Bilateral classes

W. Hässelbarth and E. Ruch
Institut für Quantenchemie, Freie Universität, Berlin, West Germany
D. J. Klein
Department of Physics, University of Texas, Austin, Texas
T. H. Seligman
Instituto de Fisica, Universidad Nacional Autónoma de México, Mexico 20, D. F.

(Received 2 October 1979; accepted for publication 14 January 1980)

We introduce "bilateral classes", a new concept for classifying the elements of groups. Bilateral classes are orbits in a group G under the action of any given subgroup of the direct product $G \times G$. The classification concept presented here encompasses conjugacy classes, cosets, double cosets, and Ree's σ-classes as particular cases. It has an interpretation as a classification scheme for bijections between G-sets under the aspect of symmetry equivalence due to symmetries in both the domain and the range. While double cosets and conjugacy classes correspond to the case of no or complete correlation between operations of the two symmetry groups, our concept also covers the general case of partial correlation. The scope of generalization corresponds to applications in physics. Expressions for the number of bilateral classes are given.

In classifying the elements of groups some particular types of classes have been proved to represent classification principles of outstanding relevance in various fields of application, i.e., cosets, double cosets, conjugacy classes and subclasses.[1] Denoting with A and B two subgroups of a group G, the corresponding definitions may be summarized as follows:

(i) right cosets $\quad Ag:=\{ag \mid a\in A\}$;
(ii) left cosets $\quad gB:=\{gb \mid b\in B\}$;
(iii) double cosets $\quad AgB:=\{agb \mid a\in A, b\in B\}$,
(iv) conjugacy classes[2] $\quad C^A_{(g)}:=\{aga^{-1} \mid a\in A\}$.

While cosets evidently are particular double cosets, as they correspond to the special choice $A=I$ or $B=I$, respectively, with I being the identity subgroup, no such relation between double cosets and conjugacy classes can be stated. There is, however, a common aspect in the generation of these classes. Elements which are equivalent to a $g\in G$ are generated by multiplying g from the right and from the left with elements of given subgroups, respectively, the pairing of left and right factors being either unrestricted or subject to a certain condition. Each of the above classifications may be characterized, by an equivalence relation which is defined by a properly chosen subset P of the direct product $G\times G$ as follows:

$$g' \sim g \Longleftrightarrow \exists (a,b)\in P: g' = agb^{-1}. \tag{1}$$

In fact, left, right, double cosets and conjugacy classes, respectively, arise if P is chosen to be $A\times I, I\times B, A\times B$, and $(A\times A)_D$, where $(A\times A)_D$ denotes the diagonal subgroup $\{(a,a) \mid a\in A\}$ of $A\times A$. However, the restriction of P to these particular choices seems quite arbitrary and it is unnecessary for defining an equivalence relation. It is sufficient for this purpose that P is a subgroup of $G\times G$. Thus, for any subgroup P, (1) represents a subdivision of G into disjoint subsets $P[g]$, which we shall call bilateral classes[3] of type P.

Definition 1:

$P[g]:=\{agb^{-1} \mid (a,b)\in P\}$.

An understanding of the characteristic features of bilateral classes in G as well as a survey of the essential particularizations into types with distinguished properties may be expected from characterizing the subgroups of $G\times G$ through properties of G. As shown by Goursat[4] a subgroup $P\subset G\times G$ is given first by a chain of subgroups $\bar{A}\subset A\subset G$ and $\bar{B}\subset B\subset G$ in each factor G where $\bar{A}$ and $\bar{B}$ are normal subgroups in A and B, respectively, satisfying the condition of isomorphic factor groups $A/\bar{A}\cong B/\bar{B}$ and second by an isomorphism $\mu: A/\bar{A}\leftrightarrow B/\bar{B}$. The subgroup P, associated with $\bar{A}, A, \bar{B}, B$, and μ, is the disjoint union of products of correlated cosets due to μ, i.e.,

$$P=\bigcup_{s=1}^{p}(\bar{A}a_s\times\bar{B}b_{\hat{\mu}(s)}), \tag{2}$$

where $a_s, b_s;\ s=1,\dots,p=|A/\bar{A}|$ are coset representatives and $\hat{\mu}: s\mapsto\hat{\mu}(s)$ denotes a one-to-one correspondence between cosets of $\bar{A}$ in A and $\bar{B}$ and B according to $\mu(\bar{A}a_s)=\bar{B}b_{\hat{\mu}(s)}$. A more symmetrical and equally useful presentation refers to a group H isomorphic with $A/\bar{A}$ and $B/\bar{B}$ and instead of μ to a pair of homomorphisms $\varphi: A\to H$ and $\psi: B\to H$ which satisfy the relation

$$\varphi(a)=\psi(b)\Longleftrightarrow\mu(\bar{A}a)=\bar{B}b,\quad \forall(a\in A, b\in B).$$

With φ and ψ we may rewrite (2) in the form (2')

$$P=\{(a,b) \mid (a,b)\in A\times B;\ \varphi(a)=\psi(b)\}. \tag{2'}$$

Using (2) and (2') we get two equivalent but more explicit formulations of Def. 1, namely:

Definition 1a:

$$P[g]:=\bigcup_{s=1}^{p}\bar{A}a_s g b^{-1}_{\hat{\mu}(s)}\bar{B},$$

Definition 1b:

$$P[g]:=\{agb^{-1} \mid a\in A, b\in B;\ \varphi(a)=\psi(b)\}.$$

According to Def. 1a bilateral classes may be considered as

 0022-2488/80/050951-03$01.00

particular unions of double cosets. Def. 1b characterizes the classification by the set of equivalence transformations $(a,b):g \mapsto agb^{-1}$ and shows that in any case the totality of left and right factors are each subgroups of G. The permitted combinations of these factors, however, are generally not unrestricted since the condition $\varphi(a) = \psi(b)$ becomes trivial only in the special case $H = I$. The properties of H allow a distinction between characteristic types of bilateral classes.

The case $H = I$ corresponds to $\overline{A} = A$, $\overline{B} = B$ and $\forall(a \in A, b \in B): \varphi(a) = \psi(b)$, i.e., it represents double cosets. Thus, double cosets are bilateral classes without any correlation between right and left factors in the set of equivalence transformations.

The case $H \cong A \cong B$ implies $\overline{A} = \overline{B} = I$, and therefore the condition $\varphi(a) = \psi(b)$ corresponds to an isomorphism $\mu: A \mapsto B$. Hence the pairing of elements of A and B is unique which means that bilateral classes of this type are distinguished by strong correlations between right and left factors in the set of equivalence transformations. The particular case $A = B$ and μ being the identity isomorphism denotes conjugacy classes. For all other cases of strong correlation we propose the term "twisted conjugacy classes".

A class concept for groups different from those mentioned in the beginning was introduced by Ree[5] under the notation σ-classes. They also prove to be a certain type of bilateral classes, characterized by $H \cong B \cong A/\overline{A}, \overline{A} \neq I, B \subset A$, or $H \cong A \cong B/\overline{B}, \overline{B} \neq I, A \subset B$.

Between the extreme situations of strong correlation and no correlation we have the general type of bilateral classes, which is distinguished by partial correlation of right and left factors in the set of equivalence transformations. We think that bilateral classes represent the natural extension from the classical types of classes in group theory to a unifying and generalizing concept.

Let χ be the mapping from $G \times G$ onto G defined by $\chi:(g_\alpha, g_\beta) \mapsto g_\alpha g_\beta^{-1}$. It follows that

$$\forall(h \in G): \chi:(g_\alpha h, g_\beta h) \mapsto g_\alpha g_\beta^{-1},$$

$$g_\alpha g_\beta^{-1} = g'_\alpha g'^{-1}_\beta \Rightarrow g'_\alpha = g_\alpha h, \quad g'_\beta = g_\beta h,$$

with

$$h = g_\alpha^{-1} g'_\alpha = g_\beta^{-1} g'_\beta.$$

Thus by the action of χ all elements of a left coset of $(G \times G)_D$ in $G \times G$ are mapped onto one element of G, and different left cosets are mapped onto different elements of G. Any such left coset, therefore, may be written as $\{(gh,h) \mid h \in G\}$ where $\{(g,e) \mid g \in G\}$ is a system of representatives in $G \times G$.

Given any subgroup $P \subset G \times G$, we have

$$\{(ag_\alpha h, bg_\beta h) \mid (a,b) \in P, b \in G\} \to \{ag_\alpha g_\beta^{-1} b^{-1} \mid (a,b) \in P\},$$

and hence the bijection

$$\{P(gh,b)(G \times G)_D \mid g \in G, b \in G\}$$
$$\mapsto \{\{agb^{-1} \mid (a,b) \in P\} \mid g \in G\}.$$

This entails the

Lemma: The mapping $\chi:(g_\alpha, g_\beta) \mapsto g_\alpha g_\beta^{-1}$ induces a one-to-one correspondence between double cosets in $G \times G$ with P as left and $(G \times G)_D$ as right factor and bilateral classes of type P in G.

The following discussion illustrates this lemma. It also provides an interpretation of bilateral classes in the symmetric group which indicates applications in combinatorics.

The concept of symmetry in a set M may be introduced with reference to a second set R of equal cardinal number using bijections from R onto M. The set of bijections is given by $\{g \circ \xi \mid g \in S\}$ where ξ is a fixed bijection $\xi: R \mapsto M$ and $g \circ \xi$ denotes the bijection ξ followed by a permutation acting on M. S denotes the unrestricted symmetric group of permutation operators. Any subgroup A of S, thus can be considered as representing a particular symmetry in M, which is characterized through symmetry equivalence of bijections as follows:

$$g' \circ \xi \sim g \circ \xi \Longleftrightarrow \exists a \in A: g' = ag.$$

Stipulating the relation $\forall g \in S: g \circ \xi = \xi \circ g$ we define the action of operators of S also on the set R. Let N be a further set of equal cardinal number and the group S be defined also as permutation operator group for N by a bijection $\eta: R \mapsto N$, $\forall g \in S: g \circ \eta = \eta \circ g$. The product $\xi \circ \eta^{-1} = \epsilon$ is then a bijection from N onto M satisfying $\forall g \in S: g \circ \epsilon = \epsilon \circ g$; i.e., S is operator group for M, N, and R. The set of bijection pairs $\{(g \circ \xi, g' \circ \eta) \mid (g,g') \in S \times S\}$ from R onto M and N classified according to subsets of pairs which represent the same bijection $g \circ \xi \circ \eta^{-1} \circ g'^{-1}: N \mapsto M$, respectively, leads to the following classes:

$$\{(gh \circ \xi, g'h \circ \eta) \mid h \in S\}, \quad (g,g') \in S \times S.$$

We introduce the concept of "correlated symmetry" in the product set $M \times N$ distinguished as a symmetry which corresponds to a subgroup of $S \times S$ and discuss it first with reference to the set R using pairs of bijections $g \circ \xi: R \mapsto M$, $g' \circ \eta: R \mapsto N$. Since correlation between properties of M and N must be invariant under permutations in the reference set R, exclusively the membership of bijection paris in classes representing a bijection between N and M respectively is of relevance for classifying according to correlated symmetry. Given a symmetry group $P \subset S \times S$, the corresponding classes are therefore determined by double cosets with $(S \times S)_D$ as right and P as left factor.

$$\{(agh \circ \xi, bg'h \circ \eta) \mid (a,b) \in P, h \in S\}, \quad (g, g') \in S \times S.$$

In terms of bijections from N onto M which we write as products $g \circ \epsilon = g'h \circ \xi \circ \eta^{-1} \circ h^{-1} g''^{-1}, g'g''^{-1} = g$ this classification corresponds to bilateral classes in S

$$\{agb^{-1} \circ \epsilon \mid (a,b) \in P\} \text{ or } \{\bar{b}\bar{g}^{-1}\bar{a}^{-1} \mid (\bar{b},\bar{a}) \in \bar{P}\},$$
$$\bar{P} \subset \overline{S} \times \overline{S},$$

where $\overline{S}$ is the operator group associated (inverse isomorphic) with S and acting on functions. According to Goursat's decomposition of subgroups (2), correlated symmetry represents total, partial or no correlation between the symmetry operations in M and N.

More generally let M and N be two G-sets associated with a given group G and $\epsilon: M \to N$ be a map with $\epsilon(gm) = g \circ \epsilon(m)$ for all $g \in G$ and $m \in M$. Let $\{\varphi\}$ be the set of maps which arises from ϵ by the combination with translations $g \in G$, i.e., $\varphi = g \circ \epsilon = \epsilon \circ g: M \to N$ and H denoting the subgroup of G for which $h \circ \epsilon = \epsilon$, $\forall h \in H$. Then, any subgroup P of $G \times G$ which contains $H \times H$ represents a correlated symmetry in $M \times N$, and classes of symmetry equivalent maps of the above type are defined by

$$\varphi \underset{P}{\sim} \psi \Longleftrightarrow \exists (a.b) \in P: a \circ \varphi \circ b^{-1} = \psi.$$

Since all these maps can be represented by left multiplication of ϵ with an element g of G we have a one-to-one correspondence between bilateral classes of the type P and symmetry equivalence classes of maps in the set $\{\varphi\}$. The foregoing result refers to the particular case of bijections between sets of equal cardinal number.

Summarizing we have:

Theorem: Any subgroup P of $G \times G$ represents a correlated symmetry in two properly chosen G-sets M and N. Bilateral classes of type P in the group G represent the classification of maps $\{\varphi\}$ between the sets M and N, which arise from a G-map ϵ, $g \circ \epsilon = \epsilon \circ g$ by translation (i.e., by composition with elements in G), according to the "correlated symmetry P".

The equivalence transformations $(a,b): g \mapsto agb^{-1}$ define an action of $P \subset G \times G$ on G. Hence bilateral classes are P-orbits of G, and we can make use of well-known results concerning the sizes and the number of orbits of a permutation group.

The order $|P[g]|$ of a bilateral class $P[g]$ containing the element g is given by the number of cosets in P of the stabilizer P_g of g. In our case, the stabilizer takes the form

$$P_g = \{(a,b) \in P \mid agb^{-1} = g\}.$$

Because of the equivalence

$$\bar{a}a_s g b_{\bar{\mu}(s)}^{-1} \bar{b}^{-1} = g \Longleftrightarrow \bar{a}a_s = g\bar{b}b_{\bar{\mu}(s)}g^{-1},$$

the presentation (2) leads to the expression

$$P_g = \bigcup_{s=1}^{p} \{(a, g^{-1}ag) \mid a \in \bar{A}a_s \cap g\bar{B}b_{\bar{\mu}(s)}g^{-1}\};$$

$$|P_g| = \sum_{s=1}^{p} |\bar{A}a_s \cap g\bar{B}b_{\bar{\mu}(s)}g^{-1}|.$$

The index of the stabilizer P_g gives the order of $P[g]$

$$|P[g]| = \frac{|P|}{|P_g|} = \frac{p|\bar{A}||\bar{B}|}{\Sigma_{s=1}^{p} |\bar{A}a_s \cap g\bar{B}b_{\bar{\mu}(s)}g^{-1}|}.$$

The number $|G/\sim P|$ of classes of type P follows from Burnside's lemma:

$$|G/\sim P| = \sum_{g \in G} \frac{1}{|P[g]|}$$

$$= \frac{1}{p|\bar{A}||\bar{B}|} \sum_{g \in G} \sum_{s=1}^{p} |\bar{A}a_s \cap \bar{B}b_{\bar{\mu}(s)}g^{-1}|.$$

Another counting formula results from the one-to-one correspondence between bilateral classes and double cosets:

$$|G/\sim P| = |P \backslash G \times G/(G \times G)_D|.$$

We recall the number $|A \backslash G/B|$ of double cosets of a group G with respect to subgroups A and B, given by a formula which goes back to Frobenius.[6]

$$|A \backslash G/B| = \frac{|G|}{|A||B|} \sum_{\rho} \frac{|C_\rho \cap A| \, |C_\rho \cap B|}{|C_\rho|}.$$

Here the sum is taken over the (ordinary) conjugacy classes of G.

Counting the number $|P \backslash G \times G/(G \times G)_D|$ requires essentially the same prerequisites since conjugacy classes of $G \times G$ are cartesian products of those of G.

$$|G/\sim P|$$

$$= \frac{|G|^2}{|P|\,|G|} \sum_{\rho} \sum_{\sigma} \frac{|C_\rho \times C_\sigma \cap P| \, |C_\rho \times C_\sigma \cap (G \times G)_D|}{|C_\rho| \, |C_\sigma|}$$

$$= \frac{|G|}{|P|} \sum_{\rho} \frac{|C_\rho \times C_\rho \cap P|}{|C_\rho|}.$$

With the presentation of the group P according to (2), we finally get the enumeration formula

$$|G/\sim P|$$

$$= \frac{|G|}{p|\bar{A}|\,|\bar{B}|} \sum_{\rho} \sum_{s=1}^{p} \frac{|C_\rho \cap \bar{A}a_s| \, |C_\rho \cap \bar{B}b_{\bar{\mu}(s)}|}{|C_\rho|}.$$

[1]See, e.g., E.P. Wigner, in *Spectroscopic and Group Theoretical Methods in Physics*, edited by F. Bloch (Wiley, New York, 1958), p. 131.

[2]Frequently the term conjugacy class is reserved for the particular case of $A = G$, while subclasses denote the case $A \neq G$.

[3]The notation "double classes" originally used by us in a preliminary note[7] on this subject, is replaced by 'bilateral classes'.

[4]E. Goursat, Annales Scientifique de l'Ecole Normale Supérieure, Paris (3), **6**, 9 (1889).

[5]R. Ree, III. J. Math. **3**, 440 (1959).

[6]G. Frobenius, SitzBer. Preuss, Akad., Berlin 163 (1895).

[7]W. Hässelbarth, E. Ruch, D.J. Klein, T.H. Seligman, in *Group Theoretical Methods in Physics*, edited by R.T. Sharp and B. Kolman (Academic, New York, 1977), p. 617.

Commutation properties of 2-parameter groups of isometries

P. Ernotte
Inst. de Physique Théorique, Université Catholique de Louvain, B-1348 Louvain-la-neuve, Belgium

(Received 13 March 1979; accepted for publication 23 October 1979)

We prove a stronger version of a result due to B. Carter: A cyclic subaction must commute with any 1-parameter subgroup of a 2-parameter group of isometries.

We prove the following improvement of a result due to B. Carter (cfr. Ref. 1)

Theorem: If a pseudo-Riemannian manifold (M, g) is invariant under a 2-parameter group of isometries G and if G contains a 1-parameter cyclic group of isometries Π^c then Π^c commutes with any other 1-parameter subgroup if at least an orbit of Π^c is nonnull at a point P of M.

The proof is a direct consequence of the following lemma. It will provide the derivative, along a Killing vector ξ_i, of the determinant of the matrix $g(\xi_i,\xi_j)$

(In the following we refer to Refs. 2 and 3 for the theorems and notation used.)

Lemma: H_1: $\xi_1,\xi_2\cdots\xi_p$ are p Killing vectors which are linearly independent at each point of an open subset U of M.

They span a p-dimensional linear subspace of $T_x(M)$ called Σ_x . H_2: $[\xi_i,\xi_j] = \Sigma_k C^k_{ij}\xi_k$. Then

(a) $\xi_i(\omega_k) = \Sigma_l\, C^l_{ik}\omega_l$, where $\omega_k \equiv g(\xi_k,\ \cdot)$ and $\xi_i(\omega_k)$ is the Lie derivative of the 1-form ω_k along the vector field ξ_i;

(b) $\xi_i(\Omega) = \Sigma_l\, C^l_{il}\Omega$ with $\Omega \equiv \omega_1\wedge\cdots\wedge\omega_p$;

(c) $\xi_i(W) = 2\,\Sigma_l\, C^l_{il} W$ with $W \equiv \det[\,g(\xi_i,\xi_j)]$.

Proof:

(a) Let (Y_α) a basis of $T_x(M)$. As Lie derivative commutes with contraction, we have:

$$\begin{aligned}
L_{\xi_i}&(\omega_k, Y_\alpha)\\
&= (\xi_i(\omega_k), Y_\alpha) + (\omega_k, L_{\xi_i} Y_\alpha)\\
&= L_{\xi_i}(\,g(\xi_k, Y_\alpha))\ \text{(by definition of } \omega_k)\\
&= (L_{\xi_i} g)(\xi_k, Y_\alpha) + g(L_{\xi_i}\xi_k, Y_\alpha) + g(\xi_k, L_{\xi_i} Y_\alpha).
\end{aligned}$$

The first term of the third line vanishes because ξ_i is a Killing vector. By comparing the first and the third lines, it remains:

$$\begin{aligned}
(\xi_i(\omega_k), Y_\alpha) &= g(L_{\xi_i}\xi_k, Y_\alpha)\\
&= g\Big(\sum_l C^l_{ik}\xi_l, Y_\alpha\Big)\\
&= \Big(\sum_l C^l_{ik}\omega_l, Y_\alpha\Big),
\end{aligned}$$

which holds for any Y_α and thus proves the assertion;

(b) Since the Lie derivative satisfies Leibniz's rule, we have:

$$\begin{aligned}
\xi_i(\Omega) &= \xi_i(\omega_1\wedge\cdots\wedge\omega_p)\\
&= \sum_j \omega_1\wedge\cdots\wedge\xi_i(\omega_j)\wedge\cdots\wedge\omega_p\\
&= \sum_j \omega_1\wedge\cdots\wedge\sum_l C^l_{ij}\omega_l\wedge\cdots\wedge\omega_p\quad \text{(by the lemma)}\\
&= \sum_l C^l_{il}\Omega;
\end{aligned}$$

(c) We remark that $W = \Omega(\xi_1,\ldots,\xi_p)$.

