribution $-Q(\lambda^2/4\pi)r^{-1}\exp(-\lambda r)$ is similarly centered to form the background of the Yukawa WS (YWS). As the ripple parameters p and λ go to zero the WS is regained, and as they go to infinity both models become empty lattices. Although these models are certainly classical and Coulombic, I shall continue to use the term classical Coulomb lattice (CCL) for the model composed of one lattice of point charges Q and another identical displaced lattice of point charges $-Q$, which gives one of the simplest models for ionic crystals. If in the GWS or YWS models one displaces the centering location of the Gaussian or Yukawa distributions, respectively, with respect to the lattice of point charges, one secures in the limit of large ripple parameter the CCL. The derivations I have provided[2] for the GWS and shall supply here for the YWS are easily transcribed for these "displaced" GWS and YWS models. Thus, this study of the GWS and YWS models provides the basis for a very wide class of Coulombic models from which one might select a model more suitable than the widely used WS. Birman[3] has used Gaussians in various ways to improve upon the CCL as a model for the ionic crystals and Medeiros and Mokross[4] have used the YWS to represent phase transitions in systems formed by polystrene particles in aqueous suspensions as observed by Williams *et al.*[5]

The purpose of this paper is to supply derivations of the Fuchs energy ϵ, the Madelung energy S, and functional relations between these two. For two competing lattices, it is the zeros of $\Delta\epsilon$, and not of ΔS, that determine when transitions occur as a function of ripple parameter. Recently, I reported[6] some of the final results for the GWS and YWS models, applied them to the cubic lattices, and pointed out that Medeiros and Mokross had correctly calculated S^{YWS} but had incorrectly assumed that $2\epsilon^{\mathrm{YWS}} = S^{\mathrm{YWS}}$. Presumably, they made this assumption in anology with the WS results $2\epsilon^{\mathrm{WS}} = S^{\mathrm{WS}}$, or perhaps in analogy with the non-Coulombic Lennard-Jones model. There are at least two ways one can see, without long derivations, that the relation between ϵ and S must vary with the ripple parameter.

First, the CCL models obeys $\epsilon^{\mathrm{CCL}} = S^{\mathrm{CCL}}$, which is suggestive. Second, in the empty lattice limit, for which the potential $\Phi(\mathbf{r})$ vanishes, both ϵ and S must diverge to negative infinity with leading terms that do not satisfy 2ϵ equal to S. A discussion of this limit for the GWS has been given in Ref. 2 and used to provide a stringent test of the final results. Here I shall also use this limit on the YWS to provide a test of the final results, but I wish first to use it to clarify the definitions of these two energies ϵ and S and to show that a functional relation between ϵ^{YWS} and S^{YWS} must vary with the ripple parameter λ.

As in Ref. 2, define $K(\mathbf{r})$ and $S(\mathbf{r})$ with $K = (\mathbf{O})$ and $S = S(\mathbf{O})$ by

$$\Phi(\mathbf{r}) = \langle\Phi\rangle + S(\mathbf{r})/Q + Q/r, \tag{1}$$

$$K(\mathbf{r}) = Q\Phi(\mathbf{r}) - Q^2/r = S(\mathbf{r}) + A, \quad A = Q\langle\Phi\rangle, \tag{2}$$

$$S = \lim_{r\to 0} S(\mathbf{r}) = \lim_{r\to 0} Q[\Phi(\mathbf{r}) - \langle\Phi\rangle - Q/r]. \tag{3}$$

Once the $\Phi(\mathbf{r})$ is defined, the $K(\mathbf{r})$ and $S(\mathbf{r})$ are defined. The $\Phi(\mathbf{r})$, $K(\mathbf{r})$, K, and A are multivalued via the arbitrary average potential $\langle\Phi\rangle$, but the $S(\mathbf{r})$, S, and ϵ are unique. Later I shall have to develop various expressions for $S(\mathbf{r})$ with $(\mathbf{r})$ not equal to zero or equal to any other vector of the lattice providing the sites for the point charges, because $S(\mathbf{r})$ is needed in the derivation of various expressions for ϵ for general λ. Howev-

 0022-2488/83/010209-06$02.50

er, the situation is much simpler for the empty lattice limit.

Until recently errors were made in applications of Ewald techniques which are equivalent to incorrectly equating K and S in the theory of S, and $K(\mathbf{r})$ and $S(\mathbf{r})$ in the theory of ϵ; Hall[7] uncovered the first error the Ihm and Cohen[8] uncovered the second for the WS case. The canceling effect[9] of these two errors for the WS and the related matter of "Evjen[10] oscillations" for the CCL model of CsCl has been discussed. From the form of Eq. (2) one sees why K has been viewed as the energy of interaction of one point charge Q with all *other* charge; it is a set of quantities corresponding to all possible finite choices of $\langle \Phi \rangle$ and contains the singlet S. The relation between $K(\mathbf{r})$ and $S(\mathbf{r})$, and hence between K and S, is not complicated by the arbitrariness of the average potential in the empty lattice limit, because $\Phi(\mathbf{r})$ vanishes in this limit and the set K reduces to the singlet S. Thus in the empty lattice limit S is given by the interaction of Q with the local Yukawa distribution which is very "bunched" up at the same site. An elementary physics calculation shows that the leading term in S is given by $-\lambda Q^2$ as λ approaches infinity.

The Fuchs energy ϵ is the interaction energy of all charge normalized to the volume Ω occupied by each point charge Q. As the local background charge near a point charge Q bunches up in the empty lattice limit to become a point charge $-Q$, the leading term in ϵ arises from two sources: the interaction of the charge Q with the local Yukawa distribution and the interaction of the Yukawa distribution with itself. The first of these is just S, and the second is readily shown to be given by $\lambda Q^2/4$. Thus we have

$$S^{\mathrm{YWS}} \sim -\lambda Q^2, \quad \epsilon^{\mathrm{YWS}} \sim -3\lambda Q^2/4, \quad 0 \ll \lambda. \tag{4}$$

The corresponding results for the GWS are given in Eqs. (16) and (38) of Ref. 2.

Thus it is seen that the Medeiros assumption that $2\epsilon = S$ for the YWS does not hold in general for the GWS or YWS, but it does hold in the WS limit of these two. Figure 1 shows that for the bcc-fcc system there are two zeros of $\Delta(2\epsilon)$ neither of which is equal to the single zero of ΔS.

Section 2 supplies derivations for S, Sec. 3 treats ϵ, and Sec. 4 gives further discussion.

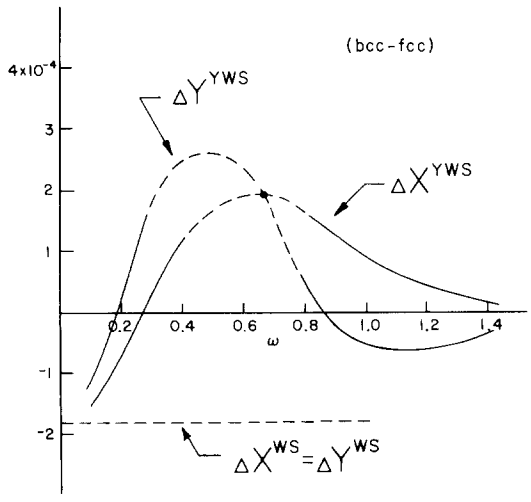


FIG. 1. Comparison of the (bcc-fcc) differences for the Madelung ($S = FX$) and twice the Fuchs ($2\epsilon = FY$) energies reduced by $F = Q^2/\Omega^{1/3}$ as a function of the "ripple" parameter $\omega = \lambda(\Omega^{1/3}/2\pi)$ for the regular (WS) and Yukawa–Wigner solids (YWS). Zeros of ΔY^{YWS} locate transitions. Middle portion (dashed) or curves is only schematic.

2. EXPRESSIONS FOR $K(\mathbf{r})$ AND $S(\mathbf{r})$ FOR THE YWS

Denote with $\{\tau\}$ a Bravais lattice with volume Ω per lattice point. Denote with $\{\gamma\}$ its reciprocal normalized by $\exp(i\gamma\cdot\tau) = 1$. The charge distribution for the YWS is given in the sense of tempered distributions[11] by

$$\rho(\mathbf{r}) = Q\sum_{\tau}\left[\delta(\mathbf{r}-\tau) - \frac{\lambda^2 \exp(-\lambda|\mathbf{r}-\tau|)}{4\pi|\mathbf{r}-\tau|}\right] \tag{5}$$

$$= \sum_{\gamma}{}' b_\gamma \exp(i\gamma\cdot\mathbf{r}), \quad b_\gamma = \frac{Q\gamma^2/\Omega}{(\gamma^2+\lambda^2)} \tag{6}$$

for $0 \leqslant \lambda < \infty$. Use units such that a point charge Q gives a potential Q/r and $\nabla^2\Phi(r) = -4\pi\rho(r)$. Through Eq. (2) one form for $S(\mathbf{r})$ is found from

$$\Phi(\mathbf{r}) = \langle\Phi\rangle - \frac{4\pi Q}{\Omega\lambda^2} + Q\sum_{\tau}\frac{\exp(-\lambda|\mathbf{r}-\tau|)}{|\mathbf{r}-\tau|}, \quad 0<\lambda. \tag{7}$$