Then:

$$\begin{aligned}
\xi_i(\Omega(\xi_1,\ldots,\xi_p)) &= \xi_i(W)\\
&= \xi_i(\Omega)(\xi_1,\ldots,\xi_p) + \Omega(L_{\xi_i}\xi_1,\xi_2,\ldots,\xi_p)\\
&\quad + \cdots + \Omega(\xi_1,\ldots,L_{\xi_i}\xi_p)\\
&= 2\sum_l C^l_{il} W. \qquad \square
\end{aligned}$$

We proceed now to the proof of the theorem.

(a) From the existence of a 2-parameter group of isometries, we deduce that there are two Killing vectors ξ_t,ξ_φ such that $[\xi_t,\xi_\varphi] = C^t_{t\varphi}\xi_t + C^\varphi_{t\varphi}\xi_\varphi$ where the coefficients C^i_{jk} are constant. From the lemma, we deduce that

$$\begin{aligned}
\xi_\varphi(W) &= 2\Big(\sum_l C^l_{\varphi l}\Big)W\\
&= 2C^t_{\varphi t} W,
\end{aligned}$$

where $W = \det(\,g(\xi_i,\xi_j))$. (1)

If $W \neq 0$, (1) can be written, in a coordinate system where $\xi_\varphi = \partial/\partial\varphi$,

$$\frac{\partial \ln W}{\partial\varphi} = 2C^t_{\varphi t}.$$

Thus $\ln W = 2C^t_{\varphi t}\cdot\varphi + \mathscr{f}$, where the function $\mathscr{f}$ is independent of φ. As Π^c: $SO(2)\times M \to M$, $\ln W$ must be periodic in φ. Thus $C^t_{\varphi t} = 0$ where $W \neq 0$.[4]

Since all C^i_{jk} are constant, it follows that $C^t_{\varphi t} = 0$.

(b) We shall prove that $C^\varphi_{\varphi t} = 0$

$$\begin{aligned}
\xi_\varphi&(\,g(\xi_t,\xi_\varphi))\\
&= g([\xi_\varphi,\xi_t],\xi_\varphi)\ \text{because } \xi_\varphi \text{ is a Killing vector}\\
&= C^\varphi_{t\varphi} g_{\varphi\varphi} + C^t_{\varphi t}\, g_{t\varphi}\\
&= C^\varphi_{t\varphi}\, g_{\varphi\varphi}
\end{aligned}$$

because of the first part of the proof. Thus $g(\xi_t,\xi_\varphi) = C^\varphi_{\varphi t}\, g_{\varphi\varphi}\varphi + h$ where h is independent of φ. To have the periodicity for $g(\xi_t,\xi_\varphi)$ we need.

$$C^\varphi_{\varphi t}\, g_{\varphi\varphi} = 0.$$

Since by hypothesis $g_{\varphi\varphi} \neq 0$ at at least a point P, $C^\varphi_{\varphi t} = 0$. $\square$

ACKNOWLEDGMENTS

The author would like to thank Professor D. Speiser who read the manuscript. He also wishes to thank B. Carter and R.G. Mc. Lenaghan for helpful discussions.

[1]B. Carter, Commun. Math. Phys. **17**, 233 (1979).
[2]Y. Choquet-Bruhat, C. Dewitt-Morette, N. Dillard-Bleick, *Analysis on Manifolds and Physics* (North-Holland, Amsterdam, 1977).
[3]J. Dieudonné, *Eléments d'analyse*, Tomes 3 et 4 (Gauthier-Villars, Paris, 1974).
[4]Here we have to suppose that $W \neq 0$ at least at one point of M.

 0022-2488/80/050954-01$01.00

On the irreducible representations of groups containing a subgroup of finite index

D. P. L. Castrigiano and U. Mutze
Sektion Physik der Universität München, D-8000 München 2, Theresienstrasse 37, Federal Republic of Germany

(Received 25 September 1979; accepted for publication 14 January 1980)

The objects under consideration are: A group G containing a subgroup S of finite index p, an irreducible representation (= multiplier representation by unitary or by unitary and antiunitary operators on a Hilbert space of arbitrary dimension) U of G, and an irreducible representation W of S. It is shown (1) that the representations $U|S$ (the restriction of U to S) and $W\uparrow G$ (the representation induced by W) are both orthogonal sums of finitely many irreducible subrepresentations, the number of which does not exceed p; (2) that the multiplicity of W in $U|S$ equals the multiplicity of U in $W\uparrow G$ if W and U are unitary representations and that these multiplicities are related in a slightly different manner for partially antiunitary representations. For the special case that S is an invariant subgroup, it is shown how the irreducible representations of G can be constructed if the irreducible representations of S and those of certain finite groups are known.

1. INTRODUCTION

Let us first indicate the physical origin of the mathematical problem that is the subject of this paper. The study of the reflection operators T, PC, and PCT in relativistic quantum (field) theory requires some information on the irreducible representations of the full Poincaré group (i.e., the group generated by the Poincaré transformations, the reversal of time direction, and any reflection at a plane). These representations were classified by many authors. However, at the very basis of all such investigations we know about, there enters an unproven hypothesis that may, in its weakest form, be stated as follows.

(i) Every representation of the proper orthochronous Poincaré group that is obtained by restriction from an irreducible representation of the full Poincaré group contains an irreducible subrepresentations.

Note that group representations on infinite dimensional spaces need not have irreducible subrepresentations: Choose, for instance, any representation of the Poincaré group with nonzero energy and restrict it to the translation group. Nevertheless, the truth of (i) is plausible: If it were false, there would exist an infinity of subspaces, all being invariant under the uncountable manifold of Poincaré transformations, whereas none of these subspaces would be invariant under time reversal and space reflection, that is, under a set of only two transformations. If this crude argument is to the point, it also supports the truth of a more general hypothesis.

(ii) is obtained from (i) by replacing the full Poincaré group by any group G and the proper orthochronous Poincaré group by any subgroup S of finite index. (The index of a subgroup S is the cardinality $|G/S|$ of the quotient space G/S.) G may then, for instance, be any Lie group with a finite number of connected components and S the unit component. (Then S is an invariant subgroup.) To prove the hypothesis (ii) is the main purpose of this paper. In fact, we will prove a little more, namely that the restricted representation decomposes into finitely many irreducible representations of S, the number of which does not exceed the index of S (Theorem 1). Under the special assumption that S is of index 2, this was proved earlier in Ref. 1 by methods that suggest no generalization to higher indices. Due to the particular structure of the full Poincaré group the index 2 case is in fact sufficient to establish (i). (The proper orthochronous Poincaré group is of index 2 in the orthochronous Poincaré group and the latter group is of index 2 in the full Poincaré group.)

Let us now say something about the mathematical framework we will work in (for details, see Sec. 2). Since our results should be directly applicable to the physically relevant representations of the full Poincaré group, we have to consider multiplier representations by unitary operators or by unitary and antiunitary operators. It will become apparent that taking into account multipliers and antiunitary operators causes only minor complications. The reader who is interested only in proper unitary representations should ignore all assumptions and statements concerning objects denoted by G_+, ω, δ, and K. No topological or measure theoretical properties of groups or representations will be required.

The main part of the paper is divided into two sections of rather different character. Sec. 3 contains our main result, Theorem 1, the content of which was indicated earlier in this introduction, and some further theorems, concerning the relation between the irreducible representations of G and certain representations of G that are induced by irreducible representations of a subgroup S of finite index. These theorems establish a generalized version of Frobenius' reciprocity theorem, and a procedure by which one can construct the irreducible representations of G if those of S are known. In Sec. 4 we assume that G contains an invariant subgroup of finite index. Then, a more practicable procedure for deduc-

 0022-2488/80/050955-06$01.00

ing the irreducible representations of G from those of the subgroup is available. The pertinent facts are well known for finite dimensional linear or partially antilinear representations (see Refs. 2–4), and for infinite dimensional unitary representations under additional topological assumptions (see Ref. 5). With Theorem 1 and the dimension-independent version of Schur's Lemma (see Sec. 2) at one's disposal, one can easily translated the proofs referring to the finite-dimensional case into general ones. Therefore, we have only stated the results and given no proofs in Sec. 4.

2. PRELIMINARIES

For reasons that are indicated in the introduction, we will consider group representations that are slightly more general than unitary representations. To be definite, we have to fix some terminology concerning these representations.

Let G be a group (no topology implied) and let H be a Hilbert space ($\equiv$ complex Hilbert space of arbitrary dimension). Suppose that U maps G to a set of operators on H such that:

(i) for any $g\in G$, the operator ($\equiv$ linear or antilinear, bounded, and everywhere defined, operator) $U(g)$ is either unitary or antiunitary;

(ii) any of the operators $U(g)U(g')(U(gg'))^{-1}$ is a complex multiple of the identity operator 1;

(iii) $U(e) = 1$, where e is the unit in G.

Then U will simply be called a *representation of* G, and H will be called the *carrier space of* U. The subset $G_+ := \{g\in G: U(g) \text{ is unitary}\}$ is an invariant subgroup of index 1 or 2 and will be called the *unitary domain of* U (or the unitary subgroup of G). The mapping $\omega: G\times G\to\mathbb{C}$ satisfying $U(g)U(g') = \omega(g,g')$ $\times U(gg')$ is called the *multiplier of* U. A function $\omega: G\times G\to\mathbb{C}$ is a multiplier of a representation of G with unitary domain G_+ if and only if

$$|\omega(g,g')| = 1, \quad \omega(e,e) = 1,$$
$$\omega(g,g'g'')\delta_g(\omega(g',g'')) = \omega(gg',g'')\omega(g,g'), \tag{2.1}$$

where, for any complex number z, the number $\delta_g(z)$ is defined to be z for $g\in G_+$ and $\bar{z}$ for $g\notin G_+$. Therefore, any solution ω of (2.1) is called a *G_+-multiplier of* G. (The if part of the preceding statement follows from the fact that the inducing formula (3.2), applied to the trivial representation of the subgroup $S = \{e\}$, gives an ω-representation of G, when the carrier space is taken to be the Hilbert space of those complex-valued functions f on G that differ from zero only on a denumerable set of points and satisfy $\Sigma_{g\in G}|f(g)|^2 < \infty$.)

We now fix once and for all a group G, a subgroup G_+ of G with $G = G_+$ or $|G/G_+| = 2$, and a G_+-multipler ω of G. We will be concerned mainly with those representations of G that have unitary domain G_+ and multiplier ω. When such a representation is restricted to a subgroup S of G it has unitary domain $S\cap G_+$ and multiplier $\omega|S\times S$. Therefore, we will employ in the sequel the following

Convention: Given a subgroup $S\subseteq G$, the term "representation of S" means "representation of S with unitary domain $S\cap G_+$." Further, the multiplier of a representation of S is $\omega|S\times S$ unless a different specification is noted.

Since the identity operator is linear, the usual "trivial representation" is a representation with trivial multiplier of G only if $G = G_+$. If we have $|G/G_+| = 2$, we obtain a similarly simple representation with trivial multiplier of G by choosing a conjugation operator C (i.e., C is antiunitary, and $C^2 = 1$) and associating the identity operator with any $g\in G_+$, and C with any $g\notin G_+$. Any representation obtained in this way will be called a *trivial representation of* G. Equivalence of representations, denoted by $\cong$, means always unitary equivalence. The equivalence class of a representation U is denoted by $[U]$.

A representation U of S is irreducible if and only if the real multiples of the identity operator are the only Hermitian linear operators that commute with $U(S)$ (Ref. 1, cf. Ref. 6). From this fact, we easily derive two important lemmas that are well known for unitary representations: (a) a linear operator that intertwines two irreducible representations of S is either zero or a real multiple of a unitary operator; (b) if a representation U is an orthogonal sum of irreducible subrepresentations, the subspaces that carry the primary components (i.e., the maximal sums of mutually equivalent subrepresentations of U) are uniquely determined by U. The set of equivalence classes of irreducible representations of S is denoted by $\hat{S}$.

3. RESTRICTION OF AN IRREDUCIBLE REPRESENTATION TO A SUBGROUP OF FINITE INDEX

Throughout this section, S will denote a subgroup of G such that $|G/S| = p < \infty$, $r_1,\dots,r_p$ will denote a system of coset representatives (i.e., $G = \cup_{i=1}^{p} r_i S$) chosen such that $r_1 = e$, and gi (for $g\in G$, $i\in\{1,\dots,p\}$) will denote the number defined by

$$gr_i S = r_{gi} S. \tag{3.1}$$

As is well known, the correspondence $(g,i)\mapsto gi$ defines a transitive action of G on the set $\{1,\dots,p\}$ (S being the stationary subgroup of the point 1). Conversely, a transitive action of G on a set of p points exists only if G possesses a subgroup of index p.

Our main result is

Theorem 1: Let U be an irreducible representation of G. Then the restriction $U|S$ of U to S is an orthogonal sum of at most p irreducible subrepresentations.

Proof: Let A be a linear operator that commutes with $U(S)$. We claim that the following linear operator commutes with $U(G)$:

$$\phi(A) := \frac{1}{p}\sum_{i=1}^{p} U(r_i)AU(r_i)^{-1}.$$

To prove this claim, we compute

$$\begin{aligned} U(g)U(r_i) &= \omega(g,r_i)U(gr_i) = \omega(g,r_i)U(r_{gi}r_{gi}^{-1}gr_i) \\ &= \omega(g,r_i)/\omega(r_{gi},r_{gi}^{-1}gr_i)U(r_{gi})U(r_{gi}^{-1}gr_i) \\ &=: \alpha_{g,i}U(r_{gi})U(s_{g,i}), \end{aligned}$$

where $\alpha_{g,i}$ is a complex number, and $s_{g,i}$ belongs to S [see (3.1)]. Therefore, and since the operator $U(r_{gi})AU(r_{gi})^{-1}$ is linear, and since the mapping $i\mapsto gi$ is a permutation, we have

$$U(g)\phi(A)U(g)^{-1} = \frac{1}{p}\sum_{i=1}^{p} U(g)U(r_i)A\,(U(g)U(r_i))^{-1}$$
$$= \frac{1}{p}\sum_{i=1}^{p} \alpha_{g,i}U(r_{gi})U(s_{g,i})AU(s_{g,i})^{-1}U(r_{gi})^{-1}\alpha_{g,i}^{-1}$$
$$= \frac{1}{p}\sum_{i=1}^{p} U(r_i)AU(r_i)^{-1} = \phi(A).$$

The following useful properties of ϕ can be simply verified:

(i) $\phi(\alpha A+\beta B) = \alpha\phi(A)+\beta\phi(B)$ for all real numbers α,β;

(ii) $\phi(A^*) = \phi(A)^*$;

(iii) $\phi(A)$ is a positive operator if A is so;

(iv) $\phi(1) = 1$.

We now use that U is irreducible. For an Hermitian linear operator A that commutes with $U(S)$, the operator $\phi(A)$ is Hermitian [by (ii)] and commutes with $U(G)$. Thus $\phi(A)$ is a real multiple, say $\varphi(A)1$, of the identity operator. Suppose, in addition, that A is positive. Then, for any normalized vector x in the carrier space of U, we have

$$\varphi(A) = \langle x|\phi(A)x\rangle = (1/p)(\langle x|Ax\rangle + \sum_{i=2}^{p}\langle U(r_i)^{-1}x|AU(r_i)^{-1}x\rangle) \geqslant (1/p)\langle x|Ax\rangle.$$

This shows

(v) $\varphi(A) \geqslant (1/p)\|A\|$.

Assume now that the theorem is false. Then there exist q non-zero Hermitian projection operators $Q_1,\ldots,Q_q$ with $q>p$, all commuting with $U(S)$, and satisfying $Q_1+\cdots+Q_q = 1$. [Obviously, a subspace X is invariant under $U(S)$ if and only if the Hermitian projection operator associated with X commutes with $U(S)$.] Combining (iv), (i), and (v) we have

$$1 = \varphi(1) = \varphi(Q_1+\cdots+Q_q) = \varphi(Q_1)+\cdots+\varphi(Q_q) \geqslant (1/p)(\|Q_1\|+\cdots+\|Q_q\|) = (q/p),$$

that is $p \geqslant q$, in contradiction to our assumption $q>p$. Therefore, the theorem is true. □

Given an irreducible representation U of G, we know that it determines certain irreducible representations of S, namely the irreducible subrepresentations of $U|S$. Changing our point of view, we are led to an interesting question: Let an irreducible representation W of S be given. To what extend an irreducible representation U of G is determined by W being a subrepresentation of $U|S$? We shall pursue this question throughout the remainder of this section. A partial answer will be, that U is equivalent to a subrepresentation of an induced representation $W\uparrow G$. Therefore, we shall need some facts on induced representations, which will be described now.

Lemma 1: Let X be a Hilbert space that carriers both a representation W of S and a trivial representation K of G (see Sec. 2). Then we have:

(a) The formula

$$(V(g)x)_i = (\omega(g,r_{g^{-1}i})/\omega(r_i,r_i^{-1}gr_{g^{-1}i})) \times K(r_i)W(r_i^{-1}gr_{g^{-1}i})K(r_{g^{-1}i})x_{g^{-1}i}, \quad (3.2)$$

defines a representation V of G on the carrier space X^p (i.e., on the Hilbert space of X-valued p-tupels).

(b) The projection operators $P_i: X^p \to X^p$, $(P_ix)_j := \delta_{ij}x_j$ satisfy

$$V(g)P_iV(g)^{-1} = P_{gi}, \quad \sum_{i=1}^{p} P_i = 1, \quad P_iP_j = \delta_{ij}P_j.$$

The set $V(G)\cup\{P_i: 1\leqslant i\leqslant p\}$ of operators is irreducible if and only if W is irreducible.

(c) Suppose that U is a representation of G such that $U|S$ contains W as a subrepresentation. Denote the carrier space of U by H. (Then X is a subspace of H.) Suppose that the subspaces $U(r_i)X$, $i\in\{1,\ldots,p\}$, are mutually orthogonal and span the space H. The the mapping

$$T: X^p \to H, \quad Tx := \sum_{i=1}^{p} U(r_i)K(r_i)x_i,$$

is unitary and the representation $T^{-1}UT$ coincides with V in (3.2).

Equation (3.2) is Shaw's and Lever's prescription for "generalized inducing" (see Ref. 4). Part (c) is analogous to Mackey's imprimitivity theorem (see Ref. 5, Theorem 6.6). It shows that Eq. (3.2), though it might appear strange at a first glance, has a very natural origin. Since the proof of Lemma 1 is straightforward and purely computational, we omit it.

Definition: Let W be as in Lemma 1. Then any representation V of G given by Eq. (3.2), with any choice of a trivial representation K and of coset representatives $r_1,\cdots,r_p$, is said to be induced by W, and is denoted by $W\uparrow G$.

It is easily shown that all representations induced by W are equivalent, and that $W\cong W'$, implies $W\uparrow G\cong W'\uparrow G$. Even if W is irreducible, we cannot expect that $W\uparrow G$ is irreducible. On the subrepresentations of $W\uparrow G$ we have

Theorem 2: Let W and $V(=W\uparrow G)$ be as in Lemma 1. Suppose, in addition, that W is irreducible. Then $W\uparrow G$ is an orthogonal sum of at most p irreducible subrepresentations.

Remark: Generalizing the hypothesis "$|G/S|<\infty$" into "G/S compact" invalidates the conclusions of Theorems 1 and 2: If we take for G the Euclidean group of motions and for S the subgroup of translations, neither $U|S$ nor $W\uparrow G$ will, in general, contain any irreducible subrepresentation.

Proof of Theorem 2: Let A be a linear operator that commutes with $V(G)$. We show that

$$\phi(A) := \sum_{i=1}^{p} P_iAP_i,$$

commutes with the irreducible set considered in part (b) of Lemma 1:

$$V(g)\phi(A) = \sum_{i=1}^{p} V(g)P_iAP_i$$
$$= \sum_{i=1}^{p} P_{gi}V(g)AP_i$$
$$= \sum_{i=1}^{p} P_{gi}AP_{gi}V(g)$$
$$= \sum_{i=1}^{p} \phi(A)V(g),$$

and

$$P_j\phi(A) = P_jAP_j = \phi(A)P_j.$$

Hence $\phi(A) = \varphi(A)1$ with a real number $\varphi(A)$ if A is Hermitian. If A is positive, we have for any normalized $x \in X^p$

$$\begin{aligned}\varphi(A) &= \langle x|\phi(A)x\rangle \\ &= \sum_{i=1}^{p} \langle P_ix|AP_ix\rangle \\ &= \sum_{i=1}^{p} \|A^{1/2}P_ix\|^2 \\ &\geqslant \frac{1}{p}\Big(\sum_{i=1}^{p} \|A^{1/2}P_ix\|\Big)^2 \\ &\geqslant \frac{1}{p}\Big\|\sum_{i=1}^{p} A^{1/2}P_i\Big\|^2 \\ &= \frac{1}{p}\langle x|Ax\rangle.\end{aligned}$$

Now, the proof becomes complete when we take over the final part of the proof of Theorem 2 [starting with (v)], with $U(S)$ replaced by $V(G)$. □

Let us now consider irreducible representations W and U of S and G respectively. Then $W\uparrow G$ and $U|S$ are representations of G and S, respectively. Since, by Theorems 2 and 1, the last two representations are orthogonal sums of finitely many irreducible representations, we may ask how many times U appears (up to equivalence) in the decomposition of $W\uparrow G$ and how many times W appears in the decomposition of $U|S$. It is a classical result on linear representations of finite groups that, there, these two multiplicities are equal (Frobenius' reciprocity theorem). Theorem 4 will establish this reciprocity also in our case as far as linear representations are concerned; for partially antilinear representations the reciprocity law will have to be slightly modified. The next theorem states an algebraic fact that is the core of the reciprocity theorem.

Theorem 3: Let W and $V(=W\uparrow G)$ be as in Lemma 1, and let U be a representation of G with carrier space H. Define the following sets of linear intertwining operators

$$\mathrm{Int}(W\uparrow G,U) = \{A{:}X^p\to H{:}AV(g) = U(g)A,\ \forall g\in G\},$$

$$\mathrm{Int}(W,U|S) = \{B{:}X\to H{:}BW(s) = U(s)B,\ \forall s\in S\},$$

and the mapping

$$F{:}X\to X^p,\ x\to(x,0,\ldots,0).$$

Then

$$\Psi{:}\mathrm{Int}(W\uparrow G,U)\to\mathrm{Int}(W,U|S),\ A\to AF,$$

is a bijective mapping; it is linear with respect to the real-linear structure of the Int-spaces and its inverse is given by

$$\Psi^{-1}(B) = \sum_{i=1}^{p} U(r_i)BK(r_i)E_i,$$

where E_i is the projection $X^p\to X$, $x\to x_i$.

Proof: Let A belong to $\mathrm{Int}(V,U)$. We show that AF belongs to $\mathrm{Int}(W,U|S)$: From (3.2) one easily infers $V(s)F = FW(s)$ for all $s\in S$, and hence, $AFW(s) = AV(s)F = U(s)AF$. Now, let B belong to $\mathrm{Int}(W,U|S)$. We show that $A{:}=\Sigma_{i=1}^{p} U(r_i)BK(r_i)E_i$ belongs to $\mathrm{Int}(V,W)$:

$$\begin{aligned}AV(g)x &= \sum_{i=1}^{p} U(r_i)BK(r_i)(V(g)x)_i \\ &= \sum_{i=1}^{p} (\omega(g,r_{g^{-1}i})/\omega(r_i,r_i^{-1}gr_{g^{-1}i})) \\ &\quad\times U(r_i)BW(r_i^{-1}gr_{g^{-1}i})K(r_{g^{-1}i})x_{g^{-1}i} \\ &= \sum_{i=1}^{p} (\omega(g,r_{g^{-1}i})/\omega(r_i,r_i^{-1}gr_{g^{-1}i}))U(r_i)U(r_i^{-1}gr_{g^{-1}i}) \\ &\quad\times BK(r_{g^{-1}i})x_{g^{-1}i} \\ &= \sum_{i=1}^{p} U(g)U(r_{g^{-1}i})BK(r_{g^{-1}i})x_{g^{-1}i} \\ &= U(g)\sum_{i=1}^{p} U(r_i)BK(r_i)x_i \\ &= U(g)Ax.\end{aligned}$$

Finally, we show that Ψ^{-1}, as defined in the theorem, is a right inverse of $\Psi{:}\Psi(\Psi^{-1}(B)) = \Sigma_{i=1}^{p} U(r_i)BK(r_i)E_iF = U(r_1)BK(r_1) = B$, and also a left inverse of $\Psi{:}\Psi^{-1}(\Psi(A)) = \Sigma_{i=1}^{p} U(r_i)AFK(r_i)E_i = A\,\Sigma_{i=1}^{p} V(r_i)FK(r_i)E_i = A$. To justify the last step we compute [occasionally writing $r(\alpha)$ for r_α and $x(i)$ for x_i] $\Sigma_{i=1}^{p}(V(r_i)FK(r_i)x_i)_j = \Sigma_{i=1}^{p}[\omega(r_i,r(r_i^{-1}j))/\omega(r_j,r_j^{-1}r_ir(r_i^{-1}j))]K(r_j)W(r_j^{-1}r_ir(r_i^{-1}j))K(r(r_i^{-1}j))\times(FK(r_i)x_i)(r_i^{-1}j) = K(r_j)W(r_1)K(r_1)K(r_j)x_j = x_j$, since $(FK(r_i)x_i)(r_i^{-1}j)$ vanishes unless $r_i^{-1}j = 1$, i.e., $j = i$. □

For linear representations of finite groups, the assertion of Theorem 3 appears in Ref. 7, p. 185, Problem 10. The next theorem is a straightforward application of Theorem 3 to irreducible representations U and W.

Theorem 4: Let U and W be irreducible representations of G and S respectively. Let $W\uparrow G$ be an induced representation. Then the multiplicity $\mu(W,U|S)$ of W in $U|S$ (i.e., the number of times that a representation equivalent to W occurs within a decomposition of $U|S$ into irreducible subrepresentations) is related to the multiplicity $\mu(U,W\uparrow G)$ of U in $W\uparrow G$ by

$$\delta(W)\mu(W,U|S) = \delta(U)\mu(U,W\uparrow G),$$

where the function δ assigns to an irreducible representation the dimension of the real-linear space of those linear operators that commute with this representation. The possible values of δ are 1, 2, and 4.

Remarks: (a) If W and U are linear representations (i.e., $S\subseteq G_+$) we have $\delta(W) = \delta(U) = 2$. (Then, the intertwining operators form not only a real-linear space but also a complex one.) Then Theorem 4 states

$$\mu(W,U|S) = \mu(U,W\uparrow G),$$

i.e., the usual reciprocity law of Frobenius.

(b) For mixed unitary antiunitary representations, the situation $\delta(W)\neq\delta(U)$ actually occurs, as can be seen from the following example. Take for S and G the proper orthochronous and the full Poincaré group. There are irreducible representations U of G that are irreducible as representations of S and whose unitary domain is S. For $W = U|S$ we then have $\delta(W) = 2$ and $\delta(U) = 1$.

Proof of Theorem 4: One easily shows that the dimensions of the real-linear spaces $\mathrm{Int}(W,U|S)$ and $\mathrm{Int}(U,W\uparrow G)$ are $\delta(W)\mu(W,U|S)$ and $\delta(U)\mu(U,W\uparrow G)$ respectively. These dimensions are equal by Theorem 3 (note that the correspondence $A \to A^*$ establishes an isomorphism between $\mathrm{Int}(U,W\uparrow G)$ and $\mathrm{Int}(W\uparrow G,U)$. We now consider the space $\mathrm{Int}(W,W)$, the dimension of which is $\delta(W)$. If W is unitary, $\mathrm{Int}(W,W)$ consists of the complex multiples of the identity operator, hence $\delta(W)=2$. If W is partially antiunitary, $S\cap G_+$ is a subgroup of index 2 in S, and, by Theorem 1, $W|S\cap G_+$ is either irreducible or an orthogonal sum of two irreducible subrepresentations. In the first case, $\mathrm{Int}(W,W)$ consists of the real multiples of the identity, hence $\delta(W)=1$. In the second case one easily finds $\delta(W)=4$ if the two subrepresentations are equivalent and $\delta(W)=2$ otherwise (cf. Ref. 4, Theorem C). Clearly, the same analysis is valid for $\delta(U)$. □

Some immediate consequences of Theorem 4 are put together in the next corollary. Part (c) of this corollary describes a procedure of constructing the irreducible representations of G from the irreducible representations of S. As one step in this construction one has to decompose an induced representation of G into its irreducible subrepresentations, which is a rather involved algebraic problem. When S is an invariant subgroup, one can avoid this step by the procedure described in Sec. 4. Part (d) of the corollary states that an irreducible representation of G has always a generalized system of imprimitivity based on G/S where the generalization consists of replacing projection operators by positive operators (cf. Refs. 8, 9 for the connection between such generalized systems of imprimitivity and usual ones).

Corollary: (a) Let U be an irreducible representation of G. Then there is an irreducible representation W of S such that U is equivalent to a subrepresentation of $W\uparrow G$.

(b) Let W be an irreducible representation of S. Then any subrepresentation U of $W\uparrow G$ has the property that W is equivalent to a subrepresentation of $U|S$.

(c) Let W be an irreducible representation of S. By decomposing $W\uparrow G$ into irreducible subrepresentations one obtains (up to equivalence all such irreducible representations U of G that satisfy $\mu(W,U|S)\neq 0$, and no others (but, generally, some of these representations more than once). There are at most p such representations. By varying W in that procedure, we obtain all irreducible representations of G.

(d) For any irreducible representation U of G there are positive operators $T_1,\ldots,T_p$ such that

$$U(g)T_iU(g)^{-1}=T_{gi},\quad \text{for all } g\in G,\ \sum_{i=1}^{p}T_i=1.$$

Proof: (a), (b): By Theorem 4 we have $\mu(W,U|S)\neq 0$ if and only if $\mu(U,W\uparrow G)\neq 0$. (c) is clear from (a) and (b). (d): By (a) it is sufficient to consider the case that U is a subrepresentation of an induced representation. Let P and I be, respectively, the projection and injection that connect the carrier spaces of U and of the induced representation. Then the operators $T_i := PP_iI$, with P_i from Lemma 1(b), have the required properties. □

4. ON THE IRREDUCIBLE REPRESENTATIONS OF GROUPS THAT CONTAIN AN INVARIANT SUBGROUP OF FINITE INDEX

For later use, the following lemma introduces a rule of transforming the representations of subgroups of G. Part (b) of the lemma shows the origin of this rule. For purely unitary representations the rule appears in Ref. 5, p. 277, and for partly antiunitary representations in Ref. 4.

Lemma 2: Let S be a subgroup of G. Let X by a Hilbert space that carries both a representation W of S and a trivial representation K of G. Then:

(a) For any $g\in G$, the formula

$$W^g(h) := (\omega(h,g^{-1})/\omega(g^{-1},ghg^{-1}))K(g)W(ghg^{-1})K(g), \quad (4.1)$$

defines a representation W^g of the group $g^{-1}Sg$. Up to equivalence, W^g does not depend on the choice of K, and $W\cong W'$ implies $W^g\cong W'^g$. For all $g,g'\in G$ we have $(W^g)^{g'} = T_{g,g'}^{-1}W^{gg'}T_{g,g'}$, where $T_{g,g'}=\omega(g'^{-1},g^{-1})1$. If S is an invariant subgroup, the correspondence $(g,[W])\to[W^g]$ defines an action of G on $\widehat{S}$.

(b) Suppose that U is a representation of G with carrier space H such that W is a subrepresentation of $U|S$ (then X is a subspace of H). Then, for any $g\in G$, the subspace $U(g^{-1})K(g)X$ carries a subrepresentation of $U|g^{-1}Sg$ that is equivalent [by means of the operator $U(g^{-1})K(g)$] to the representation W^g in (a).

The structure of the restriction of an irreducible representation of G to an invariant subgroup is described in

Theorem 5: Let N be an invariant subgroup of G such that $|G/N|=p<\infty$. Let U be an irreducible representation of G. Consider a decomposition of $U|N$ into irreducible subrepresentations, and group equivalent subrepresentations together, thus obtaining a decomposition

$$U|N=\bigoplus_{i=1}^{q}\bigoplus_{j=1}^{p_i}U_{ij},$$

where $U_{ij}\cong U_{lk}$ if and only if $i=l$, and $p_1+\cdots+p_q\leqslant p$. Then we have:

(a) The carrier space H_i of the representation $U_i := \oplus_{j=1}^{p_i}U_{ij}$ is uniquely determined by the class $[U_i]$: It does not depend on the particular decomposition of $U|N$ into irreducibles.

(b) The multiplicities p_i are all equal.

(c) For any two subrepresentations U_{ij} and U_{lk}, there is a $g\in G$ such that $U_{ij}^g\cong U_{lk}$.

(d) For any $i\in\{1,\ldots,q\}$, define $S_i := \{g\in G: U_{i1}^g\cong U_{i1}\}$. Then any subspace H_i [see (a)] carries an irreducible subrepresentations W_i of $U|S_i$ (which is obviously equivalent to p_iU_{i1} when restricted to N).

(e) The correspondence $(g,H_i)\mapsto U(g)H_i$ defines a transitive action of G on the set $\{H_1,\ldots,H_q\}$.

Note that (e), together with part (c) of Lemma 1, shows that U is equivalent to $W_i\uparrow G$. The structure of the representations W_i in (d) is analyzed in the next theorem.

Theorem 6: Let N be an invariant subgroup of G such that $|G/N|<\infty$ and $N\subseteq G_+$. Let V be an irreducible representation of N. Suppose that W is an irreducible representation of $S: = \{g\in G: V^g \cong V\} \supseteq N$ such that

(i) $W(n) = V(n) \oplus \cdots \oplus V(n) = V(n) \otimes 1$, for all $n\in N$, where the sum consists of p terms, and 1 is the identity operator in the Hilbert space $\mathbb{C}^p$. Choose coset representatives $r_1,\ldots,r_q$ in S/N such that $r_1 = e$. For any $i\in\{1,\ldots,q\}$ choose an operator $T(r_j)$ such that $T(r_1) = 1$ and $T(r_j)$ is unitary (antiunitary) for $r_j\in G_+(r_j\notin G_+)$,

(ii) $$V(n)T(r_j) = (\omega(n,r_j)/\omega(r_j,r_j^{-1}nr_j)) \times T(r_j)V(r_j^{-1}nr_j), \quad \text{for all } n\in N.$$

(Since r_j belongs to S, this is always possible; and since V is irreducible, $T(r_j)$ is determined up to a phase factor). Define $T(s)$ for any $s\in S$ by the rule

(iii) $T(r_jn) = \omega(r_j,n)^{-1}T(r_j)V(n)$, for all $n\in N$. Then we have:

(a) The mapping T is a representation of S and the multiplier τ of T satisfies $\omega(r_in,r_jn')/\tau(r_in,r_jn') = \omega(r_i,r_j)/\tau(r_i,r_j)$ for all $n,n'\in N$. Therefore the rule $\rho(r_iN,r_jN): = \omega(r_i,r_j)/\tau(r_i,r_j)$ defines a multiplier ρ of the quotient S/N.

(b) For any $s\in S$ we have $W(s) = T(s)\otimes R(sN)$, where R is an irreducible representation of S/N with unitary domain $(S\cap G_+)/N$ and multiplier ρ from (a).

If $N\subseteq G_+$ is not assumed, the operators $(T(s)\otimes 1)^{-1}W(s)$ cannot be shown to commute with $W(N)$. Even if they would commute, we were not able to derive the form $W = T\otimes R$ in that case.

Finally, we shall answer the question how to obtain the irreducible representations of G if those of an invariant subgroup of finite index are known. Let M be an invariant subgroup of G with finite index. Then $N: = M\cap G_+$ is also invariant with finite index. We now assume that the irreducible representations of N are known. To obtain, up to eqivalence, all irreducible representations of G we proceed as follows. We consider the action of G on $\hat{N}$ that is described in part (a) of Lemma 2. From each orbit, we choose one point. For each such point, we choose a representative. For each chosen representative V, we determine the group $S = \{g\in G: V^g \cong V\}$ and construct the mapping T and the multiplier ρ of S/N as in Theorem 6. The next step is to determine the irreducible representations of the finite group S/N, restricting ourselves to those representations with unitary domain $(S\cap G_+)/N$ and multiplier ρ (note that ρ will, generally, not be trivial even if ω is so). From each equivalence class of those representations, we choose a representative. For each such representative R, we construct an irreducible representation W of S by the rule $W(s) = T(s)\otimes R(sN)$. Finally, for this W, we form any induced representation $W\uparrow G$. By this procedure, we obtain just one representative from every element of $\hat{G}$.

ACKNOWLEDGMENT

The authors are indebted to Professor H.J. Meister for many useful discussions.

[1]U. Mutze, Commun. Math. Phys. **59**, 267 (1978).
[2]A.H. Clifford, Ann. Math. **38**, 535 (1937).
[3]H. Boerner, *Representations of Groups*, 2nd edition (North-Holland, Amsterdam, 1970).
[4]R. Shaw, and J. Lever, Commun. Math. Phys. **38**, 257 (1974).
[5]G.W. Mackey, Acta Math. **99**, 265 (1958).
[6]R. Ascoli and G. Teppati, Nuovo Cimento **65**, 229 (1970).
[7]A.A. Kirillov, *Elements of the Theory of Representations* (Springer, New York, 1976).
[8]D.P.L. Castrigiano, Lokalisierung relativistischer Elementarsysteme, Doctoral dissertation, Universität München, 1976.
[9]D.P.L. Castrigiano and R.W. Henrichs, "Systems of Covariance and Subrepresentations of Induced Representations," to appear in Lett. Math. Phys.

Clebsch–Gordan coefficients for corepresentations. I ⊗ I

R. Dirl
Institut für theoretische Physik, TU Wien A-1040 Wien, Karlsplatz 13, Austria

(Received 18 December 1978; accepted for publication 25 May 1979)

A general method is given for determining Clebsch–Gordan coefficients for corepresentations in terms of convenient Clebsch–Gordan coefficients for the normal subgroup, at which the considered Kronecker products are composed of corepresentations of type I only.

INTRODUCTION

An important application of group theory to physics is the problem of decomposing Kronecker products of unitary irreducible corepresentations (co-unirreps) into a direct sum of their irreducible constitutuents.[1–4] The reason for considering this problem arises from the fact that selection rules governing transitions in magnetic materials are determined by Clebsch–Gordan coefficients (CG coefficients) for corepresentations of antiunitary groups.

The aim of this series of papers is to compute CG- coefficients for corepresentations by means of a general method which has been extensively described in Ref. 5 and applied to various problems in Refs. 6–12. In contrary to the methods given in Refs. 1–4, we assume from the outset that convenient CG-coefficients for the normal subgroup are given. The reason for supposing this will be suggested not only by the special structure of the antiunitary group G, which contain a normal subgroup H of index two, but also by the special values for multiplicities, which specify how many times a given co-unirrep is contained into the considered Kronecker product of co-unirreps.

The crucial point of the present method consists of considering the columns of the unitary CG matrices for corepresentations as H-adapted vectors, i.e., vectors which transform according to unirreps of the subgroup H, but which have to satisfy additional transformation properties originating from a special representative of the antiunitary group elements. To suppose the CG-matrices for corepresentations as unitary implies no loss of generality and simplifies our considerations. This approach allows the derivation of simple defining equations for unitary transformations which link CG coefficients for corepresentations with convenient CG coefficients for the normal subgroup H. Since the structure of these defining equations depends on the three different types of co-unirreps, we consider them separately. Apart from two cases (contained in this paper) we solve these defining equations quite generally without reference to a special group. These solutions allow one to identify the multiplicity index in a very special way.

The material is organized as follows: In Sec. I we summarize the basic definitions and notations concerning co-unirreps, which are used throughout this and the following papers. The multiplicities, which specify how many times a given co-unirreps is contained in a Kronecker product whose constituents are of type I only, are written down in Sec. II. Section III is devided into three parts depending on the different types of co-unirreps. For each case a simple defining equation for those unitary matrices is derived, which link CG coefficients for corepresentations with CG coefficients for the normal subgroup H. Apart from the first two cases we give a general solution of our problem.

I. DEFINITIONS AND NOTATIONS

In this section we recall briefly the basic definitions and notations concerning corepresentations of finite groups, which are used throughout this and the following papers. Let

$$G = \{H, sH\} \tag{I.1}$$

be the coset decomposition of the finite group G with respect to the normal subgroup H. Unitary corepresentations of G are matrix representations satisfying[13–15]

$$\mathbb{R}^*(s^{-1}hs) = \mathbb{R}(s)^\dagger \mathbb{R}(h)\mathbb{R}(s), \quad \text{for all } h\in H, \tag{I.2}$$

$$\mathbb{R}(s)\mathbb{R}^*(s) = \mathbb{R}(s^2), \quad \text{with } s^2\in H, \tag{I.3}$$

$$\mathbb{R}(sh) = \mathbb{R}(s)\mathbb{R}^*(h), \quad \text{for all } h\in H, \tag{I.4}$$

where

$$\mathbb{R} = \{\mathbb{R}(h): h\in H\}, \tag{I.5}$$

is an ordinary vector representation of H. Furthermore, unitary corepresentations are equivalent if and only if

$$W^\dagger \mathbb{R}(g) W^g = \mathbb{R}'(g), \quad \text{for all } g\in G, \tag{I.6}$$

$$W^g = \begin{cases} W, & \text{for } g\in H, \\ W^*, & \text{for } g\in sH, \end{cases} \tag{I.7}$$

are satisfied, where W denotes a unitary matrix.

Concerning the co-unirreps of a given finite group G, one distinguishes three different types.