This shows that if one takes $\Phi(\mathbf{r})$ to be given by the last term only, one is tacitly assuming that the average potential is given by the second term, which is equivalent to defining $\langle\Phi\rangle$ by the limit as γ goes to zero of $(4\pi b_\gamma)/\gamma^2$. This corresponds to defining $\Phi(\mathbf{r})$ as the limit of the sequence of potentials associated with a finite nested sequence of summation cells[9] where each summation cell is a proximity cell centered about a lattice site and each cell possesses a charge distribution $Q[\delta(\mathbf{r}) - (\lambda^2/4\pi)r^{-1}\exp(-\lambda r)]$ centered about its lattice point *and extending outside the cell.* Equation (2) gives

$$S(\mathbf{r}) = Q^2\sum_{\tau}\frac{\exp(-\lambda|\mathbf{r}-\tau|)}{|\mathbf{r}-\tau|} - \frac{Q^2}{r} - \frac{4\pi Q^2}{\lambda^2\Omega}, \quad 0<\lambda \tag{8}$$

which is useful for large λ. A form useful for small λ is also needed. This could be eventually secured by using an integral transformation[12,4] on the summand of Eq. (8), but I shall derive other forms for $S(\mathbf{r})$ for the YWS by methods analogous to those used in Ref. 2 for the GWS. Then, to establish consistency, I shall show that the new expressions are equivalent to Eq. (8). This procedure has the advantage of showing how the limiting process defining $\Phi(\mathbf{r})$ and $K(\mathbf{r})$ gives rise to various expressions for $\langle\Phi\rangle$ and how the independence of $S(\mathbf{r})$ on $\langle\Phi\rangle$ arises in Ewald techniques.

It is only necessary that $\langle\Phi\rangle$ be treated equivalently everywhere. For definiteness, choose a summation cell[9] for defining $\Phi(\mathbf{r})$ to be the proximity cell and imagine a finite array of these centered on lattice sites. Let each cell contain the charge that an identical cell contains in the infinite charge array, so the charge associated with a summation cell does not extend outside the cell. The charge in each cell will have reflection symmetry (no dipole moment), and a finite array of them will possess a well-defined potential if a point charge Q is taken to contribute Q/r. The limit of an infinite nested sequence of these finite arrays defines the total potential $\Phi(\mathbf{r})$ such that $\langle\Phi\rangle$ is given by Eq. (5) of Ref. 9, which we do not need explicity here. With this limiting process understood on the τ summation, Eq. (2) gives

$$\left[\frac{K(\mathbf{r})}{Q^2}\right] = \lim_{N\to\infty} \sum_{\tau}^{N}{}' \left[\frac{1}{|\mathbf{r}-\tau|} - \int_{\tau} \frac{\lambda^2}{4\pi} \sum_{w}^{\infty} \frac{\exp(-\lambda|\mathbf{z}-\mathbf{w}|)}{|\mathbf{z}-\mathbf{r}||\mathbf{z}-\mathbf{w}|}\, d^3z\right]$$
$$- \int_0 \frac{\lambda^2}{4\pi} \sum_{w}^{\infty} \frac{\exp(-\lambda|\mathbf{r}-\mathbf{w}|)}{|\mathbf{z}-\mathbf{r}||\mathbf{z}-\mathbf{w}|}\, d^3z \tag{9}$$
$$= \lim_{N\to\infty} \sum_{\tau}^{N}{}' \left[\frac{1}{|\mathbf{r}-\tau|} - \frac{\lambda^2}{\Omega} \int_{\tau} \sum_{\gamma}^{\infty} \frac{\exp(i\gamma\cdot\mathbf{z})\, d^3z}{(\gamma^2+\lambda^2)|\mathbf{z}-\mathbf{r}|}\right]$$
$$- \frac{\lambda^2}{\Omega} \int_0 \sum_{\gamma}^{\infty} \frac{\exp(i\gamma\cdot\mathbf{z})}{(\gamma^2+\lambda^2)|\mathbf{z}-\mathbf{r}|}\, d^3z \tag{10}$$
$$= \lim_{N\to\infty} \sum_{\tau}^{N}{}' \int_0^{\infty} \frac{dv}{(\pi v)^{1/2}} \{\exp[-v(\mathbf{r}-\tau)^2]$$
$$- \frac{\lambda^2}{\Omega} \int_{\tau} \sum_{\gamma}^{\infty} \frac{\exp[i\gamma\cdot\mathbf{z} - v(\mathbf{z}-\mathbf{r})^2]}{\gamma^2+\lambda^2}\, d^3z\}$$
$$- \frac{\lambda^2}{\Omega} \int_0 \sum_{\gamma}^{\infty} \frac{\exp[i\gamma\cdot\mathbf{z} - v(\mathbf{z}-\mathbf{r})]}{\gamma^2+\lambda^2}\, d^3z, \tag{11}$$

where the Poisson summation formula (PSF)[13] is used to secure Eq. (10) and the lemma of Eq. (A3) of Ref. 7 is used to get Eq. (11). Following the analysis of Ref. 7, passing the τ-summation inside the integral in Eq. (11) brings in the average potential such that $K(\mathbf{r}) = S(\mathbf{r}) + A$ where

$$\frac{S(\mathbf{r})}{Q^2} = \int_0^{\infty} (\pi v)^{-1/2} \left\{ \sum_{\tau}^{\infty}{}' \exp[-v(\tau-\mathbf{r})^2 \right.$$
$$\left. - \frac{\lambda^2}{\Omega} \left(\frac{\pi}{v}\right)^{3/2} \sum_{\gamma}^{\infty} \frac{\exp[i\gamma\cdot\mathbf{r} - (\gamma^2/4v)]}{\gamma^2+\lambda^2} \right\} dv \tag{12}$$
$$= \frac{S^{\mathrm{WS}}(\mathbf{r})}{Q^2} - \frac{\lambda^2}{\Omega}$$
$$\times \int_0^{\infty} (\pi v)^{-1/2} \left(\frac{\pi}{v}\right)^{3/2} \sum_{\gamma}^{\infty}{}' \frac{\exp[i\gamma\cdot\mathbf{r} - (\gamma^2/4v)]}{\gamma^2+\lambda^2}\, dv \tag{13}$$
$$= \frac{S^{\mathrm{WS}}(\mathbf{r})}{Q^2} - \frac{4\pi\lambda^2}{\Omega} \int_0^{\infty} \sum_{\gamma}^{\infty}{}' \frac{\exp(i\gamma\cdot\mathbf{r} - s\gamma^2)}{\gamma^2+\lambda^2}\, ds, \tag{14}$$

where $S^{\mathrm{WS}}(\mathbf{r})$ denotes the WS limit of $\lambda = 0$. For convenience of reference in the sequel various expressions for $S^{\mathrm{WS}}(\mathbf{r})$ are given in Appendix A. In Appendix B it is shown that

$$\frac{S(\mathbf{r})}{Q^2} = \frac{S^{\mathrm{WS}}(\mathbf{r})}{Q^2} - \frac{4\pi\lambda^2}{\Omega} \sum_{\gamma}^{\infty}{}' \frac{\exp(i\gamma\cdot\mathbf{r})}{\gamma^2(\gamma^2+\lambda^2)} \tag{15}$$
$$= \frac{S^{\mathrm{WS}}(\mathbf{r})}{Q^2} - \frac{4\pi}{\Omega}$$
$$\times \int_0^{\infty} [1 - \exp(-S\lambda^2)] \sum_{\gamma}^{\infty}{}' \exp(i\gamma\cdot\mathbf{r} - s\gamma^2)\, ds. \tag{16}$$

Note that the integrand of Eq. (16) is not equal to that of Eq. (14). The form of $S(\mathbf{r})$ given by Eq. (15) is especially useful in deriving expressions for ϵ and for small λ gives the expression for $S = S(0)$

$$S = S^{\mathrm{WS}} - \frac{4\pi\lambda^2 Q^2}{\Omega} \sum_{\gamma}^{\infty}{}' \frac{1}{\gamma^2(\gamma^2+\lambda^2)}. \tag{17}$$

Next I make connections with Eq. (8) using Eq. (16) and Eq. (A4) to write

$$S(\mathbf{r}) = -\frac{Q^2}{r} + \frac{4\pi Q^2}{\Omega}$$
$$\times \int_0^{\infty} \exp(-s\lambda^2) \sum_{\gamma}^{\infty}{}' \exp(i\gamma\cdot\mathbf{r} - s\gamma^2)\, ds \tag{18}$$
$$= -\frac{Q^2}{r} - \frac{4\pi Q^2}{\Omega\lambda^2} + \frac{4\pi Q^2}{\Omega}$$
$$\times \int_0^{\infty} \exp(-s\lambda^2) \sum_{\gamma}^{\infty} \exp(i\gamma\cdot\mathbf{r} - s\gamma^2)\, ds. \tag{19}$$

Setting $s = 1/4v$ and applying the PSF gives

$$S(\mathbf{r}) = -\frac{Q^2}{r} - \frac{4\pi Q^2}{\Omega\lambda^2} + Q^2$$
$$\times \int_0^{\infty} (\pi v)^{-1/2} \exp(-\lambda^2/4v) \sum_{\tau}^{\infty} \exp[-v(\mathbf{r}-\tau)^2]\, dv. \tag{20}$$

For positive λ the order of summation and integration in this equation can be interchanged giving, with the aid of the transformation[12] mentioned earlier, Eq. (8).