Type I:

$$\mathbb{R}^\alpha(h) = R^\alpha(h), \quad \text{for all } h\in H, \tag{I.8}$$

$$\mathbb{R}^\alpha(s) = U^\alpha, \tag{I.9}$$

$$U^\alpha U^{\alpha *} = R^\alpha(s^2), \tag{I.10}$$

$$\mathbb{R}^{\alpha *}(s^{-1}hs) = U^{\alpha\dagger} R^\alpha(h) U^\alpha, \quad \text{for all } h\in H; \tag{I.11}$$

Type II:

$$\mathbb{R}^\beta(h) = \begin{bmatrix} R^\beta(h) & 0 \\ 0 & R^\beta(h) \end{bmatrix}, \quad \text{for all } h\in H, \tag{I.12}$$

$$\mathbb{R}^\beta(s) = \begin{bmatrix} 0 & U^\beta \\ -U^\beta & 0 \end{bmatrix}, \tag{I.13}$$

$$U^\beta U^{\beta *} = -R^\beta(s^2), \tag{I.14}$$

$$R^{\beta *}(s^{-1}hs) = U^{\beta\dagger} R^\beta(h) U^\beta, \quad \text{for all } h\in H; \tag{I.15}$$

 0022-2488/80/050961-07$01.00

Type III:

$$\mathbb{R}^{\gamma}(h) = \begin{bmatrix} R^{\gamma}(h) & 0 \\ 0 & Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma} \end{bmatrix}, \quad \text{for all } h \in H, \tag{I.16}$$

$$\mathbb{R}^{\gamma}(s) = \begin{bmatrix} 0 & R^{\gamma}(s^2) \\ \mathbb{1}_{\gamma} & 0 \end{bmatrix}, \tag{I.17}$$

$$R^{\gamma *}(s^{-1}hs) = Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}, \quad \text{for all } h \in H. \tag{I.18}$$

Consequently, the matrix elements of the co-unirreps can be written as

Type I:

$$\mathbb{R}^{\alpha}_{ij}(h) = R^{\alpha}_{ij}(h), \tag{I.19}$$

$$\mathbb{R}^{\alpha}_{ij}(s) = U^{\alpha}_{ij}, \quad i,j = 1,2,\ldots,n_{\alpha}; \tag{I.20}$$

Type II:

$$\mathbb{R}^{\beta}_{ai;bj}(h) = \delta_{ab}\, R^{\beta}_{ij}(h), \tag{I.21}$$

$$\mathbb{R}^{\beta}_{ai;bj}(s) = (-1)^{\Delta(a)} \delta_{a,b+1} U^{\beta}_{ij},$$

$$a,b = 1,2 \text{ and } i,j = 1,2,\ldots,n_{\beta},$$

$$\Delta(a) = \begin{cases} 0, & \text{for } a = 1, \\ 1, & \text{for } a = 2; \end{cases} \tag{I.22}$$

Type III:

$$\mathbb{R}^{\gamma}_{ai;bj}(h) = \begin{cases} R^{\gamma}_{ij}(h), & \text{for } a = b = 1, \\ R^{\gamma *}_{ij}(s^{-1}hs) = \{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}_{ij}, & \text{for } a = b = 2, \\ 0, & \text{otherwise,} \end{cases} \tag{I.23}$$

$$\mathbb{R}^{\gamma}_{ai;bj}(s) = \begin{cases} R^{\gamma}_{ij}(s^2), & \text{for } a = 1 \text{ and } b = 2, \\ \delta_{ij}, & \text{for } a = 2 \text{ and } b = 1, \\ 0, & \text{otherwise,} \end{cases} \tag{I.24}$$

$$a,b = 1,2 \quad \text{and} \quad i,j = 1,2,\ldots,n_{\gamma}.$$

The symbols α, β, and γ denote the elements of the set A_H consisting of all equivalence classes. These equivalence classes decompose into three disjoint subsets A_i, $i = \mathrm{I,II,III}$ with respect to the group elements of the supergroup G:

$$A_H = A_{\mathrm{I}} \cup A_{\mathrm{II}} \cup A_{\mathrm{III}}. \tag{I.25}$$

Thereby our notation should always imply $\alpha \in A_{\mathrm{I}}$, $\beta \in A_{\mathrm{II}}$, and $\gamma(\neq \bar{\gamma}) \in A_{\mathrm{III}}$.

II. MULTIPLICITIES FOR COREPRESENTATIONS

Within this paper we consider only Kronecker products of the kind

$$\mathbb{R}^{\alpha_1\alpha_2} = \{\mathbb{R}^{\alpha_1\alpha_2}(g) = \mathbb{R}^{\alpha_1}(g) \otimes \mathbb{R}^{\alpha_2}(g) : g \in G\}. \tag{II.1}$$

$\mathbb{R}^{\alpha_1\alpha_2}$ forms a unitary corepresentation of G which can be decomposed by a unitary $n_{\alpha_1} n_{\alpha_2}$-dimensional matrix $W^{\alpha_1\alpha_2} = W$ into a direct sum of its irreducible constituents. This unitary matrix must satisfy

$$W^{\dagger} \mathbb{R}^{\alpha_1\alpha_2}(g) W^{g} = \sum_{\alpha \in A_{\mathrm{I}}} \oplus M_{\alpha_1\alpha_2;\alpha} \mathbb{R}^{\alpha}(g) \oplus \sum_{\beta \in A_{\mathrm{II}}} \oplus M_{\alpha_1\alpha_2;\beta} \mathbb{R}^{\beta}(g) \oplus \sum_{\gamma \in A_{\mathrm{III}}} \oplus M_{\alpha_1\alpha_2;\gamma} \mathbb{R}^{\gamma}(g),$$

$$\text{for all } g \in G. \tag{II.2}$$

Using the orthogonality relations for the characters of the unirreps of H, we obtain the well-known results [13]

$$M_{\alpha_1\alpha_2;\alpha} = m_{\alpha_1\alpha_2;\alpha}, \tag{II.3}$$

$$M_{\alpha_1\alpha_2;\beta} = \tfrac{1}{2} m_{\alpha_1\alpha_2;\beta}, \tag{II.4}$$

$$M_{\alpha_1\alpha_2;\gamma} = m_{\alpha_1\alpha_2;\gamma} = M_{\alpha_1\alpha_2;\bar{\gamma}}, \tag{II.5}$$

where the symbols $m_{\ldots}$ denote multiplicities which refer to subduction with respect to the subgroup H. Equations (II.3)–(II.5) show that the representation theory with respect to the subgroup H will play a fundamental role when calculating CG coefficients for corepresentations and that $m_{\alpha_1\alpha_2;\beta}$ must be either zero or an even integer.

III. CG COEFFICIENTS FOR COREPRESENTATIONS

Because of the special structure of the CG series (II.2), it is suggestive to consider the three different cases separately depending on whether $\alpha \in A_{\mathrm{I}}$, or $\beta \in A_{\mathrm{II}}$, or $\gamma \in A_{\mathrm{III}}$. Before doing this let us recall that we assume from the outset that CG coefficients for H are known, which allow one to decompose the subduced representation

$$\mathbb{R}^{\alpha_1\alpha_2} \downarrow H = R^{\alpha_1\alpha_2} \tag{III.1}$$

into a direct sum of its irreducible constituents, i.e.,

$$M^{\dagger} R^{\alpha_1\alpha_2}(h) M = \sum_{\alpha \in A_{\mathrm{I}}} \oplus m_{\alpha_1\alpha_2;\alpha} R^{\alpha}(h) \oplus \sum_{\beta \in A_{\mathrm{II}}} \oplus m_{\alpha_1\alpha_2;\beta} \times R^{\beta}(h) \oplus \sum_{\gamma \in A_{\mathrm{III}}} \oplus m_{\alpha_1\alpha_2;\gamma} \{R^{\gamma}(h) \oplus Z^{\gamma\dagger} R^{\gamma}(h) Z^{\gamma}\}, \quad \text{for all } h \in H, \tag{III.2}$$

where the Kronecker products of unirreps of H are denoted by

$$R^{\alpha_1\alpha_2} = \{R^{\alpha_1\alpha_2}(h) = R^{\alpha_1}(h) \otimes R^{\alpha_2}(h) : h \in H\}. \tag{III.3}$$

A. CG coefficients of type I

Utilizing the unitarity of the CG matrix W, we can rewrite the defining equations for CG coefficients of corepresentations in a similar way as in Ref. 5. These equations read as

$$\mathbb{R}^{\alpha_1\alpha_2}(h) \mathbf{W}^{\alpha w}_{k} = \sum_{l=1}^{n_\alpha} R^{\alpha}_{lk}(h) \mathbf{W}^{\alpha w}_{l}, \quad \text{for all } h \in H, \tag{III.4}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(s) \mathbf{W}^{\alpha w *}_{k} = \sum_{l=1}^{n_\alpha} U^{\alpha}_{lk} \mathbf{W}^{\alpha w}_{l},$$

$$w = 1,2,\ldots,M_{\alpha_1\alpha_2;\alpha}, \quad k = 1,2,\ldots,n_{\alpha}, \tag{III.5}$$

at which we have to note that it suffices to take the transformation laws (III.5) into account, in order to be able to satisfy Eq. (II.2) for each group element. Thereby we have introduced an abbreviated notation which reads in more detail

$$\{\mathbf{W}^{\alpha w}_{k}\}_{ij} = \{\mathbf{W}^{\alpha_1\alpha_2;\alpha w}_{k}\}_{ij} = W^{\alpha_1\alpha_2}_{ij;\alpha w k}$$

$$\alpha \in A_1, \quad w = 1,2,\ldots,M_{\alpha_1\alpha_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha,$$
$$i = 1,2,\ldots,n_{\alpha_1}, \quad j = 1,2,\ldots,n_{\alpha_2}. \tag{III.6}$$

Hence, Eqs. (III.4) and (III.5) allow one to interpret the columns of the CG matrix W as H-adapted vectors of a $n_{\alpha_1} n_{\alpha_2}$-dimensional Euclidean space $\mathscr{W}^{\alpha_1\alpha_2}$, i.e., vectors which transform according to the considered unirrep of H, but which have to satisfy additionally the conditions (III.5). In order to get a unitary matrix W, it is necessary that the vectors $\mathbf{W}_j^{\alpha w}$ are orthornormal with respect to each index. The orthogonality with respect to α and j follows directly from their transformation properties, whereas orthogonality with respect to the multiplicity index w can only be achieved by further manipulations.

In order to be able to calculate systematically H-adapted vectors, we introduce like in Ref. 5 a $n_{\alpha_1} n_{\alpha_2}$-dimensional matrix representation of the group algebra $A(H)$. In this connection we have to note that it is meaningless to extend the concept of group algebras to the supergroup G^3. The corresponding units of $A(H)$ are given by

$$\mathbb{E}_{ij}^\alpha = \mathbb{E}_{ij}^{\alpha_1\alpha_2;\alpha}$$
$$= \frac{n_\alpha}{|H|} \sum_{h\in H} R_{ij}^{\alpha *}(h)\mathbb{R}^{\alpha_1\alpha_2}(h) = E_{ij}^\alpha, \tag{III.7}$$

where the symbol on the right-hand side should indicate that these units refer to the representation (II.3) of H. The units (III.7) satisfy apart from their well known rules the following important relations:

$$\mathbb{R}^{\alpha_1\alpha_2}(s)^\dagger \mathbb{E}_{ij}^\alpha \mathbb{R}^{\alpha_1\alpha_2}(s) = \sum_{k,l} U_{ik}^{\alpha *} U_{jl}^\alpha \mathbb{E}_{kl}^{\alpha *}, \tag{III.8}$$

which reflect the features of corepresentations in terms of the group algebra $A(H)$.

Since the CG matrix W is assumed to be unitary, the vectors

$$\mathbf{W}_k^{\alpha w}, \quad w = 1,2\ldots,M_{\alpha_1\alpha_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \tag{III.9}$$

form an orthonormal basis of

$$\mathscr{W}^{\alpha_1\alpha_2;\alpha} = \sum_i \mathbb{E}_{ii}^\alpha \mathscr{W}^{\alpha_1\alpha_2}, \quad \dim \mathscr{W}^{\alpha_1\alpha_2;\alpha} = n_\alpha M_{\alpha_1\alpha_2;\alpha}. \tag{III.10}$$

Presupposing a unitary CG matrix $M^{\alpha_1\alpha_2} = M$ is known which satisfies Eqs. (III.2), it is obvious that because of

$$\{\mathbf{M}_k^{\mu v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\alpha_2;\mu v}\}_{ij} = M_{ij;\mu v k}^{\alpha_1\alpha_2}$$
$$\mu \in A_H, \quad v = 1,2,\ldots,m_{\alpha_1\alpha_2;\mu}, \quad k = 1,2,\ldots,n_\mu, \tag{III.11}$$

the vectors

$$\mathbf{M}_k^{\alpha v}, \quad v = 1,2,\ldots,m_{\alpha_1\alpha_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \tag{III.12}$$

form another orthonormal basis of $\mathscr{W}^{\alpha_1\alpha_2;\alpha}$. Thereby we have to note that the multiplicity index v originates from the corresponding subduction with respect to the subgroup H and should therefore not be confused with w. Corresponding to their definition, the vectors (III.12) transform according to

$$\mathbb{R}^{\alpha_1\alpha_2}(h)\mathbf{M}_k^{\alpha v} = \sum_l R_{lk}^\alpha(h)\mathbf{M}_l^{\alpha v}, \quad \text{for all } h\in H, \tag{III.13}$$

whereas the additional conditions (III.5) are not satisfied in general.

For this reason we remember that the elements of both bases (III.9) and (III.12) must be linked by a unitary transformation. Utilizing Schur's lemma with respect to the subgroup H, the unitary transformation must be independent of the index k:

$$\mathbf{W}_k^{\alpha w} = \sum_{v=1}^{m_{\alpha_1\alpha_2;\alpha}} B_{vw}\mathbf{M}_k^{\alpha v}, \tag{III.14}$$

$$\mathbf{M}_k^{\alpha v} = \sum_{w=1}^{M_{\alpha_1\alpha_2;\alpha}} B_{vw}^* \mathbf{W}_k^{\alpha w}, \quad k = 1,2,\ldots,n_\alpha. \tag{III.15}$$

Obviously, unitarity of the $M_{\alpha_1\alpha_2;\alpha}$-dimensional matrix B implies orthonormality of the vectors $\mathbf{W}_k^{\alpha w}$ also with respect to w.

Hence, our problem is now reduced to the task of determining a unitary $M_{\alpha_1\alpha_2;\alpha}$-dimensional matrix B, so that the corresponding vectors satisfy Eqs. (III.5). For this purpose we consider

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{M}_k^{\alpha v *} = \sum_l U_{lk}^\alpha \sum_{v'} \{\sum_w B_{vw} B_{v'w}\}\mathbf{M}_l^{\alpha v'}, \tag{III.16}$$

where we have used Eqs. (III.5), (III.14), and (III.15). Introducing the notation

$$F_{v'v} = \{BB^T\}_{vv'} = \sum_w B_{v'w}B_{vw}, \quad v,v' = 1,2,\ldots,m_{\alpha_1\alpha_2;\alpha}, \tag{III.17}$$

Eqs. (III.16) turn out to be

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{M}_k^{\alpha v *} = \sum_l U_{lk}^\alpha \sum_{v'} F_{v'v}\,\mathbf{M}_l^{\alpha v'}. \tag{III.18}$$

Conversely we obtain

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U_{lk}^\alpha\, \mathbf{W}_l^{\alpha w}$$
$$= \sum_l U_{lk}^\alpha \sum_{w'} \{B^\dagger F B^*\}_{w'w} \mathbf{W}_l^{\alpha w'}, \tag{III.19}$$

by taking Eqs. (III.18), (III.14), and (III.15) into account and writing the matrix multiplication symbolically. Hence, if we can find a matrix B satisfying

$$FB^* = B, \quad \text{with} \quad BB^\dagger = B^\dagger B = \mathbb{1}_M, \tag{III.20}$$

the corresponding CG coefficients are immediately obtained from Eqs. (III.14). The symbol $\mathbb{1}_M$ denotes the $M_{\alpha_1\alpha_2;\alpha}$-dimensional unit matrix. Corresponding to its definition (III.17), the $M_{\alpha_1\alpha_2;\alpha}$-dimensional matrix F is symmetric and unitary, which can be verified by means of

$$\mathbb{R}^{\alpha_1\alpha_2}(s^2)\mathbf{M}_k^{\alpha v} = \sum_l R_{lk}^\alpha(s^2)\mathbf{M}_l^{\alpha v}. \tag{III.21}$$

Inserting Eqs. (III.18) twice into Eqs. (III.21) and utilizing Eq. (I.10), we obtain

$$FF^* = \mathbb{1}_M, \tag{III.22}$$

which verifies our proposition.

The unitary matrix F is uniquely fixed through Eqs. (III.18). Its matrix elements can therefore be determined by means of

$$\langle \mathbf{M}_k^{\alpha v'}, \ U^{\alpha_1} \otimes U^{\alpha_2}\{\sum_l U_{kl}^\alpha \mathbf{M}_l^{\alpha v}\}^* \rangle$$

$$= F_{v'v} = F_{v'v}^{\alpha(\alpha_1\alpha_2)},$$

$$v,v' = 1,2,\ldots, m_{\alpha_1\alpha_2;\alpha}, \qquad \text{(III.23)}$$

whose values may not depend on the free index k. The corresponding proof can be carried out readily by utilizing Eq. (III.8) together with the well-known properties of the units.

Presupposing the unitary CG matrix M can be calculated by means of the method given in Ref. 5, its matrix elements take the form

$$\{\mathbf{M}_k^{\alpha v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\alpha_2;\alpha(i_v j_v)}\}_{ij}$$
$$= ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\alpha(i_v j_v)}||^{-1} \frac{n_\alpha}{|H|} \sum_h R_{ii_v}^{\alpha_1}(h) R_{jj_v}^{\alpha_2}(h)\, R_{ka_0}^{\alpha *}(h),$$
$$v = 1,2,\ldots,m_{\alpha_1\alpha_2;\alpha},\quad k = 1,2,\ldots,n_\alpha, \qquad \text{(III.24)}$$

where we have used the same notation as in Ref. 5. Inserting these special values into Eq. (III.23), we obtain after a straightforward calculation as matrix elements

$$F_{v'v}^{\alpha(\alpha_1\alpha_2)} = ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\alpha(i_{v'} j_{v'})}||^{-1}\, ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\alpha(i_v j_v)}||^{-1}$$
$$\times \frac{n_\alpha}{|H|} \sum_h \mathbb{R}_{i_{v'} i_v}^{\alpha_1}(hs) \mathbb{R}_{j_{v'} j_v}^{\alpha_2}(hs) \mathbb{R}_{a_0 a_0}^{\alpha *}(hs), \qquad \text{(III.25)}$$

which are indeed independent of the free index k.

To summarize our results, the problem of calculating CG coefficients of type I for corepresentations $\mathbb{R}^{\alpha_1\alpha_2}$ is reduced to the task of solving Eq. (III.20). This problem is of course less complicated than that of calculating CG coefficients for corepresentations directly from their defining equations (III.4) and (III.5) without utilizing the properties of CG coefficients for the subgroup.

B. CG coefficients of type II

Like in the previous case we utilize the unitarity of the CG matrix W which allows one to write the defining equations for CG coefficients of type II for corepresentations in the following way:

$$\mathbb{R}^{\alpha_1\alpha_2}(h)\, \mathbf{W}_{dk}^{\beta w} = \sum_{l=1}^{n_\beta} R_{lk}^{\beta}(h)\, \mathbf{W}_{dl}^{\beta w}, \quad \text{for all } h \in H, \qquad \text{(III.26)}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\, \mathbf{W}_{dk}^{\beta w *} = (-1)^{\Delta(d+1)} \sum_{l=1}^{n_\beta} U_{lk}^{\beta}\, \mathbf{W}_{d+1,l}^{\beta w},$$
$$w = 1,2,\ldots,M_{\alpha_1\alpha_2;\beta},\quad d = 1,2, \text{ and } k = 1,2.,,,n_\beta, \qquad \text{(III.27)}$$

where the special matrix notation (I.21) and (I.22) has been already taken into account. In this connection we have to note that it suffices to consider Eqs. (III.27) since the remaining equations follow Eq. (III.26) together with (III.27). Our abbreviated notation reads in more detail

$$\{\mathbf{W}_{dk}^{\beta w}\}_{ij} = \{\mathbf{W}_{dk}^{\alpha_1\alpha_2;\beta w}\}_{ij} = W_{ij;\beta w dk}^{\alpha_1\alpha_2},$$
$$\beta \in A_{\mathrm{II}},\quad w = 1,2,\ldots,M_{\alpha_1\alpha_2;\beta},\quad d = 1,2, \quad \text{and}$$
$$k = 1,2,\ldots,n_\beta,\ i = 1,2,\ldots,n_{\alpha_1},\quad j = 1,2,\ldots,n_{\alpha_2}. \qquad \text{(III.28)}$$

Equations (III.26) and (III.27) allow one to interpret the columns of the CG matrix W as H-adapted vectors of $\mathscr{W}^{\alpha_1\alpha_2}$, but which have to satisfy additionally Eq. (III.27). In order to be able to achieve that the CG matrix W is unitary, it is necessary to require that the vectors $\mathbf{W}_{dk}^{\beta w}$ are orthonormal with respect to each index. Orthogonality with respect to β and k is automatically guaranteed through their transformation properties (III.26), whereas the orthogonality with respect to the index d and the multiplicity index w can only be achieved by further manipulations.

Therefore, the vectors

$$\mathbf{W}_{dk}^{\beta w},\quad w = 1,2,\ldots,\quad M_{\alpha_1\alpha_2;\beta},\quad d = 1,2, \text{ and}$$
$$k = 1,2,\ldots,n_\beta, \qquad \text{(III.29)}$$

form an orthonormal basis of

$$\mathscr{W}^{\alpha_1\alpha_2;\beta} = \sum_i \mathbb{E}_{ii}^{\beta}\, \mathscr{W}^{\alpha_1\alpha_2},\quad \dim\, \mathscr{W}^{\alpha_1\alpha_2;\beta} = 2n_\beta M_{\alpha_1\alpha_2;\beta}, \qquad \text{(III.30)}$$

where the corresponding units are given by

$$\mathbb{E}_{ij}^{\beta} = \mathbb{E}_{ij}^{\alpha_1\alpha_2;\beta} = \frac{n_\beta}{|H|} \sum_h R_{ij}^{\beta *}(h) \mathbb{R}^{\alpha_1\alpha_2}(h) = E_{ij}^{\beta}. \qquad \text{(III.31)}$$

The symbols on the right-hand side of Eq. (III.31) should reflect the special property (III.1). The units (III.31) satisfy the following important relations:

$$\mathbb{R}^{\alpha_1\alpha_2}(s)^\dagger \mathbb{E}_{ij}^{\beta}\, \mathbb{R}^{\alpha_1\alpha_2}(s) = \sum_{kl} U_{ik}^{\beta *} U_{jl}^{\beta}\, \mathbb{E}_{kl}^{\beta *}, \qquad \text{(III.32)}$$

which express the typical features of corepresentations in terms of the group algebra of the subgroup H. Because of Eq. (III.11), the vectors

$$\mathbf{M}_k^{\alpha v},\quad v = 1,2\ldots,m_{\alpha_1\alpha_2;\beta},\quad k = 1,2,\ldots,n_\beta, \qquad \text{(III.33)}$$

define a further orthornormal, H-adapted basis of $\mathscr{W}^{\alpha_1\alpha_2;\beta}$,i.e.,

$$\mathbb{R}^{\alpha_1\alpha_2}(h)\mathbf{M}_k^{\beta v} = \sum_l R_{lk}^{\beta}(h)\mathbf{M}_l^{\beta v}, \quad \text{for all } h \in H. \qquad \text{(III.34)}$$

Schur's lemma with respect to the subgroup H requires that the elements of the bases (III.29) and (III.33) must be linked by a unitary transformation which is independent of the free index k:

$$\mathbf{W}_{dk}^{\beta w} = \sum_{v=1}^{m_{\alpha_1\alpha_2;\beta}} B_{v;dw}\, \mathbf{M}_k^{\beta v}, \qquad \text{(III.35)}$$

$$\mathbf{M}_k^{\beta v} = \sum_{d=1}^{2} \sum_{w=1}^{M_{\alpha_1\alpha_2;\beta}} B_{v;dw}^{*}\, \mathbf{W}_{dk}^{\beta w},\quad k = 1,2,\ldots,n_\beta. \qquad \text{(III.36)}$$

Unitarity of the $2\,M_{\alpha_1\alpha_2;\beta}$-dimensional matrix B assures that the vectors $\mathbf{W}_{dk}^{\beta w}$ are also orthonormal with respect to w and d.

Therefore, the problem is now to determine a unitary $2M_{\alpha_1\alpha_2;\beta}$-dimensional matrix B, so that the corresponding vectors are satisfying Eq. (III.27). For this purpose we consider

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{M}_k^{\beta v *} = \sum_l U_{lk}^{\beta} \sum_{v'} \Big\{ \sum_{dw} B_{v;dw}$$
$$\times (-1)^{\Delta(d+1)} B_{v';d+1,w} \Big\}\, \mathbf{M}_l^{\beta v'}, \qquad \text{(III.37)}$$

where we have already used Eqs. (III.35), and (III.36). Let us introduce the following notations:

$$G_{dw;d'w'} = (-1)^{\Delta(d)} \delta_{d',d+1}\, \delta_{ww'},$$
$$d,d' = 1,2 \text{ and } w,w' = 1,2,\ldots, M_{\alpha_1\alpha_2;\beta}, \qquad \text{(III.38)}$$

$$F_{v'v} = \{BGB^T\}_{vv'} = \sum_{dw} B_{v;dw}(-1)^{\Delta(d+1)} B_{v';d+1,w},$$
$$v,v' = 1,2,\ldots,m_{\alpha_1\alpha_2;\beta}, \tag{III.39}$$

which have as a consequence of Eq. (III.37)

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\,\mathbf{M}_k^{\beta v *} = \sum_l U^{\beta}_{lk} \sum_{v'} F_{v'v}\,\mathbf{M}_l^{\beta v'}. \tag{III.40}$$

Conversely, we obtain

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{W}_{dk}^{\beta w *} = (-1)^{\Delta(d+1)} \sum_l U^{\beta}_{lk}\,\mathbf{W}^{\beta w}_{d+1,l}$$
$$= \sum_l U^{\beta}_{lk} \sum_{d'w'} \{B^\dagger F B^*\}_{d'w',dw}\,\mathbf{W}^{\beta w'}_{d'l}, \tag{III.41}$$

by taking Eqs. (III.35), (III.36), and (III.40) into account. Utilizing the orthonormality of the vectors (III.29), Eqs. (III.41) yield

$$(-1)^{\Delta(d+1)}\,\delta_{d',d+1}\,\delta_{ww'} = \sum_{vv'} B^\dagger_{d'w';v'}\,F_{v'v}\,B^*_{v;dw}, \tag{III.42}$$

or in matrix notation

$$G^T = B^\dagger F B^*. \tag{III.43}$$

Hence, if we can find a unitary matrix B which satisfies

$$B\,G^T = F B^*, \tag{III.44}$$

the corresponding CG coefficients follow immediately from Eq. (III.35). Now it is easy to verify that the matrix F is not only antisymmetric but also unitary. This can be seen from

$$\mathbb{R}^{\alpha_1\alpha_2}(s^2)\mathbf{M}_k^{\beta v} = \sum_l R^{\beta}_{lk}(s^2)\mathbf{M}_l^{\beta v}, \tag{III.45}$$

which has the nontrivial consequence

$$F F^* = -\mathbf{1}_{2M}, \tag{III.46}$$

where the symbol $\mathbf{1}_{2M}$ denotes the $2M_{\alpha_1\alpha_2;\beta}$-dimensional unit matrix. In order to verify Eq. (III.46) one has only to use Eq. (III.40) twice and Eq. (I.14).

The antisymmetric unitary matrix F is uniquely fixed through Eq. (III.40). Its matrix elements are given by

$$\left\langle \mathbf{M}_k^{\beta v'},\, U^{\alpha_1}\otimes U^{\alpha_2}\left\{\sum_l U^{\beta}_{kl}\mathbf{M}_l^{\beta v}\right\}^*\right\rangle = F_{v'v} = F^{\beta(\alpha_1\alpha_2)}_{v'v},$$
$$v,v' = 1,2,\ldots,m_{\alpha_1\alpha_2;\beta}, \tag{III.47}$$

and are of course independent of the free index k, which can be readily proven by means of Eq. (III.32).

Assuming that the CG matrix M can be calculated by means of the method given in Ref. 5, its matrix elements take the form

$$\{\mathbf{M}_k^{\beta v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\alpha_2;\beta(i_v j_v)}\}_{ij}$$
$$= ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\beta(i_v j_v)}||^{-1}\,\frac{n_\beta}{|H|}$$
$$\times \sum_h R^{\alpha_1}_{ii_v}(h) R^{\alpha_2}_{jj_v}(h)\,R^{\beta *}_{ka_0}(h),$$
$$v = 1,2,\ldots,m_{\alpha_1\alpha_2;\beta} \quad k = 1,2,\ldots,n_\beta. \tag{III.48}$$

A straightforward calculation yields for Eq. (III.47)

$$F^{\beta(\alpha_1\alpha_2)}_{v'v} = ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\beta(i_v j_v)}||^{-1}\,||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\beta(i_v j_v)}||^{-1}$$
$$\times \frac{n_\beta}{|H|} \sum_h \mathbb{R}^{\alpha_1}_{i_{v'} i_v}(hs)\mathbb{R}^{\alpha_2}_{j_{v'} j_v}(hs)\,\{R^{\beta}(h)U^{\beta}\}^*_{a_0 a_0}, \tag{III.49}$$

which are indeed independent of the index k.

Now let us return to the problem of determining a unitary matrix B which satisfies Eq. (III.44). Introducing the following notation:

$$\{\mathbf{B}^{d,w}\}_v = \mathbf{B}_{v,dw},$$
$$d = 1,2 \text{ and } w = 1,2,\ldots,M_{\alpha_1\alpha_2;\beta}, \quad v = 1,2,\ldots,m_{\alpha_1\alpha_2;\beta}, \tag{III.50}$$

Eq. (III.44) can be written as

$$F\mathbf{B}^{d,w*} = (-1)^{\Delta(d+1)}\,\mathbf{B}^{d+1,w},$$
$$d = 1,2 \text{ and } w = 1,2,\ldots,M_{\alpha_1\alpha_2;\beta}. \tag{III.51}$$

Hence the columns of the matrix B can be seen as vectors of a $2M_{\alpha_1\alpha_2;\beta}$-dimensional Euclidean space which have a special transformation law with respect to F. Because of

$$F\mathbf{B}^{d+1,w*} = F\{(-1)^{\Delta(d+1)}F\mathbf{B}^{d,w*}\}^*$$
$$= (-1)^{\Delta(d)}\,\mathbf{B}^{d,w}, \tag{III.52}$$

it suffices to determine, for example, for $d = 1$ just $M_{\alpha_1\alpha_2;\beta}$ orthonormal vectors $\mathbf{B}^{1,w}$, $w = 1,2,\ldots,M_{\alpha_1\alpha_2;\beta}$,

$$\langle \mathbf{B}^{1,w}, \mathbf{B}^{1,w'}\rangle = \delta_{ww'} \Longleftrightarrow \langle \mathbf{B}^{2,w}, \mathbf{B}^{2,w'}\rangle = \delta_{ww'}, \tag{III.53}$$

which satisfy additionally

$$\langle \mathbf{B}^{1,w}, F\mathbf{B}^{1,w'*}\rangle = 0, \text{ for all } w,w' = 1,2,\ldots,M_{\alpha_1\alpha_2;\beta}. \tag{III.54}$$

Condition (III.54) is necessary and sufficient that the corresponding matrix B is unitary.

Although $F^T = -F$ implies $F_{vv} = 0$, Eq. (III.54) represents a nontrivial condition, if $m_{\alpha_1\alpha_2;\beta} \geqslant 4$.. For the special case $m_{\alpha_1\alpha_2;\beta} = 2$ it follows from

$$F_{11} = F_{22} = 0 \tag{III.55}$$

that we can choose

$$\{\mathbf{B}^{1,w}\}_v = \delta_{vw}, \quad w = 1, \tag{III.56}$$

which implies for

$$\{\mathbf{B}^{2,w}\}_v = -F_{21}\,\delta_{v2}. \tag{III.57}$$

The corresponding CG coefficients are immediately obtained by inserting Eqs. (III.56), and (III.57) into (III.35). For this special case F_{12} must be a unimodular number, since F is a unitary matrix. In this connection we have to note that the condition $F^T = -F$ requires that the multiplicity $m_{\alpha_1\alpha_2;\beta}$ must be larger than one, if $M_{\alpha_1\alpha_2;\beta} \neq 0$.

To conclude this section we realize that Eqs. (III.52) together with Eq. (III.54) will simplify the determination of a unitary matrix B.

C. CG-coefficients of type III

Utilizing once more the unitarity of the CG matrix W, the defining equations for CG coefficients of type III for corepresentations can be written as follows:

$$\mathbb{R}^{\alpha_1\alpha_2}(h)\mathbf{W}_{1k}^{\gamma w} = \sum_{l=1}^{n_\gamma} R^{\gamma}_{lk}(h)\mathbf{W}_{1l}^{\gamma w}, \tag{III.58}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(h)\mathbf{W}_{2k}^{\gamma w} = \sum_{l=1}^{n_\gamma} \{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}_{lk}\,\mathbf{W}_{2l}^{\gamma w},$$

for all $h \in H$, (III.59)

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{W}_{1k}^{\gamma w *} = \mathbf{W}_{2k}^{\gamma w}, \quad \text{(III.60)}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{W}_{2k}^{\gamma w *} = \sum_{l=1}^{n_\gamma} R_{lk}^{\gamma}(s^2)\,\mathbf{W}_{1l}^{\gamma w}, \quad w = 1,2\ldots,M_{\alpha_1\alpha_2;\gamma}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(III.61)}$$

where the special matrix notation (I.23), and (I.24) has already been taken into account. Analogous to the previous cases, it suffices to consider Eqs. (III.60) and (III.61), since the remaining are then automatically satisfied:

$$\{\mathbf{W}_{dk}^{\gamma w}\}_{ij} = \{\mathbf{W}_{dk}^{\alpha_1\alpha_2;\gamma w}\}_{ij} = W_{ij;\gamma w dk}^{\alpha_1\alpha_2}, \quad \gamma \in A_{III},$$
$$w = 1,2,\ldots,M_{\alpha_1\alpha_2;\gamma}, d = 1,2 \text{ and } k = 1,2,\ldots,n_\gamma = n_{\bar\gamma},$$
$$i = 1,2,\ldots,n_{\alpha_1}, \quad j = 1,2,\ldots,n_{\alpha_2}. \quad \text{(III.62)}$$

However, contrary to the previous cases, one has to be more careful when considering the transformation properties of the columns of the CG matrix W with respect to the subgroup H. Of course they can be seen as H-adapted vectors $\mathscr{W}^{\alpha_1\alpha_2}$, but one must be aware of the different transformation laws (III.58) and (III.59). Nevertheless, the columns of the CG matrix W are H-adapted vectors which have to satisfy additionally Eqs. (III.60) and (III.61). Unitarity of W requires orthonormality of the vectors $\mathbf{W}_{dk}^{\gamma w}$ with respect to each index. Since the unirreps R^{γ} and $R^{\bar\gamma}$ are inequivalent, the vectors $\mathbf{W}_{dk}^{\gamma w}$ are automatically orthogonal with respect to the indices γ,k,d.

This property can also be shown by means of the corresponding units

$$\mathbb{E}_{ij}^{\gamma} = \mathbb{E}_{ij}^{\alpha_1\alpha_2;\gamma} = \frac{n_\gamma}{|H|}\sum_h R_{ij}^{\gamma *}(h)\mathbb{R}^{\alpha_1\alpha_2}(h) = E_{ij}^{\gamma}, \quad \text{(III.63)}$$

$$E_{ij}^{\bar\gamma} = E_{ij}^{\alpha_1\alpha_2;\bar\gamma} = \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma\dagger} R^{\bar\gamma}(h)\, Z^{\gamma}\}_{ij}^{*}\mathbb{R}^{\alpha_1\alpha_2}(h) = E_{ij}^{\bar\gamma}, \quad \text{(III.64)}$$

which satisfy furthermore the following important relations:

$$\mathbb{R}^{\alpha_1\alpha_2}(s)^\dagger\, \mathbb{E}_{ij}^{\gamma}\, \mathbb{R}^{\alpha_1\alpha_2}(s) = \sum_{kl} R_{ik}^{\gamma *}(s^2)\, R_{jl}^{\gamma}(s^2)\, \mathbb{E}_{kl}^{\bar\gamma *}, \quad \text{(III.65)}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(s)^\dagger\, \mathbb{E}_{ij}^{\bar\gamma}\, \mathbb{R}^{\alpha_1\alpha_2}(s) = \mathbb{E}_{ij}^{\gamma *}. \quad \text{(III.66)}$$

These relations reflect the typical features of corepresentations in terms of the corresponding group algebra of the subgroup . Note that the symbols E_{ij}^{γ} and $E_{ij}^{\bar\gamma}$ refer to $R^{\alpha_1\alpha_2}$.

Since the CG matrix W is presupposed as unitary, the vectors

$$\mathbf{W}_{dk}^{\gamma w}, \quad w = 1,2,\ldots,M_{\alpha_1\alpha_2;\gamma}, \quad d = 1,2, \text{ and } k = 1,2,\ldots,n_\gamma \quad \text{(III.67)}$$

are an orthonormal H-adapted basis of

$$\mathscr{W}^{\alpha_1\alpha_2;\gamma} = \sum_i \{\mathbb{E}_{ii}^{\gamma} + \mathbb{E}_{ii}^{\bar\gamma}\}\mathscr{W}^{\alpha_1\alpha_2},$$
$$\dim \mathscr{W}^{\alpha_1\alpha_2;\gamma} = 2n_\gamma\, M_{\alpha_1\alpha_2;\gamma}. \quad \text{(III.68)}$$

Because of Eq. (III.11), the vectors

$$\mathbf{M}_{k}^{\gamma v}, \quad v = 1,2,\ldots,m_{\alpha_1\alpha_2;\gamma}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(III.69)}$$

$$\mathbf{M}_{k}^{\bar\gamma v}, \quad v = 1,2,\ldots,m_{\alpha_1\alpha_2;\gamma}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(III.70)}$$

form a further orthonormal H-adapted basis of $\mathscr{W}^{\alpha_1\alpha_2;\gamma}$.

These vectors transform according to

$$\mathbb{R}^{\alpha_1\alpha_2}(h)\mathbf{M}_{k}^{\gamma v} = \sum_l R_{lk}^{\gamma}(h)\mathbf{M}_{l}^{\gamma v}, \quad \text{(III.71)}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(h)\mathbf{M}_{k}^{\bar\gamma v} = \sum_l \{Z^{\gamma\dagger} R^{\bar\gamma}(h) Z^{\gamma}\}_{lk}\mathbf{M}_{l}^{\bar\gamma v}, \quad \text{for all } h \in H. \quad \text{(III.72)}$$

Due to Schur's lemma with respect ot the subgroup H, it follows that the elements of the bases (III.67) and (III.69), and (III.70) must be linked by special unitary transformations which may not depend on the index k:

$$\mathbf{W}_{1k}^{\gamma w} = \sum_{v=1}^{M_{\alpha_1\alpha_2;\gamma}} B_{vw}\, \mathbf{M}_{k}^{\gamma v}, \quad \text{(III.73)}$$

$$\mathbf{M}_{k}^{\gamma v} = \sum_{w=1}^{M_{\alpha_1\alpha_2;\gamma}} B_{vw}^{*}\, \mathbf{W}_{dk}^{\gamma w}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(III.74)}$$

$$\mathbf{W}_{2k}^{\gamma w} = \sum_{v=1}^{M_{\alpha_1\alpha_2;\gamma}} C_{vw}\, \mathbf{M}_{k}^{\bar\gamma v}, \quad \text{(III.75)}$$

$$\mathbf{M}_{k}^{\bar\gamma v} = \sum_{w=1}^{M_{\alpha_1\alpha_2;\gamma}} C_{vw}^{*}\mathbf{W}_{2k}^{\gamma w}, \quad k = 1,2,\ldots,n_\gamma. \quad \text{(III.76)}$$

Hence, if the matrices B and C are unitary, the corresponding vectors $\mathbf{W}_{dk}^{\gamma w}$ are also orthonormal with respect to the multiplicity index w.

Thus, our problem is now reduced to the task of determining $M_{\alpha_1\alpha_2;\gamma}$-dimensional unitary matrices B and C, so that Eqs. (III.60) and (III.61) are satisfied. For this purpose we consider the following expressions:

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{M}_{k}^{\gamma v *} = \sum_{v'}\left\{\sum_{w} C_{v'w}\, B_{vw}\right\}\mathbf{M}_{k}^{\bar\gamma v'}, \quad \text{(III.77)}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{M}_{k}^{\bar\gamma v *} = \sum_l R_{lk}^{\gamma}(s^2)\sum_{v'}\left\{\sum_{w} B_{v'w} C_{vw}\right\}\mathbf{M}_{l}^{\gamma v'}, \quad \text{(III.78)}$$

which can be readily verified by means of Eqs. (III.74), (III.60), and (III.75), respectively Eqs. (III.76), (III.61), and (III.73). Introducing the notation

$$F_{v'v} = \{CB^{T}\}_{v'v} = \sum_{w} C_{v'w}B_{vw}, \quad v,v' = 1,2,\ldots,m_{\alpha_1\alpha_2;\gamma}, \quad \text{(III.79)}$$

Eqs. (III.77), and (III.78) read as

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{M}_{k}^{\gamma v *} = \sum_{v'} F_{v'v}\, \mathbf{M}_{k}^{\bar\gamma v'}, \quad \text{(III.80)}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{M}_{k}^{\bar\gamma v *} = \sum_l R_{lk}^{\gamma}(s^2)\sum_{v'} F_{vv'}\, \mathbf{M}_{l}^{\gamma v'}. \quad \text{(III.81)}$$

Conversely we obtain

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{W}_{1k}^{\gamma w *} = \mathbf{W}_{2k}^{\gamma w} = \sum_{w'}\{C^{\dagger}FB^{*}\}_{w'w}\, \mathbf{W}_{2k}^{\gamma w'}, \quad \text{(III.82)}$$

$$\mathbb{R}^{\alpha_1\alpha_2}(s)\mathbf{W}_{2k}^{\gamma w *} = \sum_l R_{lk}^{\gamma}(s^2)\mathbf{W}_{1l}^{\gamma w} = \sum_l R_{lk}^{\gamma}(s^2)\sum_{w'}\{C^{\dagger}FB^{*}\}_{w'w}^{T}\, \mathbf{W}_{1l}^{\gamma w'} \quad \text{(III.83)}$$

from what follows:

$$C^{\dagger}FB^{*} = \mathbb{1}_m \quad \text{(III.84)}$$

and which represents the defining equation for B and C. Thereby we have to note that F is a unitary matrix. This can be shown, for example, by means of

$$\mathbb{R}^{\alpha_1\alpha_2}(s^2)\mathbf{M}_k^{\gamma v} = \sum_l R_{lk}^{\gamma}(s^2)\mathbf{M}_l^{\gamma v}, \tag{III.85}$$

by using Eqs. (III.80), and (III.81)

Since B and C must be unitary matrices satisfying Eq. (III.84), it is obvious to choose as a special solution

$$B = \mathbb{1}_m \Longleftrightarrow F = C, \tag{III.86}$$

from what follows that the vectors

$$\mathbf{W}_{1k}^{\gamma w} = \mathbf{M}_k^{\gamma w},$$
$$w = 1,2,...,m_{\alpha_1\alpha_2;\gamma}, \quad k = 1,2,...,n_\gamma, \tag{III.87}$$
$$\mathbf{W}_{2k}^{\gamma w} = \sum_v F_{vw}\mathbf{M}_k^{\bar{\gamma} v},$$
$$w = 1,2,...,m_{\alpha_1\alpha_2;\gamma}, \quad k = 1,2,...,n_\gamma, \tag{III.88}$$

are the desired columns of the CG matrix. Hence, it follows that the multiplicity problem is solved in a special way, since the multiplicity index w can be identified with v.