Equations (13)–(17) together with Appendix A provide checks of the correctness of the derivations in that $S(\mathbf{r})$ reduces to $S^{\mathrm{WS}}(\mathbf{r})$ as λ goes to zero. For a check in the empty lattice limit note that Eq. (8) gives

$$S = \lambda Q^2 - \frac{4\pi Q^2}{\Omega\lambda^2} + Q^2 \sum_{\tau}^{\infty}{}' \frac{\exp(-\lambda\tau)}{\tau}, \tag{21}$$

from which follows the first expression in Eq. (4); instead of using Eq. (8), one could also use Eq. (20) to secure the first of Eq. (4).

Properties of ΔS for two lattices

In order to calculate ΔS for two Bravais lattices, it is convenient to use Eq. (21) for very large λ and Eq. (17) for very small λ. These two suffice to locate the zeros of ΔS (and later $\Delta\epsilon$) for the bcc-fcc system as shown in Ref. 6 and the present Fig. 1. The mutual reciprocity of the bcc and fcc lattices can be exploited by defining mutually-reciprocal, unit lattices (a,b) by $\tau = \Omega^{1/3}\mathbf{a}$, $\gamma = 2\pi\mathbf{b}/\Omega^{1/3}$, and $\exp(2\pi i\mathbf{a}\cdot\mathbf{b}) = 1$. Define $S = FX$, $F = Q^2/\Omega^{1/3}$, $\omega = \lambda(\Omega^{1/3}/2\pi)$, and $\Delta X = X(a) - X(b)$. Then Eqs. (21) and (17) give, respectively,

$$\Delta X = \sum_{a}{}' \frac{\exp(-2\pi\omega a)}{a} - \sum_{b}{}' \frac{\exp(-2\pi\omega b)}{b} \tag{22}$$
$$= \Delta X^{\mathrm{WS}} + M, \tag{23}$$

where

$$M = \left(\frac{\omega}{\pi}\right)^2 \left\{ \sum_{a}{}' \frac{1}{a^2(a^2+\omega^2)} - \sum_{b}{}' \frac{1}{b^2(b^2+\omega^2)} \right\}. \tag{24}$$

Equation (24) can be expanded[6] in a Taylor's series in ω^2 with the coefficients involving Lennard-Jones[14] sums. This was done as described in Ref. 6 to compute values of ΔS for small ω as shown here in Fig. 1. For systems other than the bcc-fcc it may be necessary to calculate ΔX at intermediate values of ω, and then it would be necessary to use the theta function method as employed by Medeiros and Mokross.[4] From work by Foldy[15] the X^{WS} for the cubic lattices have been calculated to ten significant figures.

Let us next consider comparing the S's for two Bravais lattices that are not reciprocally related. Denote the associated unit lattices with $\mathbf{a}$ and $\mathbf{a}'$, chosen such that $\Delta X = X(a) - X(a')$ yields ΔX^{WS} as negative. In Eq. (22) replace $\mathbf{b}$ and $\mathbf{a}'$, and in Eq. (24) replace $\mathbf{a}$ by $\mathbf{a}'$. One can still use Eq. (24), so modified, with a Taylor's expansion to sketch ΔX

for small ω. Using just first-neighbor contributions in the modified Eq. (22), one can quickly determine whether ΔX approaches zero from above or below as ω approaches infinity. In the latter case ΔS probably does not possess a zero for finite ω; this was found to be the case for the fcc-sc and bcc-sc systems in the GWS model,[2,6] where SC denotes simple cubic.

3. THE FUCHS ENERGY

The Fuchs energy ϵ is independent of the particular summation cell[9] one uses except that the same summation cell must be used everywhere. I shall use the same summation cell that I used in Eq. (9), and my derivation parallels that given in Ref. 2 for the GWS which in turn parallels that given in Ref. 9 for the WS. The definition of the Fuchs energy is

$$\frac{2\epsilon}{Q^2} = \lim_{N\to\infty}\frac{1}{N}\left\{\sum_{\tau}^{N}\sum_{\tau'\neq\tau}^{N}\frac{1}{\Delta\tau} + \sum_{\tau}^{N}\sum_{\tau'}^{N}\frac{\lambda^4}{(4\pi)^2}\int_0\int\sum_{w}^{\infty}\sum_{w'}^{\infty}\frac{\exp[-\lambda(|\mathbf{z}-\mathbf{w}|+|\mathbf{z}'-\mathbf{w}'|)]}{|\Delta\tau+\Delta\mathbf{z}||\mathbf{z}-\mathbf{w}||\mathbf{z}'-\mathbf{w}'|}\,d^3z\,d^3z'\right.$$
$$\left. - 2\sum_{\tau}^{N}\sum_{\tau'}^{N}\frac{\lambda^2}{4\pi}\int_0\sum_{w}^{\infty}\frac{\exp(-\lambda|\mathbf{z}-\mathbf{w}|)}{|\Delta\tau+\mathbf{z}||\mathbf{z}-\mathbf{w}|}\,d^3z\right\} \tag{25}$$

$$= \lim_{N\to\infty}\frac{1}{N}\left[\sum_{\tau}^{N}\sum_{\tau'\neq\tau}^{N}\frac{1}{\Delta\tau} + \sum_{\tau}^{N}\sum_{\tau'}^{N}\frac{\lambda^4}{\Omega^2}\int_0\int\sum_{\gamma}^{\infty}\sum_{\gamma'}^{\infty}\frac{\exp(i\gamma\cdot\mathbf{z}+i\gamma'\cdot\mathbf{z}')}{(\gamma^2+\lambda^2)(\gamma'^2+\lambda^2)|\Delta\tau+\Delta\mathbf{z}|}\,d^3z\,d^3z'\right.$$
$$\left. - 2\sum_{\tau}^{N}\sum_{\tau'}^{N}\frac{\lambda^2}{\Omega}\int_0\sum_{\gamma}^{\infty}\frac{\exp(i\gamma\cdot\mathbf{z})}{(\gamma^2+\lambda^2)|\Delta\tau+\mathbf{z}|}\right], \tag{26}$$

where $\mathbf{w}$ and $\mathbf{w}'$ are τ lattices and PSF has been used to secure Eq. (26) from (25). Setting λ equal to zero regains Eq. (6) of Ref. 9 for the WS. Next group together the terms given by $\tau=\tau'$ in Eq. (26) to give

$$\frac{2\epsilon}{Q^2} = \frac{\lambda^4}{\Omega^2}\int_0\int\sum_{\gamma}^{\infty}\sum_{\gamma'}^{\infty}\frac{\exp(i\gamma\cdot\mathbf{z}+i\gamma'\cdot\mathbf{z}')}{\Delta z(\gamma^2+\lambda^2)(\gamma'^2+\lambda^2)}\,d^3z\,d^3z' - \frac{2\lambda^2}{\Omega}\int_0\sum_{\gamma}^{\infty}\frac{\exp(i\gamma\cdot\mathbf{z})}{z(\gamma^2+\lambda^2)}\,d^3z$$
$$+ \lim_{N\to\infty}\frac{1}{N}\sum_{\tau}^{N}\sum_{\tau'\neq\tau}^{N}\frac{1}{\Delta\tau}\,(\text{remaining terms}). \tag{27}$$

In the next step I shall change the summation limit on the τ' summation to infinity. The proof of its validity is exactly the same as given for the GWS in its Appendix and will not be repeated here. With this step justified, one can with a change to summation over $\Delta\tau$ write the last term in Eq. (27) as

$$\sum_{\tau}{}'\left[\frac{1}{\tau} - \frac{2\lambda^2}{\Omega}\int_0\sum_{\gamma}^{\infty}\frac{\exp i\gamma\cdot\mathbf{z})}{(\gamma^2+\lambda^2)|\tau+z|}\,d^3z\right.$$
$$\left. + \frac{\lambda^4}{\Omega^2}\int_0\int\sum_{\gamma}^{\infty}\sum_{\gamma'}^{\infty}\frac{\exp(i\gamma\cdot\mathbf{z}+i\gamma'\cdot\mathbf{z})}{(\gamma^2+\lambda^2)(\gamma'^2+\lambda^2)|\tau+\Delta\tau|}\,d^3z\,d^3z'\right]. \tag{28}$$

Putting this into Eq. (27) and rearranging the terms gives

$$2\epsilon = K - \frac{\lambda^2}{\Omega}\int_0\sum_{\gamma}^{\infty}\frac{K(z)\exp(i\gamma\cdot\mathbf{z})}{\gamma^2+\lambda^2}\,d^3z$$
$$- Q^2\frac{\lambda^2}{\Omega}\int_0\sum_{\gamma}^{\infty}\frac{\exp(i\gamma\cdot\mathbf{z})}{z(\gamma^2+\lambda^2)}d^3z. \tag{29}$$

This may be simplified, and ϵ shown to be independent of the average potential, by substituting $K=S+A$ and $K(\mathbf{r})=S(\mathbf{r})+A$, which leads to cancellation of the two terms containing A. Thus Eq. (29) holds with K replaced by S and $K(\mathbf{r})$ replaced by $S(\mathbf{r})$, giving

$$2\epsilon = S - \sum_{\gamma}{}'\frac{\lambda^2}{(\gamma^2+\lambda^2)}\left[\frac{1}{\Omega}\int_0 S(\mathbf{z})\exp(i\gamma\cdot\mathbf{z})\,d^3z\right]$$
$$- Q^2\sum_{\gamma}{}'\frac{\lambda^2}{(\gamma^2+\lambda^2)}\left[\frac{1}{\Omega}\int_0\frac{\exp(i\gamma\cdot\mathbf{z})}{z}\,d^3z\right]. \tag{30}$$

The second term in Eq. (30) can be simplified by using Eq. (15) to give

$$\frac{1}{\Omega}\int_0 S(\mathbf{z})\exp(i\gamma\cdot\mathbf{z})\,d^3z$$
$$= \frac{1}{\Omega}\int_0 S^{\mathrm{ws}}(z)\exp(i\gamma\cdot\mathbf{z})\,dz - \frac{4\pi Q^2\lambda^2}{\Omega\gamma^2(\gamma^2+\lambda^2)}, \quad \gamma\neq 0$$
$$= \frac{1}{\Omega}\int_0 S^{\mathrm{ws}}(z)\,dz, \quad \gamma=0. \tag{31}$$

Note that the integral of $S^{\mathrm{ws}}(z)$ over the centered summation cell is equal to the negative of the integral of z^{-1} over the cell, substitute this into Eq. (30), and find that

$$2\epsilon = S + \frac{4\pi Q^2}{\Omega}\sum_{\gamma}{}'\frac{\lambda^4}{\gamma^2(\gamma^2+\lambda^2)}$$
$$- \sum_{\gamma}{}'\frac{\lambda^2}{(\gamma^2+\lambda^2)}\left[\frac{1}{\Omega}\int_0 S^{\mathrm{ws}}(z)\exp(i\gamma\cdot\mathbf{z})\,d^3z\right]$$
$$- Q^2\sum_{\gamma}{}'\frac{\lambda^2}{(\gamma^2+\lambda^2)}\left[\frac{1}{\Omega}\int_0\frac{\exp(i\gamma\cdot\mathbf{z})}{z}\,d^3z\right]. \tag{32}$$

To simplify the last two terms, set λ equal to zero in Eq. (18) to give

$$S^{\mathrm{ws}}(\mathbf{r}) = \frac{4\pi Q^2}{\Omega}\int_0^{\infty}\sum_{\gamma}{}'\exp(s\gamma^2+i\gamma\cdot\mathbf{r})\,ds - \frac{Q^2}{r}, \tag{33}$$

and substitute this into Eq. (32) to secure

$$2\epsilon = S + \frac{4\pi Q^2}{\Omega}\sum_{\gamma}{}'\frac{1}{\gamma^2}\left[\left(\frac{\lambda^2}{\gamma^2+\lambda^2}\right)^2 - \left(\frac{\lambda^2}{\gamma^2+\lambda^2}\right)\right] \tag{34}$$

$$= S^{ws} - 2\left(\frac{4\pi Q^2}{\Omega}\sum_{\gamma}{}' \frac{\lambda^2}{\gamma^2(\gamma^2+\lambda^2)}\right) + \frac{4\pi Q^2}{\Omega}\sum_{\gamma}{}' \frac{\lambda^4}{\gamma^2(\gamma^2+\lambda^2)^2}, \quad (35)$$

where Eq. (17) is used to get Eq. (35), which is useful for small λ.

One can find directly from Eq. (35) an alternate form useful for large λ,but I shall first give a functional relation between ϵ and S and then use it to secure the alternate expression for ϵ from Eq. (21) for S.

4. THE FUNCTIONAL RELATION AND DISCUSSION

Define $2\epsilon = FY$ and use the notation of Eq. (22) to write

$$Y = X - \sum_{b}{}' \frac{\omega^2}{(\pi b^2)(b^2+\omega^2)} + \sum_{b}{}' \frac{\omega^4}{(\pi b^2)(b^2+\omega^2)^2}. \quad (36)$$

Also substitute into Eq. (17) rewritten as

$$X = X^{ws} - \sum_{b}{}' \frac{\omega^2}{(\pi b^2)(b^2+\omega^2)}, \quad (37)$$

which yields

$$Y = X^{ws} - 2\sum_{b}{}' \frac{\omega^2}{(\pi b^2)(b^2+\omega^2)} + \sum_{b}{}' \frac{\omega^4}{(\pi b^2)(b^2+\omega^2)^2}. \quad (38)$$

Now it is evident that Y and X obey the functional equation

$$Y = X + \frac{\omega}{2}\frac{dX}{d\omega}. \quad (39)$$

If we apply this functional equation to Eq. (21) for S rewritten in terms of X as

$$X = -2\pi\omega - \frac{1}{\pi\omega^2} + \sum_{a}{}' \frac{\exp(-2\pi\omega a)}{a}, \quad (40)$$

we find

$$Y - 3\pi\omega + \sum_{a}{}' \frac{\exp(-2\pi\omega a)}{a} - \pi\omega\sum_{a}{}'\exp(-2\pi\omega a). \quad (41)$$

An immediate check on the accuracy of Eq. (39) and (41) is given by the empty lattice limit for which

$$Y \sim -3\pi\omega, \quad \text{or } \epsilon \sim -3\lambda Q^2/4, \quad (42)$$

in agreement with Eq. (4) found by elementary means.

Additional expressions for Y follow from setting r equal to zero in Eqs. (12)–(16) and (18)–(20) and applying Eq. (39) also holds with X replaced by ΔX and Y replaced by ΔY, where the Δ refers to the difference of a quantity evaluated on two lattices.

For the bcc-fcc system, Fig. 1 shows how ΔY varies with ω. Note Y has two zeros whereas ΔX has only one. Medeiros and Mokross[4] correctly found the single zero of ΔX but incorrectly assumed that $\Delta Y = \Delta X$. I used Eqs. (37) and (38) for small ω and Eqs. (40) and (41) for large ω; it was not necessary in finding the zeros to calculate the curves in the intermediate domain indicated schematically by the dashed lines.

For other systems the zeros may fall in the intermediate domain, and then it may be necessary to use the theta function method (TFM), which for the YWS is slightly more complicated than for the WS. The TFM as used to evaluate $S^{ws}(r)$ and S^{ws} is described briefly at the end of Appendix A; one way to use the TFM to evaluate S (hence X or ΔX) is given by Medeiros and Mokross, which one could also use on Y or ΔY. For ΔS, their procedure amounts to working with Eqs. (18) and (20). I should add that at small ω it may be advantageous to separate out ΔS^{ws} in both equations, i.e., work with Eq. (16) instead of Eq. (18).

From Eq. (39) it is seen that ΔY equals ΔX at the extrema of ΔX, and

$$\frac{d\Delta Y}{d\omega} = \frac{\omega}{2}\frac{d^2\Delta X}{d\omega^2}, \text{ at extrema of } \Delta X. \quad (43)$$

Thus at the single maxima of of ΔX for the bcc-fcc system in Fig. 1 one has that ΔY equals ΔX and ΔY has a negative slope there. Also at the extrema of ΔY one has

$$3\frac{d\Delta X}{d\omega} = -\omega\frac{d^2\Delta X}{d\omega^2}, \text{ at extrema of } \Delta Y. \quad (44)$$

Finally note that although the Y's and X's are tightly coupled for the GWS and YWS models through functional equations, the functional equations differ.

APPENDIX A

This appendix contains various expressions for $S^{ws}(\mathbf{r})$ from which $S^{ws}(\mathbf{0}) = S^{ws}$ may be evaluated.

Take Eq. (12) and set λ equal to zero to write

$$\frac{S^{ws}(\mathbf{r})}{Q^2} = \int_0^\infty (\pi v)^{-1/2} \times\left\{\sum_{\tau}{}'\exp[-v(\gamma-\mathbf{r})^2] - \frac{1}{\Omega}\left(\frac{\pi}{v}\right)^{3/2}\right\} dv \quad (A1)$$

$$= \int_0^\infty (\pi v)^{-1/2} \times\left\{\frac{1}{\Omega}\left(\frac{\pi}{v}\right)^{3/2}\sum_{\gamma}{}'\exp\left(i\gamma\cdot\mathbf{r} - \frac{\gamma^2}{4v}\right) - \exp(-vr^2)\right\} dv, \quad (A2)$$

where Eq. (A2) follows by the PSF.[13] Next set $v = 1/4s$, which gives

$$S^{ws}(\mathbf{r}) = \frac{4\pi Q^2}{\Omega} \times\int_0^\infty \left[\sum_{\gamma}{}'\exp(i\gamma\cdot\mathbf{r} - s\gamma^2) - \frac{\Omega\exp(-r^2/4s)}{(4\pi s)^{3/2}}\right] ds. \quad (A3)$$

If $\mathbf{r}\notin\{\tau\}$ one can also write

$$S^{ws}(\mathbf{r}) = -\frac{Q^2}{r} + \frac{4\pi Q^2}{\Omega}\int_0^\infty \sum_{\gamma}{}'\exp(i\gamma\cdot\mathbf{r} - s\gamma^2)\, ds, \quad (A4)$$

but one cannot interchange the summation and integration in Eq. (A4), because that would give a conditionally convergent expression. To obtain S^{ws} just set $\mathbf{r} = \mathbf{0}$ in Eqs. (A1)–(A3).

To evaluate $S^{ws}(\mathbf{r})$ and S^{ws} use the integrand of Eq. (A1) over the domain $0 < d \leqslant v < \infty$ and the integrand of Eq.