In order to be able to write Eq. (III.88) explicitly, it is necessary to calculate the matrix elements of F. This has to be done by means of

$$\langle \mathbf{M}_k^{\bar{\gamma} v'}, U^{\alpha_1} \otimes U^{\alpha_2}\mathbf{M}_k^{\gamma v *}\rangle = F_{v'v} = F_{v'v}^{\bar{\gamma}(\alpha_1\alpha_2)},$$
$$v,v' = 1,2,...,m_{\alpha_1\alpha_2;\gamma}, \tag{III.89}$$

whose values must be independent of the free index k. This proposition can be proven with the aid of Eq. (III.66).

Presupposing that the CG matrix M can be computed by means of the method described in Ref. 5, the corresponding matrix elements are given by

$$\{\mathbf{M}_k^{\gamma v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\alpha_2;\gamma(i_v j_v)}\}_{ij}$$
$$= ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\gamma(i_v j_v)}||^{-1}\frac{n_\gamma}{|H|}\sum_h R_{ii_v}^{\alpha_1}(h)R_{jj_v}^{\alpha_2}(h)$$
$$\times R_{ka_0}^{\gamma *}(h),$$
$$v = 1,2,...,m_{\alpha_1\alpha_2;\gamma}, \quad k = 1,2,...,n_\gamma, \tag{III.90}$$
$$\{\mathbf{M}_k^{\bar{\gamma} v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\alpha_2;\bar{\gamma}(i_v j_v)}\}_{ij}$$
$$= ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\bar{\gamma}(i_v j_v)}||^{-1}\frac{n_\gamma}{|H|}\sum_h R_{ii_v}^{\alpha_1}(h)R_{jj_v}^{\alpha_2}(h)$$
$$\times\{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{ka_0}^{*}$$
$$v = 1,2,...,m_{\alpha_1\alpha_2;\gamma}, \quad k = 1,2,...,n_\gamma, \tag{III.91}$$

where the index sets $\{(i_v, j_v)\}$ occurring in Eqs. (III.90), and (III.91) are of the same order, but in general not identical. Inserting Eqs. (III.90) and (III.91) into Eq. (III.89), we obtain

$$F_{v'v}^{\bar{\gamma}(\alpha_1\alpha_2)} = ||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\bar{\gamma}(i_{v'} j_{v'})}||^{-1}\,||\mathbf{B}_{a_0}^{\alpha_1\alpha_2;\gamma(i_v j_v)}||^{-1}$$
$$\times\frac{n_\gamma}{|H|}\sum_h \mathbb{R}_{i_{v'}i_v}^{\alpha_1}(hs)\mathbb{R}_{j_{v'}j_v}^{\alpha_2}(hs)\{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{a_0a_0}^{*}, \tag{III.92}$$

which are of course independent of the free index k.

Summarizing our results, CG coefficients of type III for corepresentations are obtained by simple formulas in terms of CG coefficients for the subgroup H. Hence, if the corresponding CG coefficients for H are known, the only problem is to compute the matrix elements of F.

SUMMARY

The aim of the first of this series of papers was to calculate for Kronecker products being composed of co-unirreps of type I only, CG coefficients in terms of ones for the normal (unitary) subgroup H. The first step of the present method requires one to determine a suitable CG matrix M which provides a decomposition of $\mathbb{R}^{\alpha_1\alpha_2}\downarrow H = R^{\alpha_1\alpha_2}$ into a direct sum of unirreps of H. Provided this has been done, CG coefficients for corepresentations can be determined as follows.

In the case of CG coefficients of type I, one has to compute the matrix elements (III.23) of the $M_{\alpha_1\alpha_2;\alpha}$-dimensional unitary matrix F, whose property to be symmetric should be utilized in any case. In order to obtain CG coefficients of type I for G, it suffices to find a unitary matrix B satisfying $FB^* = B$, since B connects CG coefficients for G with them for H.

CG coefficients of type II have to be computed as follows: Calculate the matrix elements (III.47) of the $2M_{\alpha_1\alpha_2;\beta}$-dimensional antisymmetric unitary matrix F and determine a unitary matrix B obeying $FB^* = BG^T$, where G is a special antisymmetric unitary matrix. Due to the special form of the matrix G, this equation can be rewritten as an eigenvalues equation, which leads to further simplifications. Any solution of the previous equation yields corresponding CG coefficients for G in terms of such ones for the normal subgroup [use Eq. (III.35)]. If $M_{\alpha_1\alpha_2;\beta} = 1$, a special solution is given by Eqs. (III.56) and (III.57).

In the last case, i.e., CG coefficients of type III, it is only necessary to calculate the matrix elements (III.89) of the $M_{\alpha_1\alpha_2;\gamma}$-dimensional unitary matrix F, since because of the special solution $B = \mathbb{1}_M$ and $C = F$ of Eq. (III.84), the corresponding CG coefficients for G are given by Eqs. (III.87), and (III.88) in terms of CG coefficients for H.

Concluding this paper, we summarize that we succeeded in deriving simple defining equations for those unitary transformations which link CG coefficients for corepresentations with ones for the unitary subgroup. Apart from the first two cases, we were able to give special solutions of these equations which lead us to a simple solution for the multiplicity problem without reference to a special magnetic group.

[1] A. Aviran and J. Zak, J. Math. Phys. **9**, 2138 (1968).
[2] A. Aviran and D.B. Litvin, J. Math. Phys. **14**, 1491 (1973).
[3] P. Rudra, J. Math. Phys. **15**, 2031 (1974).
[4] P.M. van den Broek, "Clebsch–Gordan Coefficients of Finite Magnetic Groups," preprint (1978).
[5] R. Dirl, J. Math. Phys. **20**, 659 (1979).
[6] R. Dirl, J. Math. Phys. **20**, 664 (1979).
[7] R. Dirl, J. Math. Phys. **20**, 679 (1979).
[8] R. Dirl, J. Math. Phys. **20**, 671 (1979).
[9] R. Dirl, J. Math. Phys. **20**, 1562 (1979).
[10] R. Dirl, J. Math. Phys. **20**, 1566 (1979).
[11] R. Dirl, J. Math. Phys. **20**, 1570 (1979).
[12] R. Dirl, J. Math. Phys. **20**, 1575 (1979).
[13] C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[14] L. Jansen and M. Boon, *Theory of Finite Groups. Applications in Physics* (North-Holland, Amsterdam, 1967).
[15] J.L. Birman, "Theory of Crystal Space Groups and Infra-red Raman Processes of Insulated Crystals," in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1974).

Clebsch–Gordan coefficients for corepresentations. I ⊗ II

R. Dirl

Institut für theoretische Physik, TU Wien A-1040 Wien, Karlsplatz 13, Austria

(Received 18 December 1978; accepted for publication 25 May 1979)

A general method is given to determine quite generally Clebsch–Gordan coefficients for corepresentations in terms of such ones of the normal subgroup, where the considered Kronecker products are composed of corepresentations of type I and II.

INTRODUCTION

This paper, being the second of a series of papers, deals with the problem of decomposing Kronecker products of co-unirreps of type I and II (as indicated by the title) into their irreducible constituents. The main point of the present method is to utilize the representation theory of the nonunitary group G insofar as to assume that, for the normal subgroup H of index 2, CG coefficients are known. This leads to the much easier task of determining unitary transformations which link CG coefficients for corepresentations with ones of the subgroup H.

The material is organized as follows: In Sec. I we state our problem, and write down the multiplicities which are needed for this subduction. Section II is devided into three parts according to the possible types of co-unirreps. For each case we derive simple defining equations for those unitary transformations which link CG coefficients for corepresentations with ones of the subgroup H. These unitary transformations will be determined for each type without reference to a special group G. This leads us to a special solution of the multiplicity problem.

I. MULTIPLICITIES FOR COREPRESENTATIONS

As already pointed out, this paper deals with the task of decomposing Kronecker products of the kind

$$\mathbb{R}^{\alpha_1\beta_2} = \{\mathbb{R}^{\alpha_1\beta_2}(g) = \mathbb{R}^{\alpha_1}(g) \otimes \mathbb{R}^{\beta_2}(g) : g \in G\}, \qquad \text{(I.1)}$$

into a direct sum of their irreducible constituents. Since $\mathbb{R}^{\alpha_1\beta_2}$ forms in general a reducible correpresentation of G, there must exist a unitary $2n_{\alpha_1}n_{\beta_2}$-dimensional matrix $W^{\alpha_1\beta_2} = W$ which leads to the desired decomposition of $\mathbb{R}^{\alpha_1\beta_2}$:

$$\begin{aligned} W^\dagger \mathbb{R}^{\alpha_1\beta_2}(g) W^g = & \sum_{\alpha\in A_\mathrm{I}} \oplus M_{\alpha_1\beta_2;\alpha}\mathbb{R}^{\alpha}(g) \\ & \oplus \sum_{\beta\in A_{\mathrm{II}}} \oplus M_{\alpha_1\beta_2;\beta}\mathbb{R}^{\beta}(g) \\ & \oplus \sum_{\gamma\in A_{\mathrm{III}}} \oplus M_{\alpha_1\beta_2;\gamma}\mathbb{R}^{\gamma}(g), \\ & \text{for all } g\in G. \end{aligned} \qquad \text{(I.2)}$$

The multiplicities $M_{...}$ occurring in Eq. (I.2) are well known[1] and read as

$$M_{\alpha_1\beta_2;\alpha} = 2m_{\alpha_1\beta_2;\alpha}, \qquad \text{(I.3)}$$

$$M_{\alpha_1\beta_2;\beta} = m_{\alpha_1\beta_2;\beta}, \qquad \text{(I.4)}$$

$$M_{\alpha_1\beta_2;\gamma} = 2m_{\alpha_1\beta_2;\gamma} = M_{\alpha_1\beta_2;\bar{\gamma}}, \qquad \text{(I.5)}$$

where the multiplicities $m_{...}$ refer to corresponding subductions with respect to the subgroup H.

II. CG COEFFICIENTS FOR COREPRESENTATIONS

The structure of the CG series (I.2) suggests that it is resonable to discuss the three different cases separately, depending on whether $\alpha\in A_\mathrm{I}$, or $\beta\in A_{\mathrm{II}}$, or $\gamma\in A_{\mathrm{III}}$. Before starting this discussion we remember that a unitary $n_{\alpha_1}n_{\beta_2}$-dimensional matrix M is known from the outset, which satisfies

$$\begin{aligned} M^\dagger R^{\alpha_1\beta_2}(h) M = & \sum_{\alpha\in A_\mathrm{I}} \oplus m_{\alpha_1\beta_2;\alpha} R^{\alpha}(h) \\ & \oplus \sum_{\beta\in A_{\mathrm{II}}} \oplus m_{\alpha_1\beta_2;\beta} R^{\beta}(h) \\ & \oplus \sum_{\gamma\in A_{\mathrm{III}}} \oplus m_{\alpha_1\beta_2;\gamma} \{R^{\gamma}(h) \oplus Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}, \\ & \text{for all } h\in H. \end{aligned} \qquad \text{(II.1)}$$

Thereby Kronecker products of unirreps of H are denoted by

$$R^{\alpha_1\beta_2} = \{R^{\alpha_1\beta_2}(h) = R^{\alpha_1}(h) \otimes R^{\beta_2}(h) : h\in H\}. \qquad \text{(II.2)}$$

In this connection we have to note that $R^{\alpha_1\beta_2}$ occurs twice in $\mathbb{R}^{\alpha_1\beta_2}$, which is in accordance with the formulas (I.3)–(I.5), i.e.,

$$\mathbb{R}^{\alpha_1\beta_2}\downarrow H = (\oplus 2) R^{\alpha_1\beta_2}. \qquad \text{(II.3)}$$

A. CG coefficients of type I

Due to our general procedure which has been described in Ref. 2 for projective unitary representations and generalized to corepresentations in Ref. 3, we can write the defining equations for CG coefficients of type I in the following way:

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{W}_k^{\alpha w} = \sum_l R^{\alpha}_{lk}(h)\mathbf{W}_l^{\alpha w}, \quad \text{for all } h\in H, \qquad \text{(II.4)}$$

$$\begin{aligned} &\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U^{\alpha}_{lk}\mathbf{W}_l^{\alpha w}, \\ &w = 1,2,\ldots,M_{\alpha_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \end{aligned} \qquad \text{(II.5)}$$

by utilizing the unitarity of the CG matrix W. Thereby, our notation has to be understood as

$$\begin{aligned} &\{\mathbf{W}_k^{\alpha w}\}_{i,bj} = \{\mathbf{W}_k^{\alpha_1\beta_2;\alpha w}\}_{i,bj} = W^{\alpha_1\beta_2}_{i,bj;\alpha w k}, \\ &\alpha\in A_\mathrm{I}, \quad w = 1,2,\ldots,M_{\alpha_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \\ &i = 1,2,\ldots,n_{\alpha_1}, \quad b = 1,2, \text{ and } j = 1,2,\ldots,n_{\beta_2}, \end{aligned} \qquad \text{(II.6)}$$

 0022-2488/80/050968-07$01.00

where the double index b,j is necessary by virtue of Eqs.. (I.21) and (I.22) of Ref. 3 and indicates that the matrix W is $2n_{\alpha_1}n_{\beta_2}$-dimensional. Equations (II.4) allow one to interpret the columns of the CG matrix W as H-adapted vectors of a $2n_{\alpha_1}n_{\beta_2}$-dimensional Euclidean space $\mathscr{W}^{\alpha_1\beta_2}$, but which have to satisfy additionally Eq. (II.5). As already pointed out in Ref. 3, unitarity of W requires orthonormality of the vectors $\mathbf{W}_k^{\alpha w}$ with respect to each index. Whereas the orthogonality of these vectors with respect to α and k is automatically guaranteed because of their transformation properties with respect to H, the orthogonality with respect to the multiplicity index w can only be achieved by further manipulations.

Obviously, the vectors

$$\mathbf{W}_k^{\alpha w};\quad w=1,2,...,M_{\alpha_1\beta_2;\alpha}\ ,\quad k=1,2,...,n_\alpha \tag{II.7}$$

form an orthonormal, H-adapted basis of

$$\mathscr{W}^{\alpha_1\beta_2;\alpha}=\sum_i \mathbb{E}_{ii}^{\alpha}\mathscr{W}^{\alpha_1\beta_2},$$

$$\dim\mathscr{W}^{\alpha_1\beta_2;\alpha}=n_\alpha M_{\alpha_1\beta_2;\alpha}, \tag{II.8}$$

where the units $\mathbb{E}_{ij}^{\alpha}$ are defined as usually, but can be written in this case as the direct sum of the matrix E_{ij}^{α} which refers to the Kronecker product $R^{\alpha_1\beta_2}$, i.e.,

$$\mathbb{E}_{ij}^{\alpha}=\mathbb{E}_{ij}^{\alpha_1\beta_2;\alpha}=(\oplus 2)E_{ij}^{\alpha_1\beta_2;\alpha}=(\oplus 2)E_{ij}^{\alpha}, \tag{II.9}$$

$$E_{ij}^{\alpha}=\frac{n_\alpha}{|H|}\sum_h R_{ij}^{\alpha *}(h)R^{\alpha_1\beta_2}(h). \tag{II.10}$$

By means of the following definitions:

$$\{\mathbf{Q}_k^{\alpha v a}\}_{i,bj}=\delta_{ab}\{\mathbf{M}_k^{\alpha v}\}_{ij};\quad a=1,2, \tag{II.11}$$

we can introduce a further H-adapted basis of $\mathscr{W}^{\alpha_1\beta_2;\alpha}$, namely,

$$\mathbf{Q}_k^{\alpha v a};$$

$$a=1,2,;\quad v=1,2,...,m_{\alpha_1\beta_2;\alpha}\ ,\quad k=1,2,...,n_\alpha\ , \tag{II.12}$$

where the vectors $\mathbf{M}_k^{\alpha v}$ represent columns of the CG matrix M. The vectors (II.12) are orthonormal with respect to each index. In this connection we have to note that the multiplicity index v originates from subductions with respect to H and should therefore not be confused with w. Although the transformation law

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{Q}_k^{\alpha v a}=\sum_l R_{lk}^{\alpha}(h)\mathbf{Q}_l^{\alpha v a},\quad \text{for all } h\in H, \tag{II.13}$$

is automatically satisfied, we cannot except that these vectors are already a solution of Eq. (II.5).

Hence, our problem is now reduced to the task of determining a unitary matrix which link the elements of the bases (II.7) and (II.12). Due to Schur's lemma with respect to H, this unitary transformation may not depend on the free index k, i.e.,

$$\mathbf{W}_k^{\alpha w}=\sum_{av}B_{av;w}\mathbf{Q}_k^{\alpha v a}, \tag{II.14}$$

$$\mathbf{Q}_k^{\alpha v a}=\sum_w B^*_{av;w}\mathbf{W}_k^{\alpha w},\quad k=1,2,...,n_\alpha. \tag{II.15}$$

In order to be able to determine a suitable $M_{\alpha_1\beta_2;\alpha}$-dimensional unitary matrix B, let us proceed in the same way as in Ref. 3. By using

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{Q}_k^{\alpha v a *}=\sum_l U_{lk}^{\alpha}\sum_{a'v'}F_{a'v';av}\mathbf{Q}_l^{\alpha v' a'}, \tag{II.16}$$

where

$$F_{a'v';av}=\{BB^T\}_{av;a'v'}=\sum_w B_{av;w}B_{a'v';w},$$

$$a,a'=1,2\ \text{and}\ v,v'=1,2,...,m_{\alpha_1\beta_2;\alpha}, \tag{II.17}$$

we obtain

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}_k^{\alpha w *}=\sum_l U_{lk}^{\alpha}\sum_{w'}\{B^\dagger F B^*\}_{w'w}\mathbf{W}_l^{\alpha w'}. \tag{II.18}$$

Obviously, any unitary matrix B satisfying

$$FB^*=B,\quad \text{with}\quad BB^\dagger=B^\dagger B=\mathbb{1}_M, \tag{II.19}$$

leads immediately to the desired CG coefficients for corepresentations by inserting the matrix elements of B into Eq. (II.14). The matrix F is not only symmetric, which follows from Eq. (II.17), but also unitary, which can be verified with the aid of Eq. (II.13) by taking the special group element $h=s^2$ and utilizing Eq. (I.10) of Ref. 3. This lead us to

$$FF^*=\mathbb{1}_M=\mathbb{1}_{2m}, \tag{II.20}$$

which proves our assertion.

In order to be able to solve Eq. (II.19), it is necessary to compute the matrix elements of F, at which of course its property to be symmetric is of practical use. The matrix elements follow directly from Eq. (II.16):

$$\left\langle \mathbf{Q}_k^{\alpha v' a'},\mathbb{R}^{\alpha_1\beta_2}(s)\left\{\sum_l U_{kl}^{\alpha}\mathbf{Q}_l^{\alpha v a}\right\}^*\right\rangle=F_{a'v';av}\ ,$$

$$a,a'=1,2,\quad v,v'=1,2,...,m_{\alpha_1\beta_2;\alpha}, \tag{II.21}$$

whose values must be independent of the free index k. This can be readily verified with the aid of

$$\mathbb{R}^{\alpha_1\beta_2}(s)^\dagger\mathbb{E}_{ij}^{\alpha}\ \mathbb{R}^{\alpha_1\beta_2}(s)=\sum_{kl}U_{ik}^{\alpha *}U_{jl}^{\alpha}\ \mathbb{E}_{kl}^{\alpha *}. \tag{II.22}$$

Because of Eq. (II.11) we write the matrix elements as

$$F_{a'v';av}=(-1)^{\Delta(a')}\delta_{a',a+1}F_{v'v}^{\alpha(\alpha_1\beta_2)}, \tag{II.23}$$

$$F_{v'v}^{\alpha(\alpha_1\beta_2)}=\left\langle \mathbf{M}_k^{\alpha v'},U^{\alpha_1}\otimes U^{\beta_2}\left\{\sum_l U_{kl}^{\alpha}\mathbf{M}_l^{\alpha v}\right\}^*\right\rangle, \tag{II.24}$$

at which the scalar product on the right-hand side of Eq. (II.24) refers to the corresponding subspace. Obviously, the matrix elements (II.24) must also be independent of the free index k. This can be shown by means of

$$\{U^{\alpha_1}\otimes U^{\beta_2}\}^\dagger E_{ij}^{\alpha}\ \{U^{\alpha_1}\otimes U^{\beta_2}\}=\sum_{kl}U_{ik}^{\alpha *}U_{jl}^{\alpha}\ E_{kl}^{\alpha *}, \tag{II.25}$$

and the well-known properties of units. In matrix notation, F reads as

$$F=\begin{bmatrix} 0 & F^{\alpha(\alpha_1\beta_2)} \\ -F^{\alpha(\alpha_1\beta_2)} & 0\end{bmatrix}, \tag{II.26}$$

where $F^{\alpha(\alpha_1\beta_2)}$ denotes a $m_{\alpha_1\beta_2;\alpha}$-dimensional matrix. Since F is a symmetric unitary matrix, it follows immediately that

$$F^{\alpha(\alpha_1\beta_2)^T}=-F^{\alpha(\alpha_1\beta_2)}, \tag{II.27}$$

$$F^{\alpha(\alpha_1\beta_2)}F^{\alpha(\alpha_1\beta_2)*}=-\mathbb{1}_m, \tag{II.28}$$

and that $F^{\alpha(\alpha_1\beta_2)}$ must be a antisymmetric unitary matrix. Furthermore, Eqs. (II.27) and (II.28)

$$m_{\alpha_1\beta_2;\alpha} > 1, \quad \text{if } m_{\alpha_1\beta_2;\alpha} \neq 0, \tag{II.29}$$

which is an nontrivial byproduct of these considerations.