(A3) over the domain $(1/4d)\leqslant s<\infty$. Then in each integral interchange the order of summation and integration. The sum of the two parts can be readily evaluated with a few terms from each part provided d is chosen to "balance" the contributions from each part. A fairly good choice is given by $d=\pi\Omega^{2/3}$.

APPENDIX B

The purpose of this appendix is to prove Eqs. (15) and (16) follow from Eq. (14), i.e., to prove for positive λ that the defining expression

$$I=\int_0^\infty \sum_\gamma{}' \frac{\lambda^2 \exp(i\gamma\cdot\mathbf{r}-s\gamma^2)}{\gamma^2+\lambda^2} \tag{B1}$$

can be written also in the two ways

$$I=\sum_\gamma{}' \frac{\lambda^2 \exp(i\gamma\cdot\mathbf{r})}{\gamma^2(\gamma^2+\lambda^2)}, \tag{B2}$$

$$I=\int_0^\infty \sum_\gamma{}' [1-\exp(s\lambda^2)]\exp(i\gamma\cdot\mathbf{r}-s\gamma^2)\,ds. \tag{B3}$$

Note that the integrands of Eqs. (B1) and (B3) are not equal.

If one can establish that interchange of the order of summation and integration is legitimate in Eqs. (B1) and (B3), the proof is immediate; then Eq. (B2) follows directly from Eq. (B1), and with the aid of partial fractions Eq. (B3) follows from Eq. (B2). This device is, if justified, much simpler than applying the PSF to Eq. (B1).

Divide the domain of integration into $0\leqslant s\leqslant\delta$ and $\delta\leqslant s<\infty$. The interchange is clearly justified for the second domain. Thus it is necessary to prove that in the limit of δ approaching zero the following two integrals vanish:

$$\int_0^\delta \sum_\gamma{}' \frac{\lambda^2 \exp(i\gamma\cdot\mathbf{r}-s\gamma^2)}{\gamma^2+\lambda^2}\,ds, \tag{B4}$$

$$\int_0^\delta \sum_\gamma{}' [1-\exp(-s\lambda^2)\exp(i\gamma\cdot\mathbf{r}-s\gamma^2)]\,ds. \tag{B5}$$

It suffices to treat the notationally simpler case of $r=0$ and prove that in this limit J and L vanish where

$$J=\int_0^\delta \sum_\gamma{}' \frac{\lambda^2}{\gamma^2+\lambda^2}\exp(-s\gamma^2)\,ds, \tag{B6}$$

$$L=\int_0^\delta \sum_\gamma{}' [1-\exp(-s\lambda^2)]\exp(-s\gamma^2)\,ds. \tag{B7}$$

Consider J and set $s=\delta x$ giving

$$J=\delta\int_0^1 \left[-1+\sum_\gamma{}' \frac{\lambda^2}{\gamma^2+\lambda^2}\exp(-\delta x\gamma^2)\right] dx, \tag{B8}$$

$$\lim_{\delta\to 0} J=\lim_{\delta\to 0}\delta\int_0^1 \sum_\gamma{}' \frac{\lambda^2}{\gamma^2+\lambda^2}\exp(-\delta x\gamma^2)\,dx. \tag{B9}$$

Instead of deriving the PSF expression for the sum inside the integral of Eq. (B9), it is only necessary to recall that the "term at the origin" in a PSF expression is given by the integral approximation to the sum. Thus for the present purposes replace the sum by

$$\frac{1}{v}\int\int\int^\infty \frac{\exp(-\delta x u^2)}{u^2+\lambda^2}\,d^3u, \tag{B10}$$

where $v=8\pi^3/\Omega$. Then it follows easily that

$$\lim_{\delta\to 0} J=\lim_{\delta\to 0}\frac{4\pi\lambda^2\sqrt{\delta}}{v}=0. \tag{B11}$$

Next consider L and again set $s=\delta x$ giving

$$L=\delta\int_0^1 [1-\exp(-\delta\lambda^2 x)]\sum_\gamma{}' \exp(-\delta\gamma^2 x)\,dx. \tag{B12}$$

Since λ is positive but otherwise arbitrary, fix it and then choose $\delta\lambda^2\ll 1$. Then

$$\lim_{\delta\to 0} L=\lim_{\delta\to 0}\lambda^2\delta^2\int_0^1 x\sum_\gamma{}' \exp(-\delta\gamma^2 x)\,dx \tag{B13}$$

$$=\lim_{\delta\to 0}\lambda^2\delta^2\int_0^1 \frac{x}{v}\left(\frac{\pi}{\delta x}\right)^{3/2} dx=0, \tag{B14}$$

when the first term in the PSF expression for the sum in Eq. (B13) has been used in Eq. (15), i.e., the integral approximation to the sum.

This completes the proof.

[1] E. P. Wigner, Phys. Rev. **46**, 1002 (1934); Trans. Faraday Soc. **34**, 678 (1936).
[2] G. L. Hall, unpublished.
[3] J. Birman, Phys. Rev. **97**, 897 (1955); J. Phys. Chem. Solids **6**, 65 (1958).
[4] J. Medeiros e Silva and B. J. Mokross, Phys. Rev. B **21**, 2972 (1980).
[5] R. Williams, R. S. Crandall, and P. Wojtowicz, Phys. Rev. Lett. **37**, 348 (1976).
[6] G. L. Hall, Phys. Rev. B **24**, 2881 (1981).
[7] G. L. Hall, Phys. Rev. B **19**, 3921 (1979).
[8] J. Ihm and M. L. Cohen, Phys. Rev. B **21**, 3754 (1980).
[9] G. L. Hall and T. R. Rice, Phys. Rev. B **21**, 3757 (1980).
[10] H. M. Evjen, Phys. Rev. **39**, 675 (1932). Also see S. K. Roy, Can. J. Phys. **32**, 509 (1934).
[11] H. Bremermann, *Distributions, Complex Variables, and Fourier Transforms* (Addison-Wesley, Reading, Mass., 1965), p. 86.
[12] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965), p. 307, No. 3.325.
[13] E. M. Stein and G. Weiss, *Fourier Analysis on Euclidean Spaces* (Princeton U. P., Princeton, N.J., 1971), p. 253.
[14] J. O. Hirschfelder, C. F. Curtis, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954), p. 1038.
[15] L. L. Foldy, Phys. Rev. B **17**, 4889 (1978).

Comments on the static spherically symmetric cosmologies of Ellis, Maartens, and Nel

C. B. Collins
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

(Received 9 November 1981; accepted for publication 6 August 1982)

Ellis, Maartens, and Nel have discussed the viability of static spherically symmetric (SSS) cosmologies in general relativity, and in doing so they have studied some of the mathematical aspects of the field equations in that situation. We investigate further these mathematical aspects. Since the field equations correspond to those studied for stellar models, this question is related to previous investigations in that context. In particular, it is shown that conditions at the center of symmetry do not always uniquely determine the space-time geometry; this has relevance to the numerical investigation of stellar systems. Finally, in view of the need to generalize SSS models, some remarks are made on the possibility of relaxing the staticity condition in the case of models that are shear-free.

PACS numbers: 98.80.Dr

1. INTRODUCTION

In a recent thought-provoking article, Ellis, Maartens, and Nel[1] consider the consequence of adopting an unconventional interpretation of the observed isotropy of the galactic redshifts. Can the observations, they wonder, be explained by means of simple static spherically symmetric models, without invoking any of the unverifiable philosophical arguments (such as the usual cosmological principles) which lead to spatially homogeneous cosmologies? They consider general relativistic static spherically symmetric models, and examine theoretical aspects of observations that could be performed, for various possible values of the somewhat repungnant cosmological constant, Λ. The line-element is of the form

$$ds^2 = -g^2(r)\,dt^2 + dr^2 + f^2(r)(d\theta^2 + \sin^2\theta\, d\phi^2),$$

and the source of the gravitational field is assumed to be a perfect fluid. Under these circumstances, the field equations of general relativity,[2] viz.,

$$G_{ij} + \Lambda g_{ij} = T_{ij},$$

become

$$\frac{f''}{f} + \frac{g''}{g} + \frac{f'g'}{fg} = p - \Lambda, \tag{1.1a}$$

$$\frac{2f''}{f} + \frac{f'^2}{f^2} - \frac{1}{f^2} = -\mu - \Lambda, \tag{1.1b}$$

and

$$\frac{2f'g'}{fg} + \frac{f'^2}{f^2} - \frac{1}{f^2} = p - \Lambda, \tag{1.1c}$$

where a prime (') denotes differentiation with respect to r; the fluid flowlines are tangent to the unit vector $u^i = g^{-1}\delta^i_0$, and μ and p are, respectively, the energy density and pressure of the fluid. Equations (1.1) are compatible whenever the Bianchi identities $G^{ij}{}_{;j} = 0$ are satisfied, i.e., when

$$(\mu + p)(g'/g) + p' = 0, \tag{1.2}$$

which is the radial equation of hydrostatic support.

Ellis, Maartens, and Nel[1] next consider the mathematical features of the differential equation system (1.1), observing in particular the invariance under the 2-parameter group of transformations

$$g \to \lambda g, \tag{1.3a}$$

$$r \to \kappa^{-1} r, \quad f \to \kappa^{-1} f, \quad \mu \to \kappa^2 \mu, \quad \Lambda \to \kappa^2 \Lambda, \quad p \to \kappa^2 p, \tag{1.3b}$$

where κ and λ are nonzero real numbers. For physical reasons discussed in their article, they suppose that the matter has an equation of state $p = \frac{1}{3}\mu$. The authors state, somewhat ambiguously, that, with this equation of state, it does not seem possible either to obtain general analytic solutions to Eqs. (1.1), or to obtain the qualitative behavior of the solutions by use of phase-plane methods "except in the cases $\Lambda = 0$; $\mu = 0$, or $\Lambda = \mu$". Accordingly, the system is examined numerically. Some ambiguity arises, first because, from this remark alone, it is not clear as to which possibility (analytic solution or qualitative behavior) the exceptional cases refer, and secondly because of the manner of labelling the three exceptional cases. The situation is somewhat clarified later on, when the authors consider the nature of the solutions as being dependent on two quantities, viz., Λ and μ_0, the value of the energy density along the central world-line $r = 0$ (for physical reasons, μ_0 is assumed to be nonnegative; in the mathematical analysis it is tacitly assumed that μ_0 is finite, and indeed when the physical applications are considered this is explicitly stated). First, three specific cases (the expectional ones) are examined. They are at this stage more precisely labelled as

(1) $\Lambda = \mu_0 = 0$. This is Minkowski space-time.

(2) $\Lambda \neq 0$, $\mu_0 = 0$. This is de Sitter space-time if $\Lambda > 0$, and anti-de Sitter space-time if $\Lambda < 0$ (see, e.g., Hawking and Ellis[3]: this corrects the statement made in Ref. 1).

(3) $\Lambda = \mu_0 > 0$. This is claimed to be the (generalized) Einstein static universe.

The authors state "*in the other cases, we have to rely on numerical integration.*" It now becomes clear that the authors' first manner of labelling ($\Lambda = 0$; $\mu = 0$, and $\Lambda = \mu$) was imprecise, that they regarded numerical integration as being

 0022-2488/83/010215-05$02.50

required in *all* other cases, and that they believed that cases (1)–(3) were the only ones capable of analysis *either* by exact solution *or* by phase-plane methods (or both). Furthermore, the authors' more precise labelling involves the quantity μ_0, and it is tacitly claimed that if $\mu_0 = 0$ then $\mu \equiv 0$, and if $\mu_0 = \Lambda$ then $\mu \equiv \Lambda$. The authors now resume their analysis of the system (1.1) by means of numerical methods, leading to an exhaustive classification of all possible choices of ordered pairs $(\mu_0, \Lambda; \mu_0 \geqslant 0)$, and invoking the additional cases

(4) $\Lambda > \mu_0 > 0$,

(5) $\mu_0 > \Lambda > 0$,

and

(6) $\Lambda \leqslant 0, \mu_0 \neq 0$.

From the physical point of view of applicability to the SSS cosmologies, Ellis, Maartens, and Nel require that the central energy density μ_0 be strictly positive, and they also require that initial values of the red-shift near the center be strictly positive [from which follows the inequality $\Lambda > \mu_0 > 0$, and so case (4) is the only candidate for application]. Nevertheless, Ellis, Maartens, and Nel take pains to study the properties of all six cases, because at that stage they are concerned (as I am here) with purely mathematical aspects of those cases.

In this article, I wish to make some further comments regarding the mathematical analysis of the system (1.1). This system is first reduced to a system of three first-order ordinary differential equations. Since $p = p(r)$ and $\mu = \mu(r)$, either μ is identically constant or there is an equation of state $p = p(\mu)$. The case where μ is identically constant includes the interior Schwarzschild solution, and is discussed elsewhere.[4] I shall be concerned chiefly with the case where $p = p(\mu)$, with the additional assumptions $\mu + p \neq 0$ and $dp/d\mu \neq 0$ [the former represents a reasonable energy condition; the latter is enforced to avoid the case where $g(r) \equiv$ const, since then the fluid flow is geodesic and the metric is Friedmann–Robertson–Walker,[5] and is in fact a generalized Einstein static model]. An exception to this, which will be discussed, involves generalizations of cases (1) and (2) above, in which, under special conditions, the constancy of μ and p is *proved* (rather than assumed). It is pointed out that if the equation of state is of the form $p = (\gamma - 1)\mu$ (a γ-law equation of state; γ is a constant, $\gamma \neq 0$, $\gamma \neq 1$), a qualitative (phase-plane) treatment *is* obtainable in the special case when $\Lambda = 0$, and that further exact solutions *are* known in that case. From the viewpoint of ultimate application to the SSS models alone, this special case is admittedly unphysical, and the exact solutions exhibited are inapplicable, since they possess only one center of symmetry, which is *irregular*. Nevertheless, in keeping with the purely mathematical aspects of the analysis of Ellis, Maartens, and Nel,[1] it is of interest to examine the $\Lambda = 0$ case [cases (1) and (6) of Ellis, Maartens, and Nel], and the role of the special exact solutions. Indeed, the results are also of interest to considerations outside the realm of SSS cosmologies, i.e., to the study of static stars. It is also possible to replace an irregular central region of the special solutions by matching appropriately to solutions that are regular at the center, and from this viewpoint the special solutions have more physical significance.

In Sec. 3, we consider how conditions at the center of symmetry ($r = 0$) determine the solution elsewhere (at $r > 0$). The tacit claim that if $p = \frac{1}{3}\mu$, $\mu_0 = 0$ implies $\mu \equiv 0$ is easily proved, but it is not at all clear how the condition $\mu_0 = \Lambda$ implies $\mu \equiv \Lambda$, in which case the solution is the Einstein static model, generalized to the equation of state $p = \frac{1}{3}\mu$. We do show that there are certain equations of state for which the appropriate central condition on the energy density does not uniquely generate the corresponding generalized Einstein static model, although the situation when $p = \frac{1}{3}\mu$ remains obscure. This question is not trivial, since the usual coordinate and tetrad bases are not defined at the center, and as a result the usual theorems relating to the uniqueness of solutions of ordinary differential equations do not apply.

Some further remarks, concerning the application of nonstatic spherically symmetric models to the observation-based cosmology of Ellis, Maartens, and Nel, are provided in Sec. 4.

2. THE SYSTEM OF EQUATIONS (1.1)

The question of static spherically symmetric geometries in general relativity is often considered in the context of stellar, rather than cosmological, situations. It is convenient to use the function $f(r)$ as an alternative radial coordinate. Before doing so, however, we briefly discuss the case when this is not possible, i.e., when $f' \equiv 0$. We first note that if $\mu + p \equiv 0$, it follows from the field equations that the space-time is an Einstein space, and hence[3] is either Minkowski, de Sitter, or anti-de Sitter space-time. In particular, if $f' \equiv 0$, we obtain from (1.1b) and (1.1c) that $\mu + p \equiv 0$, and de Sitter space-time results. Henceforth we shall assume $\mu + p \not\equiv 0$ so, *a fortiori*, $f' \not\equiv 0$. Following Kramer *et al.*,[4] but including the cosmological constant, we have from (1.1b),

$$\left(\frac{df}{dr}\right)^2 = 1 - \frac{2m(f)}{f} \quad \text{where} \quad \frac{dm}{df} = \frac{1}{2}(\mu + \Lambda) f^2 \tag{2.1a}$$

and, by (1.1c) and (1.2),

$$2f(f - 2m)\frac{dp}{dr} = -(\mu + p)(2m + (p - \Lambda)f^3). \tag{2.1b}$$

Let $M = m/f$, $D = \frac{1}{3}\mu f^2$, $P = \frac{1}{3}pf^2$ and $\lambda = \frac{1}{3}\Lambda f^2$. Then

$$\frac{dM}{df} = \frac{1}{f}(D + \lambda - M),$$

and

$$\frac{d\lambda}{df} = \frac{1}{f} 2\lambda.$$

Moreover, from (2.1b),

$$\frac{dD}{df} = \frac{1}{f}\frac{1}{1 - 2M}\left[D(2 - 4M) - \frac{(D + P)(P + M - \lambda)}{dp/d\mu}\right]$$

whenever $dp/d\mu \not\equiv 0$. This now provides a system of three first-order ordinary differential equations. If we further assume that $p = (\gamma - 1)\mu$, where γ is a constant satisfying $\gamma \neq 0$, $\gamma \neq 1$, and if we write $t = \ln f$, we obtain the *autonomous* system

$$\frac{dD}{dt} = \frac{D}{1-2M}\left[2 - \left(\frac{5\gamma-4}{\gamma-1}\right)M - \gamma D + \frac{\lambda\gamma}{\gamma-1}\right], \tag{2.2a}$$

$$\frac{dM}{dt} = D + \lambda - M, \tag{2.2b}$$

and

$$\frac{d\lambda}{dt} = 2\lambda. \tag{2.2c}$$

This system generalizes that considered previously,[6,7] wherein $\Lambda = 0$ (so $\lambda = 0$) and the system is examinable by phase-plane analysis. The use of the variables D, M, and λ in deriving Eqs. (2.2) is suggested by the invariance property (1.3b) above, since D, M, and λ are each invariant under (1.3b). The equation of state $p = (\gamma - 1)\mu$ is also of a form which respects (1.3b). [Note that it is possible to obtain a similar system of autonomous equations directly out of the system (1.1), using variables constructed from f and g and their derivatives, which are invariant under the transformations (1.3). The alternative approach used herein is thought to provide more physical insight, since the quantity $m(f)$ is related to the total mass within a radius f.] The denominator $1 - 2M$ in Eq. (2.2a) does not vanish identically, for, if it did, then $2m \equiv f$ in (2.1a), and so $f' \equiv 0$, a contradiction.