Assuming that the CG coefficients for H can be computed by means of the method given in Ref. 2, they have the form

$$\{\mathbf{M}_k^{\alpha v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\beta_2;\alpha(i_v j_v)}\}_{ij} = \|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\alpha(i_v j_v)}\|^{-1}\frac{n_\alpha}{|H|} \times \sum_h R_{ii_v}^{\alpha_1}(h) R_{jj_v}^{\beta_2}(h) R_{ka_0}^{\alpha *}(h),$$

$$v = 1,2,\ldots,m_{\alpha_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha. \tag{II.30}$$

Inserting these special values into Eq. (II.24), we obtain after a straightforward calculation

$$F_{v'v}^{\alpha(\alpha_1\beta_2)} = \|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\alpha(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\alpha(i_v j_v)}\|^{-1} \times \frac{n_\alpha}{|H|}\sum_h \mathbb{R}_{i_{v'} i_v}^{\alpha_1}(hs)\{R^{\beta_2}(h)U^{\beta_2}\}_{j_{v'} j_v} \mathbb{R}_{a_0 a_0}^{\alpha *}(hs). \tag{II.31}$$

Now let us return to the task of determining a unitary matrix B which satisfies Eq. (II.19). If taking the special structure of Eq. (II.26) into account, it is suggestive to make the following ansatz:

$$B = \begin{bmatrix} \mathbf{A} & F^{\alpha(\alpha_1\beta_2)}\mathbf{B}^* \\ -F^{\alpha(\alpha_1\beta_2)}\mathbf{A}^* & \mathbf{B} \end{bmatrix}, \tag{II.32}$$

where $\mathbf{A}$ and $\mathbf{B}$ shall be proportional by numerical factors to unitary $m_{\alpha_1\beta_2;\alpha}$-dimensional matrices, but otherwise arbitrary matrices. The last condition implies no loss of generality. Furthermore, it can be easily shown that, for every pair $\mathbf{A}$ and $\mathbf{B}$, the matrix (II.32) is a solution of Eq. (II.19). Since B is required to be unitary, we obtain as conditions

$$\mathbf{A}\mathbf{A}^\dagger + \mathbf{B}\mathbf{B}^\dagger = \mathbb{1}_m, \tag{II.33}$$

$$(\mathbf{A}\mathbf{A}^\dagger)F^{\alpha(\alpha_1\beta_2)\dagger} = F^{\alpha(\alpha_1\beta_2)*}(\mathbf{B}\mathbf{B}^T)^*, \tag{II.34}$$

which suggest that we choose as special solutions of Eqs. (II.33) and (II.34) the following matrices:

$$\mathbf{A} = \frac{i}{\sqrt{2}}\mathbb{1}_m \quad \text{and} \quad \mathbf{B} = \frac{1}{\sqrt{2}}F^{\alpha(\alpha_1\beta_2)}. \tag{II.35}$$

Hence, the corresponding unitary matrix B reads as

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbb{1}_m & -\mathbb{1}_m \\ iF^{\alpha(\alpha_1\beta_2)} & F^{\alpha(\alpha_1\beta_2)} \end{bmatrix}, \tag{II.36}$$

which allows one to identify the multiplicity index w with the pair (a,v), i.e.,

$$w = (a,v), \quad a = 1,2 \quad \text{and} \quad v = 1,2,\ldots,m_{\alpha_1\beta_2;\alpha}. \tag{II.37}$$

The corresponding CG coefficients are immediately obtained from Eqs. (II.14):

$$\mathbf{W}_k^{\alpha(1v)} = \frac{i}{\sqrt{2}}\left\{\mathbf{Q}_k^{\alpha v 1} + \sum_{v'} F_{v'v}^{\alpha(\alpha_1\beta_2)}\mathbf{Q}_k^{\alpha v' 2}\right\}, \tag{II.38}$$

$$\mathbf{W}_k^{\alpha(2v)} = \frac{1}{\sqrt{2}}\left\{-\mathbf{Q}_k^{\alpha v 1} + \sum_{v'} F_{v'v}^{\alpha(\alpha_1\beta_2)}\mathbf{Q}_k^{\alpha v' 2}\right\},$$

$$v = 1,2,\ldots,m_{\alpha_1\beta_2;\alpha}. \tag{II.39}$$

To summarize our results, we have shown that CG coefficients of type I for corepresentations can be traced back by simple formulas to CG coefficients for the subgroup H. Hence, if the corresponding CG coefficients for H are known, the only problem is to compute the antisymmetric $m_{\alpha_1\beta_2;\alpha}$-dimensional unitary matrix $F^{\alpha(\alpha_1\beta_2)}$. Thereby, we obtained as a byproduct that the multiplicity $m_{\alpha_1\beta_2;\alpha}$ must be larger than one, if $m_{\alpha_1\beta_2;\alpha} \neq 0$, which however cannot be verified by means of the representation theory for the subgroup H.

B. CG coefficients of type II

The defining equations for CG coefficients of type II for corepresentations are rewritten as

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{W}_{dk}^{\beta w} = \sum_l R_{lk}^{\beta}(h)\mathbf{W}_{dl}^{\beta w}, \quad \text{for all } h \in H, \tag{II.40}$$

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}_{dk}^{\beta w *} = (-1)^{\Delta(d+1)}\sum_l U_{lk}^{\beta}\mathbf{W}_{d+1,l}^{\beta w},$$

$$w = 1,2,\ldots,M_{\alpha_1\beta_2;\beta}, \quad d = 1,2, \quad \text{and} \quad k = 1,2,\ldots,n_\beta, \tag{II.41}$$

by using the unitarity of the CG matrix W, which allow one to consider the columns of the CG matrix as H-adapted vectors of $\mathscr{W}^{\alpha_1\beta_2}$. The abbreviated notation used in Eqs. (II.40) and (II.41) reads in more detail as

$$\{\mathbf{W}_{dk}^{\beta w}\}_{i,bj} = \{\mathbf{W}_{dk}^{\alpha_1\beta_2;\beta w}\}_{i,bj} = W_{i,bj;\beta w dk}^{\alpha_1\beta_2},$$

$$\beta \in A_{\text{II}}, \quad w = 1,2,..,M_{\alpha_1\beta_2;\beta} \quad d = 1,2, \text{ and } k = 1,2,\ldots,n_\beta,$$

$$i = 1,2,\ldots,n_{\alpha_1}, \quad b = 1,2, \text{ and } j = 1,2,\ldots,n_{\beta_2}, \tag{II.42}$$

where the double index b,j has been already explained in the previous part of this paper.

Provided that the CG matrix W is unitary, the vectors

$$\mathbf{W}_{dk}^{\beta w}, \quad w = 1,2,\ldots,M_{\alpha_1\beta_2;\beta}, \quad d = 1,2, \text{ and } k = 1,2,\ldots,n_\beta \tag{II.43}$$

form an orthonormal H-adapted basis of

$$\mathscr{W}^{\alpha_1\beta_2;\beta} = \sum_i \mathbb{E}_{ii}^{\beta}\mathscr{W}^{\alpha_1\beta_2},$$

$$\dim \mathscr{W}^{\alpha_1\beta_2;\beta} = 2n_\beta M_{\alpha_1\beta_2;\beta}, \tag{II.44}$$

where the corresponding units $\mathbb{E}_{ij}^{\beta}$ are defined as usually and decompose into the direct sum of the submatrices E_{ij}^{β}:

$$\mathbb{E}_{ij}^{\beta} = \mathbb{E}_{ij}^{\alpha_1\beta_2;\beta} = (\oplus 2)E_{ij}^{\alpha_1\beta_2;\beta} = (\oplus 2)E_{ij}^{\beta}, \tag{II.45}$$

$$E_{ij}^{\beta} = \frac{n_\beta}{|H|}\sum_h R_{ij}^{\beta *}(h)R^{\alpha_1\beta_2}(h). \tag{II.46}$$

An other orthonormal basis of $\mathscr{W}^{\alpha_1\beta_2;\beta}$ can be introduced by means of the following definitions:

$$\{\mathbf{Q}_k^{\beta v a}\}_{i,bj} = \delta_{ab}\{\mathbf{M}_k^{\beta v}\}_{ij}, \quad a = 1,2, \tag{II.47}$$

where the vectors $\mathbf{M}_k^{\beta v}$ are the corresponding columns of the CG matrix M satisfying Eq. (II.1). The vectors

$$\mathbf{Q}_k^{\beta v a}, \quad a = 1,2, \quad v = 1,2,\ldots,m_{\alpha_1\beta_2;\beta}, \quad k = 1,2,\ldots,n_\beta \tag{II.48}$$

are orthonormal and transform according to

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{Q}_k^{\beta va} = \sum_l R^{\beta}_{lk}(h)\mathbf{Q}_l^{\beta va},$$
$$\text{for all } h\in H, \qquad \text{(II.49)}$$

but are in general not a solution of Eq. (II.41).

In order to be able to satisfy Eq. (II.41), we remember that the elements of the bases (II.43) and (II.48) must be connected by a unitary transformation which may not depend on the free index k in accordance to Schur's lemma with respect to H:

$$\mathbf{W}_{dk}^{\beta w} = \sum_{av} B_{av;dw}\mathbf{Q}_k^{\beta va}, \qquad \text{(II.50)}$$

$$\mathbf{Q}_k^{\beta va} = \sum_{dw} B^*_{av;dw}\mathbf{W}_{dk}^{\beta w}, \quad k = 1,2,...,n_\beta. \qquad \text{(II.51)}$$

Since the vectors $\mathbf{Q}_k^{\beta va}$ are orthonormalized, the corresponding vectors $\mathbf{W}_{dk}^{\beta w}$ have the same property if the matrix B is unitary.

Hence, the problem is now reduced to the task of determining a $2M_{\alpha_1\beta_2;\beta}$-dimensional unitary matrix B, so that the corresponding vectors (II.50) satisfy the conditions (II.41). For this reason we derive

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{Q}_k^{\beta va*} = \sum_l U^{\beta}_{lk} \sum_{a'v'} F_{a'v';av}\mathbf{Q}_l^{\beta v'a'}, \qquad \text{(II.52)}$$

where we have introduced the notations

$$G_{dw;d'w'} = (-1)^{\Delta(d+1)}\delta_{d',d+1}\delta_{ww'},$$
$$d,d' = 1,2, \quad w,w' = 1,2,...,M_{\alpha_1\beta_2;\beta}, \qquad \text{(II.53)}$$
$$F_{a'v';av} = \{BGB^T\}_{av;a'v'}$$
$$= \sum_{dw} B_{av;dw}(-1)^{\Delta(d+1)}B_{a'v';d+1,w},$$
$$a,a' = 1,2; \quad v,v' = 1,2,...,m_{\alpha_1\beta_2;\beta}. \qquad \text{(II.54)}$$

Equations (II.52) allow one to transcribe Eq. (II.41) as follows:

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}_{dk}^{\beta w*} = \sum_l U^{\beta}_{lk} \sum_{d'w'} \{B^\dagger FB^*\}_{d'w';dw}\mathbf{W}_{d'l}^{\beta w'}, \qquad \text{(II.55)}$$

which lead us immediately to

$$BG^T = FB^*. \qquad \text{(II.56)}$$

Now, if we can find a unitary matrix B satisfying Eq. (II.56), the corresponding CG coefficients of type II are readily obtained from Eq. (II.50). Before attacking this problem let us remark that F must be an antisymmetric unitary matrix. This can be easily verified by considering Eq. (II.49) for the special group element $h = s^2$ and using Eq. (I.14) of Ref. 3. Thereby, we obtain

$$FF^* = -\mathbb{1}_{2M}, \qquad \text{(II.57)}$$

which proves our assertion.

In order to be able to solve Eq. (II.56), it is necessary to compute the matrix elements of F which are uniquely fixed through Eq. (II.52):

$$\left\langle \mathbf{Q}_k^{\beta v'a'}, \mathbb{R}^{\alpha_1\beta_2}(s)\left\{\sum_l U^{\beta}_{kl}\mathbf{Q}_l^{\beta va}\right\}^*\right\rangle = F_{a'v';av},$$
$$a,a' = 1,2, \quad v,v' = 1,2,...,m_{\alpha_1\beta_2;\beta}. \qquad \text{(II.58)}$$

These matrix elements must be independent of the free index k, which can be shown by means of

$$\mathbb{R}^{\alpha_1\beta_2}(s)^\dagger\mathbb{E}^{\beta}_{ij}\,\mathbb{R}^{\alpha_1\beta_2}(s) = \sum_{kl} U^{\beta*}_{ik}U^{\beta}_{jl}\,\mathbb{E}^{\beta*}_{kl}. \qquad \text{(II.59)}$$

Corresponding to the special structure of the vectors $\mathbf{Q}_k^{\beta va}$, the matrix elements (II.58) turn out to be

$$F_{a'v';av} = (-1)^{\Delta(a')}\delta_{a',a+1}F^{\beta(\alpha_1\beta_2)}_{v'v}, \qquad \text{(II.60)}$$

$$F^{\beta(\alpha_1\beta_2)}_{v'v} = \left\langle \mathbf{M}_k^{\beta v'}, U^{\alpha_1}\otimes U^{\beta_2}\left\{\sum_l U^{\beta}_{kl}\mathbf{M}_l^{\beta v}\right\}^*\right\rangle, \qquad \text{(II.61)}$$

where of course the matrix elements (II.61) may also not depend on the index k. Thereby, the scalar product which occurs on the right-hand side of Eq. (II.61) refers to the corresponding subspace. The previous proposition can be shown by means of

$$\{U^{\alpha_1}\otimes U^{\beta_2}\}^\dagger E^{\beta}_{ij}\{U^{\alpha_1}\otimes U^{\beta_2}\} = \sum_{kl} U^{\beta*}_{ik}U^{\beta}_{jl}\,E^{\beta*}_{kl} \qquad \text{(II.62)}$$

and the fact that the vectors $\mathbf{M}_k^{\beta v}$ are H adapted. Because of Eq. (II.60), the matrix F takes the form

$$F = \begin{bmatrix} 0 & F^{\beta(\alpha_1\beta_2)} \\ -F^{\beta(\alpha_1\beta_2)} & 0 \end{bmatrix}, \qquad \text{(II.63)}$$

where the $m_{\alpha_1\beta_2;\beta}$-dimensional submatrix $F^{\beta(\alpha_1\beta_2)}$ must be symmetric and unitary, since F is itself antisymmetric and unitary, i.e.,

$$F^{\beta(\alpha_1\beta_2)^T} = F^{\beta(\alpha_1\beta_2)}, \qquad \text{(II.64)}$$

$$F^{\beta(\alpha_1\beta_2)}F^{\beta(\alpha_1\beta_2)*} = \mathbb{1}_m. \qquad \text{(II.65)}$$

If the corresponding columns of the CG matrix M can be determined with the aid of the method given in Ref. 2, there components are given by

$$\{\mathbf{M}_k^{\beta v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\beta_2;\beta(i_vj_v)}\}_{ij}$$
$$= \|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\beta(i_vj_v)}\|^{-1}\frac{n_\beta}{|H|}$$
$$\times\sum_h R^{\alpha_1}_{ii_v}(h)R^{\beta_2}_{jj_v}(h)R^{\beta*}_{ka_0}(h),$$
$$v = 1,2,...,m_{\alpha_1\beta_2;\beta}, \quad k = 1,2,...,n_\beta. \qquad \text{(II.66)}$$

Inserting these special values into Eq. (II.61), we obtain after a straightforward calculation

$$F^{\beta(\alpha_1\beta_2)}_{v'v} = \|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\beta(i_{v'}j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\beta(i_vj_v)}\|^{-1}$$
$$\times\frac{n_\beta}{|H|}\sum_h \mathbb{R}^{\alpha_1}_{i_{v'}i_v}(hs)\{R^{\beta_2}(h)U^{\beta_2}\}_{j_{v'}j_v}$$
$$\times\{R^{\beta}(h)U^{\beta}\}^*_{a_0a_0}. \qquad \text{(II.67)}$$

Apart from these special values for Eq. (II.61), we are now in the position to determine unitary matrices B which satisfy Eq. (II.56). For this purpose we define by

$$\{\mathbf{B}^{d,w}\}_{av} = B_{av;dw},$$
$$a = 1,2, \text{ and } v = 1,2,...,m_{\alpha_1\beta_2;\beta},$$
$$d = 1,2, \text{ and } w = 1,2,...,M_{\alpha_1\beta_2;\beta} \qquad \text{(II.68)}$$

vectors representing the columns of B. Consequently, Eq. (II.56) can be replaced by

$$F\mathbf{B}^{d,w*} = (-1)^{\Delta(d+1)}\mathbf{B}^{d+1,w};$$
$$d = 1,2 \text{ and } w = 1,2,...,...,M_{\alpha_1\beta_2;\beta}, \qquad \text{(II.69)}$$

which show that if we choose the vectors $\mathbf{B}^{1,w}$,

$w=1,2,\ldots,M_{\alpha_1\beta_2;\beta}$, the corresponding vectors $\mathbf{B}^{2,w}$, $w=1,2,\ldots,M_{\alpha_1\beta_2;\beta}$ are automatically fixed through Eq. (II.69). In order to achieve the requirement that B is unitary, we must require that the vectors are orthonormal with respect to both indices d and w. Therefore, it is obvious to make the following ansatz:

$$\{\mathbf{B}^{1,w}\}_{av}=\delta_{a1}\delta_{vw}, \tag{II.70}$$

which implies

$$\{\mathbf{B}^{2,w}\}_{av}=-\{F\mathbf{B}^{1,w*}\}_{av}=\delta_{a2}F^{\beta(\alpha_1\beta_2)}_{vw}, \tag{II.71}$$

in matrix notation

$$B=\begin{bmatrix}\mathbb{1}_m & 0\\ 0 & F^{\beta(\alpha_1\beta_2)}\end{bmatrix}. \tag{II.72}$$

Now it is easy to verify that B is unitary and a solution of Eq. (II.56). Hence, the corresponding CG coefficients of type II are given by

$$\mathbf{W}^{\beta w}_{1k}=\mathbf{Q}^{\beta w1}_{k}, \tag{II.73}$$

$$\mathbf{W}^{\beta w}_{2k}=\sum_v F^{\beta(\alpha_1\beta_2)}_{vw}\mathbf{Q}^{\beta v2}_{k},\quad w=1,2,\ldots,m_{\alpha_1\beta_2;\beta}, \tag{II.74}$$

which make it obvious that we can choose as multiplicity index w the index v refering to subductions with respect to H, i.e.,

$$w=v,\quad v=1,2,\ldots,m_{\alpha_1\beta_2;\beta}\,. \tag{II.75}$$

Consequently, we have shown that CG coefficients of type II for corepresentations can be traced back by simple formulas to CG coefficients for H, at which the only problem is to compute the symmetric unitary submatrix $F^{\beta(\alpha_1\beta_2)}$.