It is of interest to observe that the system of Eqs. (2.2) has exactly two fixed points, where the right-hand sides vanish identically. These are given by

(i) $D = M = \lambda = 0$. This corresponds to Minkowski space-time, since by (1.1c) and (2.1a), $f'^2 \equiv 1$ and $g' \equiv 0$.

(ii) $D = M = 2(\gamma - 1)/[(\gamma + 2)^2 - 8]$; $\lambda = 0$. In this case the energy density μ is

$$\mu = \frac{2D}{f^2} = \frac{4(\gamma-1)}{(\gamma+2)^2 - 8}\,\frac{1}{f^2}.$$

Whereas the solution (ii) has been ascribed to Misner and Zapolsky,[6,8,9] it is a special case of the class VI solutions of Tolman[10] and has been discussed by Wyman[11] and others.[4] It does not have a regular center at $f = 0$ (since $\mu \to \infty$ as $f \to 0$), and it has only one center, since

$$\left(\frac{df}{dr}\right)^2 = 1 - 2M \neq 0. \tag{2.3}$$

Astrophysically, it is the relativistic analog of a special singular solution of Chandrasekhar[12] for certain Newtonian polytropes. The functions f and g in the metric can be determined directly by integrating (1.2) and (2.3).

The two solutions (i) and (ii) above are special solutions valid when $\Lambda = 0$. In the general $\Lambda = 0$ case, the system (2.2) reduces to one that has been examined by phase-plane methods.[6] It has been shown that there are only two solutions of interest, i.e., in which $m = 0$ when $f = 0$. One is the "Misner–Zapolsky" solution, and the other is one which, as far as I am aware, is not known in exact analytic form. It possesses the property that μ is finite and nonzero at $f = 0$, and it extends out to infinite values of f (and infinite proper distance), where it approaches the Misner–Zapolsky solution. It corresponds to the usual static spherically symmetric stellar model with a *regular* center [and with an equation of state $p = (\gamma - 1)\mu$]. Details are provided in Ref. 6. Treated as a cosmological model, it represents a complete unbounded universe with negative red-shifts as observed from the central world-line $f = 0$ (cf. Refs. 1 and 13).

A full qualitative investigation of the system (2.2) would be of great interest, although presently available mathematical techniques appear to be incapable of providing this. Alternatively, we could search for situations in which the system (2.2) can be reduced to a subsystem of two equations, as in the $\Lambda = 0$ case we have just discussed. One example of this occurs if $\gamma = 0$, for then the variables $D + \lambda$ and M can be used; however, this case is not of physical interest since the matter would then satisfy the unrealistic equation of state $\mu + p \equiv 0$. In fact, I suspect that there are no tractable subsystems of (2.2), in addition to those already mentioned.

3. CONDITIONS AT THE CENTER

From Eq. (1.2), it follows that if $p = \frac{1}{3}\mu$ then $\mu g^4 = \text{const}$, and so it is clear that if μ_0, the central value of the energy density, is zero, then $\mu \equiv 0$. This result readily generalizes to γ-law equations of state, but its generalization to other equations of state remains obscure. The difficulty arising here is due to the fact that the system of field Eqs. (1.1) and (1.2) is not regular at the center, $r = 0$. As a result of this, the usual uniqueness theorems for solutions of systems of ordinary differential equations no longer hold. It does not seem possible to "regularize" the equations in such a way as to avoid this difficulty, which is also very apparent when one considers the orthonormal tetrad formulation of the problem, there being no regular orthonormal tetrad field adapted to the spherical symmetry in a neighborhood of a central point.

I am not aware of any proof that when $p = \frac{1}{3}\mu$, $\mu_0 = \Lambda$ implies $\mu \equiv \Lambda$. Again, the same difficulties are encountered at $r = 0$. When $p = \frac{1}{3}\mu$, the case $\mu \equiv \Lambda$ is the Einstein static model, generalized to the given equation of state. For a general equation of state, $p = p(\mu)$, and given cosmological constant, the Einstein static model would be characterized by the condition $\mu + 3p \equiv 2\Lambda$, and it is not clear when the central condition $\mu_0 + 3p_0 = 2\Lambda$ necessarily gives rise to only this case. It is possible to show that, for certain equations of state, the central condition $\mu_0 + 3p_0 = 2\Lambda$ is satisfied by solutions other than the Einstein static solution, i.e., at least for certain equations of state, *the usual conditions at the center* ($r = 0$) *do not uniquely determine the space-time geometry elsewhere* (*at* $r > 0$). This is an important point, because the standard numerical procedures used in analyzing static spherically symmetric systems in general relativity (see, e.g., Ref. 14) involve integration from the center, and it is clearly highly desirable to be able to demarcate those cases in which uniqueness ensues. The following argument shows that there are certainly situations when, for a given equation of state, the space-time is not uniquely determined by the usual conditions at the center.

Equations (1.1a)–(1.1c) are equivalent to (1.1b), (1.1c), and (1.2). Let $f(r)$ be an arbitrary function of r, analytic and odd on some interval $(-R, R)$, $R > 0$; we suppose that f satisfies the conditions of regularity at the center $r = 0$, so that[1]

$$f(r) \to 0,\ f'(r) \to 1 \text{ and } f''(r)/f(r) \to \text{finite limit as } r \to 0$$

(so f is expanded about the origin in the form $f(r) = r + \Sigma_{k=1}^{\infty} a_{2k+1} r^{2k+1}$). Thus $(f'^2 - 1)/f^2 \to 6a_3$ as $r \to 0$, and $(f'^2 - 1)/f^2$ is analytic (and even) on $(-R,R)$. By Eq. (1.1b), it follows that μ is analytic (and even) on $(-R,R)$. We may expand μ about the origin in the form $\mu(r) = \Sigma_{k=0}^{\infty} \mu_{2k} r^{2k}$, where $\mu_0 = -\Lambda - 18a_3$ and $\mu_2 = 5(3a_3^2 - 10a_5)$. Eliminating g'/g between (1.1c) and (1.2), we have

$$p' = \left(\frac{\mu + p}{2}\right) \frac{f}{f'} \left(\frac{f'^2 - 1}{f^2} - p - \Lambda\right). \tag{3.1}$$

Since the right-hand side of (3.1) is analytic in p and r, it follows[15] that p is analytic, with p_0 arbitrarily specifiable. Choose p_0 such that $\mu_0 + p_0 \neq 0$. Then, writing (1.2) in the form

$$g' = -\frac{gp'}{\mu + p}, \tag{3.2}$$

we can apply the analyticity argument to (3.2), and conclude that $g(r)$ is analytic in r, with $g(0)$ freely specifiable. We may set $g(0) = 1$, without loss of generality. Clearly g satisfies the conditions of regularity at the center $r = 0$, so that[1]

$$g(r) \to \text{finite nonzero limit}, \quad g'(r) \to 0 \text{ as } r \to 0,$$

the latter following by letting $r \to 0$ in (3.1) and (3.2).

We now suppose that p_0 is chosen so that $\mu_0 + 3p_0 = 2\Lambda$, and that $10a_5 \neq 3a_3^2$, so that $\mu_2 \neq 0$. Then, fixing the values of all a_{2k+1}, we have a solution of Einstein's field equations in which $\mu' \not\equiv 0$ (since $\mu_2 \neq 0$), and in which $p' \not\equiv 0$ [since $p' \equiv 0$ requires, by (1.1) and (1.2), $\mu + 3p \equiv 2\Lambda$, implying $\mu' \equiv 0$, a contradiction]. The matter is inhomogeneous, with an equation of state $p = p(\mu)$ obtained by eliminating r from the relationships $p = p(r)$ and $\mu = \mu(r)$. On the other hand, given this specific equation of state, there is an Einstein static model with the same central values of μ and p, and with the same value of the cosmological constant, and this is of course spatially homogeneous, and therefore distinct from the previous solution.

Thus, for a fixed equation of state and cosmological constant Λ, it does not always follow that the central values of the energy density and pressure will uniquely determine the solution. It would be valuable to understand the clear-cut circumstances under which, for any fixed equation of state and cosmological constant Λ, one obtains uniqueness, but such an investigation is beyond the scope of the present paper.

4. NONSTATIC SPHERICALLY SYMMETRIC COSMOLOGIES

Ellis, Maartens, and Nel[1] speculate that some of the interesting features of static spherically symmetric models might be preserved in expanding models, which would moreover be more realistic. The simplest generalization which preserves spherical symmetry and yet introduces expansion occurs when the world-lines of the galaxies have no distortion (i.e., no shear), since the static case is characterized by requiring that both the shear and the expansion vanish. Now Mansouri[16] has recently shown that when the cosmological constant Λ is zero, shear-free spherically symmetric perfect fluid general-relativistic space-times in which $p = p(\mu)$, $\mu + p \neq 0$, *and in which there is a comoving timelike hypersurface of zero pressure* are necessarily either static or spatially homogeneous and isotropic Friedmann–Robertson–Walker models. Glass[17] has presented an alternative proof of Mansouri's results, in which the arguments are made somewhat more transparent. Further clarification and generalizations of Mansouri's result have been obtained by Collins and Wainwright.[18] It may be suspected as a consequence of Mansouri's result that the condition related to the timelike hypersurface of zero pressure could be relaxed, i.e., that (with $\Lambda = 0$) there could be no expanding shear-free spherically symmetric perfect fluid general-relativistic models with an equation of state $p = p(\mu)$, other than the Friedmann–Robertson–Walker models. However, this is not the case. One such set of solutions, having $\Lambda = 0$ and ascribed to Wyman,[19] appears in the book of exact solutions by Kramer, Stephani, MacCallum, and Herlt.[4] Contrary to the claims in Ref. 4, this set is not the most general one [the special case $A(t) \equiv 0$ in (14.35) of Ref. 4 has been overlooked (in 14.57) when $p = p(\mu)$, although Wyman[19] considered it[20]]. These solutions are easily generalized to the case when $\Lambda \neq 0$ (see Refs. 18 and 19). In addition, there is an analogous set of solutions are easily generalized to the case when $\Lambda \neq 0$ (see Refs. 18 and 19). In addition, there is an analogous set of solutions with plane symmetry. *Together with the spatially homogeneous and isotropic Friedmann–Robertson–Walker models, these plane symmetric and spherically symmetric models comprise the only expanding irrotational shear-free perfect fluid general relativistic models with an equation of state $p = p(\mu)$ in which $\mu + p \neq 0$.* Proofs of these results appear in Ref. 18, and rely to some extent on an article by Barnes,[21] which treats shear-free irrotational perfect fluids in general relativity. What is not clear from Barnes' paper is whether there actually exist anisotropic solutions *which admit an equation of state, $p = p(\mu)$*, and, if so, what those solutions are. This question is completely answered in Ref. 18. ticularly with regard to the role that the inhomogeneous spherically symmetric family could play in the observation-based philosophy-free study of cosmology initiated by Ellis, Maartens, and Nel.[1]

ACKNOWLEDGMENTS

This work was supported by an operating grant from the Natural Sciences and Engineering Research Council of Canada. Dr. J. Wainwright first proved that expanding irrotational shear-free perfect fluid general relativistic models, with an equation of state $p = p(\mu)$ in which $\mu + p \neq 0$, necessarily possess spherical, planar, or hyperbolic symmetry, and he first established the existence of the plane symmetric models. I am grateful to him for many useful discussions. I am also grateful to Helen Warren for her careful typing of the manuscript.

[1] G. F. R. Ellis, R. Maartens, and S. D. Nel, Mon. Not. R. Astron. Soc. **184**, 439 (1978).

[2] We choose geometrical units, in which $8\pi G = c = 1$, where G is the Newtonian gravitational constant, and c is the velocity of light in a vacuum. The signature of the space-time metric is $(-+++)$, and the conventions for the Riemann tensor, Ricci tensor, and Ricci scalar are defined,

respectively, by $v^i_{;l;k} - v^i_{;k;l} = R^i_{\ jkl}v^j$, $R_{ij} = R^k_{\ ikj}$, and $R = R^i_{\ i}$, where v^i is any (sufficiently differentiable) vector field. The tensors g_{ij}, $G_{ij} = R_{ij} - \frac{1}{2}Rg_{ij}$, and T_{ij} are, respectively, the metric tensor, the Einstein tensor, and the energy-momentum tensor.

[3]S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge U.P., Cambridge, 1973).

[4]D. Kramer, H. Stephani, M. A. H. MacCallum, and E. Herlt, *Exact Solutions of Einstein's Field Equations* (Cambridge U.P., Cambridge, New York, New Rochelle, Melbourne, and Sydney, 1980), and references cited.

[5]G. F. R. Ellis, "Relativistic Cosmology" in *General Relativity and Cosmology*, Proc. Int. Sch. Phys. "Enrico Fermi" Course XLVII, edited by R. K. Sachs (Academic, New York and London, 1971).

[6]C. B. Collins, J. Math. Phys. **18**, 1374 (1977).

[7]Equation (2.2a), with $\lambda = 0$, corrects a misprint in Eq. (3.1) of Ref. 6. We have changed the notation to conform with that of Refs. 1 and 4.

[8]C. W. Misner and H. S. Zapolsky, Phys. Rev. Lett. **12**, 635 (1964).

[9]S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972).

[10]R. C. Tolman, Phys. Rev. **55**, 364 (1939). The special "Misner–Zapolsky" solution is obtained by putting $\Lambda = 0$ and $AB = 0$ (with $A^2 + B^2 \neq 0$) in Eq. (4.6) of Tolman's paper.

[11]M. Wyman, Phys. Rev. **75**, 1930 (1949).

[12]S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (Dover, New York, 1969).

[13]B. Schmidt, Kugelsymmetrische statische Materielösungen der Einsteinschen Feldgleichungen (Diplomarbeit, Hamburg University, 1966).

[14]J. B. Hartle, "Relativistic Stars, Gravitational Collapse and Black Holes" in *Relativity, Astrophysics and Cosmology*, edited by W. Israel (Reidel, Boston, 1973).

[15]G. Birkhoff and G.-C. Rota, *Ordinary Differential Equations* (Blaisdell, Waltham, Mass., 1969).

[16]R. Mansouri, Ann. Inst. Henri Poincaré **27**, 175 (1977).

[17]E. N. Glass, J. Math. Phys. **20**, 1508 (1979).

[18]C. B. Collins and J. Wainwright, "On the Role of Shear in General Relativistic Cosmological and Stellar Models," preprint, University of Waterloo, Canada (1982).

[19]M. Wyman, Phys. Rev. **70**, 396 (1946).

[20]Note that the special case $A(t) \equiv 0$ in Eq. (14.35) of Ref. 4 *is* considered. However, we are concerned here with solutions which not only have $A(t) \equiv 0$, but also possess an equation of state $p = p(\mu)$. In Ref. 4, the final reduction in the case $p = p(\mu)$ is valid only if $A(t) \not\equiv 0$, whereas if $A(t) \equiv 0$ another solution (not given) is arrived at. Further details are provided in Ref. 18.

[21]A. Barnes, Gen. Relativ. Gravit. **4**, 105 (1973).

respectively, by $v^i_{;l;k} - v^i_{;k;l} = R^i_{\ jkl}v^j$, $R_{ij} = R^k_{\ ikj}$, and $R = R^i_{\ i}$, where v^i is any (sufficiently differentiable) vector field. The tensors g_{ij}, $G_{ij} = R_{ij} - \frac{1}{2}Rg_{ij}$, and T_{ij} are, respectively, the metric tensor, the Einstein tensor, and the energy-momentum tensor.

[3]S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge U.P., Cambridge, 1973).

[4]D. Kramer, H. Stephani, M. A. H. MacCallum, and E. Herlt, *Exact Solutions of Einstein's Field Equations* (Cambridge U.P., Cambridge, New York, New Rochelle, Melbourne, and Sydney, 1980), and references cited.

[5]G. F. R. Ellis, "Relativistic Cosmology" in *General Relativity and Cosmology*, Proc. Int. Sch. Phys. "Enrico Fermi" Course XLVII, edited by R. K. Sachs (Academic, New York and London, 1971).

[6]C. B. Collins, J. Math. Phys. **18**, 1374 (1977).

[7]Equation (2.2a), with $\lambda = 0$, corrects a misprint in Eq. (3.1) of Ref. 6. We have changed the notation to conform with that of Refs. 1 and 4.

[8]C. W. Misner and H. S. Zapolsky, Phys. Rev. Lett. **12**, 635 (1964).

[9]S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (Wiley, New York, 1972).

[10]R. C. Tolman, Phys. Rev. **55**, 364 (1939). The special "Misner–Zapolsky" solution is obtained by putting $\Lambda = 0$ and $AB = 0$ (with $A^2 + B^2 \neq 0$) in Eq. (4.6) of Tolman's paper.

[11]M. Wyman, Phys. Rev. **75**, 1930 (1949).

[12]S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (Dover, New York, 1969).

[13]B. Schmidt, Kugelsymmetrische statische Materielösungen der Einsteinschen Feldgleichungen (Diplomarbeit, Hamburg University, 1966).

[14]J. B. Hartle, "Relativistic Stars, Gravitational Collapse and Black Holes" in *Relativity, Astrophysics and Cosmology*, edited by W. Israel (Reidel, Boston, 1973).

[15]G. Birkhoff and G.-C. Rota, *Ordinary Differential Equations* (Blaisdell, Waltham, Mass., 1969).

[16]R. Mansouri, Ann. Inst. Henri Poincaré **27**, 175 (1977).

[17]E. N. Glass, J. Math. Phys. **20**, 1508 (1979).

[18]C. B. Collins and J. Wainwright, "On the Role of Shear in General Relativistic Cosmological and Stellar Models," preprint, University of Waterloo, Canada (1982).

[19]M. Wyman, Phys. Rev. **70**, 396 (1946).

[20]Note that the special case $A(t) \equiv 0$ in Eq. (14.35) of Ref. 4 *is* considered. However, we are concerned here with solutions which not only have $A(t) \equiv 0$, but also possess an equation of state $p = p(\mu)$. In Ref. 4, the final reduction in the case $p = p(\mu)$ is valid only if $A(t) \not\equiv 0$, whereas if $A(t) \equiv 0$ another solution (not given) is arrived at. Further details are provided in Ref. 18.

[21]A. Barnes, Gen. Relativ. Gravit. **4**, 105 (1973).