C. CG coefficients of type III

Due to our procedure we rewrite the defining equations for CG coefficients of type III as follows:

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{W}^{\gamma w}_{1k}=\sum_l R^{\gamma}_{lk}(h)\mathbf{W}^{\gamma w}_{1l}, \tag{II.76}$$

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{W}^{\gamma w}_{2k}=\sum_l\{Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}_{lk}\mathbf{W}^{\gamma w}_{2l},$$

$$\text{for all } h\in H, \tag{II.77}$$

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}^{\gamma w*}_{1k}=\mathbf{W}^{\gamma w}_{2k}, \tag{II.78}$$

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}^{\gamma w*}_{2k}=\sum_l R^{\gamma}_{lk}(s^2)\mathbf{W}^{\gamma w}_{1l},$$

$$w=1,2,\ldots,M_{\alpha_1\beta_2;\gamma}\quad k=1,2,\ldots,n_\gamma, \tag{II.79}$$

where the symbols have to be understood as

$$\{\mathbf{W}^{\gamma w}_{dk}\}_{i,bj}=\{\mathbf{W}^{\alpha_1\beta_2;\gamma w}_{dk}\}_{i,bj}=W^{\alpha_1\beta_2}_{i,bj;\gamma w dk},$$

$$\gamma\in A_{\rm III},\quad w=1,2,\ldots,M_{\alpha_1\beta_2;\gamma},$$

$$d=1,2,\ \text{and}\ k=1,2,\ldots,n_\gamma,$$

$$i=1,2,\ldots,n_{\alpha_1},\quad b=1,2,\quad\text{and}\quad j=1,2,\ldots,n_{\beta_2}. \tag{II.80}$$

Provided that the CG matrix W is unitary, the vectors

$$\mathbf{W}^{\gamma w}_{dk},$$

$$w=1,2,\ldots,M_{\alpha_1\beta_2;\gamma},\quad d=1,2,\quad\text{and}\quad k=1,2,\ldots,n_\gamma, \tag{II.81}$$

define an orthonormal H-adapted basis of

$$\mathscr{W}^{\alpha_1\beta_2;\gamma}=\sum_i\{\mathbb{E}^{\gamma}_{ii}+\mathbb{E}^{\bar\gamma}_{ii}\}\mathscr{W}^{\alpha_1\beta_2},$$

$$\dim\mathscr{W}^{\alpha_1\beta_2;\gamma}=2n_\gamma M_{\alpha_1\beta_2;\gamma}, \tag{II.82}$$

where the corresponding units are defined as usually and decompose into a direct sum of the submatrices E^{γ}_{ij} and $E^{\bar\gamma}_{ij}$:

$$\mathbb{E}^{\gamma}_{ij}=\mathbb{E}^{\alpha_1\beta_2;\gamma}_{ij}=(\oplus 2)E^{\alpha_1\beta_2;\gamma}_{ij}=(\oplus 2)E^{\gamma}_{ij}, \tag{II.83}$$

$$E^{\gamma}_{ij}=\frac{n_\gamma}{|H|}\sum_h R^{\gamma *}_{ij}(h)R^{\alpha_1\beta_2}(h), \tag{II.84}$$

$$\mathbb{E}^{\bar\gamma}_{ij}=\mathbb{E}^{\alpha_1\beta_2;\bar\gamma}_{ij}=(\oplus 2)E^{\alpha_1\beta_2;\gamma}_{ij}=(\oplus 2)E^{\bar\gamma}_{ij}, \tag{II.85}$$

$$E^{\bar\gamma}_{ij}=\frac{n_\gamma}{|H|}\sum_h\{Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}^*_{ij}R^{\alpha_1\beta_2}(h). \tag{II.86}$$

By means of the following definitions:

$$\{\mathbf{Q}^{\gamma va}_{k}\}_{i,bj}=\delta_{ab}\{\mathbf{M}^{\gamma v}_{k}\}_{ij},\quad a=1,2, \tag{II.87}$$

$$\{\mathbf{Q}^{\bar\gamma va}_{k}\}_{i,bj}=\delta_{ab}\{\mathbf{M}^{\bar\gamma v}_{k}\}_{ij},\quad a=1,2, \tag{II.88}$$

we introduce a further orthonormal basis of $\mathscr{W}^{\alpha_1\beta_2;\gamma}$, namely,

$$\mathbf{Q}^{\gamma va}_{k},\quad a=1,2,\quad v=1,2,\ldots,m_{\alpha_1\beta_2;\gamma},\quad k=1,2,\ldots,n_\gamma, \tag{II.89}$$

$$\mathbf{Q}^{\bar\gamma va}_{k},\quad a=1,2,\quad v=1,2,\ldots,m_{\alpha_1\beta_2;\gamma},\quad k=1,2,\ldots,n_\gamma. \tag{II.90}$$

Although they transform according to

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{Q}^{\gamma va}_{k}=\sum_l R^{\gamma}_{lk}(h)\mathbf{Q}^{\gamma va}_{l}, \tag{II.91}$$

$$\mathbb{R}^{\alpha_1\beta_2}(h)\mathbf{Q}^{\bar\gamma va}_{k}=\sum_l\{Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}_{lk}\mathbf{Q}^{\bar\gamma va}_{l},$$

$$\text{for all } h\in H, \tag{II.92}$$

we cannot expect that they are already a solution of Eqs. (II.78) and (II.79).

Nevertheless, the elements of the bases (II.81), (II.89), and (II.90) must be linked by unitary transformations which are due to Schur's lemma with respect to H independent of the index k:

$$\mathbf{W}^{\gamma w}_{1k}=\sum_{av}B_{av;w}\mathbf{Q}^{\gamma va}_{k}, \tag{II.93}$$

$$\mathbf{Q}^{\gamma va}_{k}=\sum_w B^*_{av;w}\mathbf{W}^{\gamma w}_{1k},\quad k=1,2,\ldots,n_\gamma, \tag{II.94}$$

$$\mathbf{W}^{\gamma w}_{2k}=\sum_{av}C_{av;w}\mathbf{Q}^{\bar\gamma va}_{k}, \tag{II.95}$$

$$\mathbf{Q}^{\bar\gamma va}_{k}=\sum_w C^*_{av;w}\mathbf{W}^{\gamma w}_{2k},\quad k=1,2,\ldots,n_\gamma. \tag{II.96}$$

Hence, we are now confronted with the less complicated task to determine $M_{\alpha_1\beta_2;\gamma}$-dimensional unitary matrices B and C so that the corresponding vectors (II.93) and (II.95) are solutions of Eqs. (II.78) and (II.79), respectively. Let us proceed as in the previous paper by deriving

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{Q}^{\gamma va*}_{k}=\sum_{a'v'}F_{a'v';av}\mathbf{Q}^{\bar\gamma v'a'}_{k}, \tag{II.97}$$

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{Q}^{\bar\gamma va*}_{k}=\sum_l R^{\gamma}_{lk}(s^2)\sum_{a'v'}F_{av;a'v'}\mathbf{Q}^{\gamma v'a'}_{l}, \tag{II.98}$$

where

$$F_{a'v';av}=\{CB^T\}_{a'v';av}=\sum_w C_{a'v';w}B_{av;w},$$

$$a,a'=1,2,\quad\text{and}\quad v,v'=1,2,\ldots,m_{\alpha_1\beta_2;\gamma}. \tag{II.99}$$

These equations allow one to transform Eqs. (II.78) and (II.79) as follows:

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}_{1k}^{\gamma w*} = \sum_{w'} \{C^\dagger F B^*\}_{w'w} \mathbf{W}_{2k}^{\gamma w'}, \tag{II.100}$$

$$\mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{W}_{2k}^{\gamma w*} = \sum_{l} R^{\gamma}_{lk}(s^2) \times \sum_{w'} \{C^\dagger F B^*\}^T_{w'w} \mathbf{W}_{1l}^{\gamma w'}, \tag{II.101}$$

respectively, which lead immediately to

$$C^\dagger F B^* = \mathbf{1}_M = \mathbf{1}_{2m}. \tag{II.102}$$

Before solving this equation, let us remark that F is a unitary matrix, which can be readily verified with the aid of Eq. (II.91) by setting $h = s^2$ and utilizing Eqs. (II.97) and (II.98).

Now it is obvious to choose

$$B = \mathbf{1}_M \Longleftrightarrow C = F \tag{II.103}$$

as a special solution of Eq. (II.102) which yield immediately the corresponding CG coefficients of type III, namely,

$$\mathbf{W}_{1k}^{\gamma(av)} = \mathbf{Q}_k^{\gamma va}, \tag{II.104}$$

$$\mathbf{W}_{2k}^{\gamma(av)} = \sum_{a'v'} F_{a'v';av} \mathbf{Q}_k^{\bar{\gamma}v'a'},$$

$$a = 1,2 \text{ and } v = 1,2,\ldots,m_{\alpha_1\beta_2;\gamma}. \tag{II.105}$$

Consequently, the multiplicity index w can be identified with the double index (a,v):

$$w = (a,v), \quad a = 1,2 \quad \text{and} \quad v = 1,2,\ldots,m_{\alpha_1\beta_2;\gamma}. \tag{II.106}$$

Now there remains the task of computing the matrix elements of the unitary matrix F. This can be done by means of

$$\langle \mathbf{Q}_k^{\bar{\gamma}v'a'}, \mathbb{R}^{\alpha_1\beta_2}(s)\mathbf{Q}_k^{\gamma va*}\rangle = F_{a'v';av},$$

$$a,a' = 1,2, \quad v,v' = 1,2,\ldots,m_{\alpha_1\beta_2;\gamma}. \tag{II.107}$$

These matrix elements must be independent of the free index k. This can be readily verified with the aid of

$$\mathbb{R}^{\alpha_1\beta_2}(s)^\dagger \mathbb{E}_{ij}^{\bar{\gamma}}\, \mathbb{R}^{\alpha_1\beta_2}(s) = \mathbb{E}_{ij}^{\gamma*}, \tag{II.108}$$

and the transformation law (II.91) and (II.92). If taking the special structure of the vectors (II.89) and (II.90) into account, the matrix elements (II.107) can be simplified to

$$F_{a'v';av} = (-1)^{\Delta(a')}\delta_{a',a+1}\, F_{v'v}^{\bar{\gamma}(\alpha_1\beta_2)}, \tag{II.109}$$

$$F_{v'v}^{\bar{\gamma}(\alpha_1\beta_2)} = \langle \mathbf{M}_k^{\bar{\gamma}v'}, U^{\alpha_1}\otimes U^{\beta_2}\mathbf{M}_k^{\gamma v*}\rangle, \tag{II.110}$$

where the scalar product on the right-hand side of Eq. (II.110) refers to the corresponding subspace. The matrix elements (II.110) must also be independent of the index k. In order to verify this proposition, we have to use the following relations:

$$\{U^{\alpha_1}\otimes U^{\beta_2}\}^\dagger E_{ij}^{\bar{\gamma}}\{U^{\alpha_1}\otimes U^{\beta_2}\} = E_{ij}^{\gamma*}, \tag{II.111}$$

and the fact that the vectors $\mathbf{M}_k^{\gamma v}$, and $\mathbf{M}_k^{\bar{\gamma}v}$ are H-adapted vectors, whose transformation properties are fixed through Eq. (II.1). In this connection we have to note that the $m_{\alpha_1\beta_2;\gamma}$-dimensional matrix $F^{\bar{\gamma}(\alpha_1\beta_2)}$ is also unitary.

Provided that the corresponding columns of the CG matrix M can be computed by means of the method described in Ref. 2, their components take the form

$$\{\mathbf{M}_k^{\gamma v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\beta_2;\gamma(i_v j_v)}\}_{ij} = \|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\gamma(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|} \times \sum_h R^{\alpha_1}_{ii_v}(h) R^{\beta_2}_{jj_v}(h) R^{\gamma*}_{ka_0}(h),$$

$$v = 1,2,\ldots,m_{\alpha_1\beta_2;\gamma}, \quad k = 1,2,\ldots,n_\gamma, \tag{II.112}$$

$$\{\mathbf{M}_k^{\bar{\gamma}v}\}_{ij} = \{\mathbf{M}_k^{\alpha_1\beta_2;\bar{\gamma}(i_v j_v)}\}_{ij} = \|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\bar{\gamma}(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R^{\alpha_1}(h) \times R^{\beta_2}_{jj_v}(h)\{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{ka_0},$$

$$v = 1,2,\ldots,m_{\alpha_1\beta_2;\gamma}, \quad k = 1,2,\ldots,n_\gamma, \tag{II.113}$$

and give rise to the following values for Eq. (II.110):

$$F_{v'v}^{\bar{\gamma}(\alpha_1\beta_2)} = \|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\bar{\gamma}(i_{v'}j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\alpha_1\beta_2;\gamma(i_v j_v)}\|^{-1} \times \frac{n_\gamma}{|H|}\sum_h \mathbb{R}^{\alpha_1}_{i_{v'}i_v}(hs) \times \{R^{\beta_2}(h)U^{\beta_2}\}_{j_{v'}j_v}\{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{a_0a_0}, \tag{II.114}$$

where the sets $\{(i_v, j_v)\}$ in Eqs. (II.112) and (II.113) in general are not equal.

To summarize the results of this part, we have shown that CG coefficients of type III for corepresentations are obtainable by simple formulas in terms of CG coefficients for H, namely,

$$\mathbf{W}_{1k}^{\gamma(av)} = \mathbf{Q}_k^{\gamma va}, \tag{II.115}$$

$$\mathbf{W}_{2k}^{\gamma(av)} = (-1)^{\Delta(a+1)}\sum_{v'} F_{v'v}^{\bar{\gamma}(\alpha_1\beta_2)}\mathbf{Q}_k^{\bar{\gamma}v'a+1},$$

$$a = 1,2 \quad \text{and} \quad v = 1,2,\ldots,m_{\alpha_1\beta_2;\gamma}. \tag{II.116}$$

Hence, the problem reduces to the task of computing the unitary $m_{\alpha_1\beta_2;\gamma}$-dimensional matrix $F^{\bar{\gamma}(\alpha_1\beta_2)}$ in order to obtain the desired CG coefficients of type III, provided corresponding CG coefficients for H are known.

SUMMARY

Within this paper we considered Kronecker products which are composed of co-unirreps of type I and II. Due to our general procedure the first step must be the calculation of a suitable CG matrix M decomposing the Kronecker product $R^{\alpha_1\beta_2}$, which is contained twice in $\mathbb{R}^{\alpha_1\beta_2}\downarrow H = (\oplus 2)R^{\alpha_1\beta_2}$, into a direct sum of its irreducible constituents. Provided this task has been solved, CG coefficients for G have to be calculated as follows.

CG coefficients of type I are immediately obtained, if the $m_{\alpha_1\beta_2;\alpha}$-dimensional unitary submatrix $F^{\alpha(\alpha_1\beta_2)}$ of F is computed, since the corresponding CG coefficients are because of

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbf{1}_m & \mathbf{1}_m \\ iF^{\alpha(\alpha_1\beta_2)} & F^{\alpha(\alpha_1\beta_2)}\end{bmatrix},$$

given by Eqs. (II.38) and (II.39) in which the definitions (II.11) have to be taken into account. When calculating the matrix elements of $F^{\alpha(\alpha_1\beta_2)}$ by means of Eq. (II.24), one should utilize its property to be antisymmetric.

In the case of the calculation of CG coefficients of type II, it is only necessary to compute the $m_{\alpha_1\beta_2;\beta}$-dimensional symmetric unitary submatrix $F^{\beta(\alpha_1\beta_2)}$ of F, since the special solution

$$B=\begin{bmatrix} \mathbb{1}_m & 0 \\ 0 & F^{\beta(\alpha_1\beta_2)} \end{bmatrix}$$

of Eq. (II.56) leads immediately to the corresponding CG coefficients (II.73) and (II.74) in which the definitions (II.47) have to be used. The matrix elements of $F^{\beta(\alpha_1\beta_2)}$ are obtainable from Eq. (II.61).

Because of the special solution (II.103) of Eq. (II.102), i.e.,

$$B=\mathbb{1}_M \quad \text{and} \quad C=\begin{bmatrix} 0 & F^{\tilde{\gamma}(\alpha_1\beta_2)} \\ F^{\tilde{\gamma}(\alpha_1\beta_2)} & 0 \end{bmatrix},$$

the corresponding CG coefficients of type III are given by Eqs. (II.115) and (II.116), in which the definitions (II.87) and (II.88) have to be used. Thus, it suffices to compute the $m_{\alpha_1\beta_2;\gamma}$-dimensional unitary submatrix $F^{\tilde{\gamma}(\alpha_1\beta_2)}$ of F by means of Eq. (II.110).

Summarizing our results, we succeeded not only in deriving simple equations for those unitary transformations, which link CG coefficients for G with convenient ones for H, but also in solving these equations. This lead us to simple solutions for the multiplicity problem, at which the multiplicity index referring to subductions with respect to H plays an essential role.

[1]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).

[2]R. Dirl, J. Math. Phys. **20**, 659 (1979).

[3]R. Dirl, J. Math. Phys. **21**, ××× (1980).

Clebsch–Gordan coefficients for corepresentations. I ⊗ III

R. Dirl
Institut für theoretische Physik, TU Wien A-1040 Wien, Karlsplatz 13, Austria

(Received 18 December 1978; accepted for publication 25 May 1979)

A general method is applied to compute Clebsch–Gordan coefficients for corepresentations in terms of such coefficients for the normal subgroup. The considered Kronecker products are composed of corepresentations of type I and III.

INTRODUCTION

The present paper deals with the task of computing CG matrices for corepresentations, at which the considered Kronecker products are composed of co-unirreps of type I and III. An essential simplification of this problem is achieved by utilizing the representation theory of the normal subgroup insofar as one presupposes the knowledge of convenient CG matrices for this subgroup. This leads to a much easier task of determining unitary transformations which link CG coefficients for the supergroup G with those of the subgroup H.

We organize the material as follows: In Sec. I we formulate our problem and derive useful symmetry relations for the multiplicities. Section II is divided into three parts due to the different types of co-unirreps, which have to be distinguished. For each case we are able to compute quite generally those unitary transformations which link CG coefficients for corepresentations with appropriate CG coefficients for the normal subgroup H.

I. MULTIPLICITIES FOR COREPRESENTATIONS

Throughout this paper we dicuss the problem of decomposing the following Kronecker product:

$$\mathbb{R}^{\alpha_1\gamma_2} = \{\mathbb{R}^{\alpha_1\gamma_2}(g) = \mathbb{R}^{\alpha_1}(g) \otimes \mathbb{R}^{\gamma_2}(g) : g \in G\}, \tag{I.1}$$

into a direct sum of its irreducible constituents. Since $\mathbb{R}^{\alpha_1\gamma_2}$ forms a $2n_{\alpha_1}n_{\gamma_2}$-dimensional corepresentation of G, which is in general reducible, there must exist a unitary matrix $W^{\alpha_1\gamma_2} = W$ which provides such a decomposition:

$$\begin{aligned} W^\dagger \mathbb{R}^{\alpha_1\gamma_2}(g) W^{g} = \sum_{\alpha\in A_{\rm I}} &\oplus M_{\alpha_1\gamma_2;\alpha}\mathbb{R}^{\alpha}(g) \\ &\oplus \sum_{\beta\in A_{\rm II}} \oplus M_{\alpha_1\gamma_2;\beta}\mathbb{R}^{\beta}(g) \\ &\oplus \sum_{\gamma\in A_{\rm III}} \oplus M_{\alpha_1\gamma_2;\gamma}\mathbb{R}^{\gamma}(g), \\ &\qquad \text{for all } g\in G. \end{aligned} \tag{I.2}$$

By utilizing the orthogonality relations for the characters of the unirreps of H, we obtain the following well-known results[1]:

$$M_{\alpha_1\gamma_2;\alpha} = m_{\alpha_1\gamma_2;\alpha} + m_{\alpha_1\bar\gamma_2;\alpha}, \tag{I.3}$$

$$M_{\alpha_1\gamma_2;\beta} = \tfrac{1}{2}\{m_{\alpha_1\gamma_2;\beta} + m_{\alpha_1\bar\gamma_2;\beta}\}, \tag{I.4}$$

$$M_{\alpha_1\gamma_2;\gamma} = m_{\alpha_1\gamma_2;\gamma} + m_{\alpha_1\bar\gamma_2;\gamma}, \tag{I.5}$$

where the multiplicities $m_{\dots}$ refer to the corresponding subductions with respect to the subgroup H. Considering in more detail the multiplicity formulas

$$m_{\alpha_1\gamma_2;\mu} = \frac{1}{|H|}\sum_{h\in H} X^{\alpha_1}(h)X^{\gamma_2}(h)X^{\mu}(h)^*, \quad \mu\in A_H, \tag{I.6}$$

we attain the following symmetry relations:

$$m_{\alpha_1\gamma_2;\mu} = m_{\alpha_1\bar\gamma_2;\bar\mu}, \quad \mu\in A_H, \tag{I.7}$$

$$m_{\alpha_1\gamma_2;\alpha} = m_{\alpha_1\bar\gamma_2;\alpha}, \tag{I.8}$$

$$m_{\alpha_1\gamma_2;\beta} = m_{\alpha_1\bar\gamma_2;\beta}, \tag{I.9}$$

$$m_{\alpha_1\gamma_2;\gamma} = m_{\alpha_1\bar\gamma_2;\bar\gamma}, \tag{I.10}$$

if using Eqs. (I.11), (I.15), and (I.18) of Ref. 2. Hence, we arrive at the final formulas

$$M_{\alpha_1\gamma_2;\alpha} = 2m_{\alpha_1\gamma_2;\alpha}, \tag{I.11}$$

$$M_{\alpha_1\gamma_2;\beta} = m_{\alpha_1\gamma_2;\beta}, \tag{I.12}$$

$$M_{\alpha_1\gamma_2;\gamma} = M_{\alpha_1\bar\gamma_2;\bar\gamma} = M_{\alpha_1\gamma_2;\bar\gamma}. \tag{I.13}$$

Thereby, we have to note that we cannot expect, despite the symmetry relations (I.13), that the multiplicities $m_{\alpha_1\gamma_2;\gamma}$ and $m_{\alpha_1\bar\gamma_2;\gamma}$ are equal, although it seems to be obvious.

II. CG COEFFICIENTS FOR COREPRESENTATIONS

From the outset it is assumed that appropriated determined CG matrices for H are known. For this reason we consider

$$\mathbb{R}^{\alpha_1\gamma_2}\downarrow H = R^{\alpha_1\gamma_2} \oplus R^{\alpha_1\bar\gamma_2}, \tag{II.1}$$

where we have introduced the notation

$$R^{\alpha_1\gamma_2} = \{R^{\alpha_1\gamma_2}(h) = R^{\alpha_1}(h)\otimes R^{\gamma_2}(h) : h\in H\}, \tag{II.2}$$

$$R^{\alpha_1\bar\gamma_2} = \{R^{\alpha_1\bar\gamma_2}(h) = R^{\alpha_1}(h)\otimes Z^{\gamma_2\dagger}R^{\bar\gamma_2}(h)Z^{\gamma_2} : h\in H\}, \tag{II.3}$$

which refer to Kronecker products of unirreps of H and whose dimensions are just $n_{\alpha_1}n_{\gamma_2} = n_{\alpha_1}n_{\bar\gamma_2}$. Hence, for our procedure it is necessary to know the CG matrices $M^{\alpha_1\gamma_2} = M$ and $N^{\alpha_1\gamma_2} = N$ which satisfy

$$\begin{aligned} M^\dagger R^{\alpha_1\gamma_2}(h)M &= \sum_{\alpha\in A_{\rm I}} \oplus m_{\alpha_1\gamma_2;\alpha}R^{\alpha}(h) \oplus \sum_{\beta\in A_{\rm II}} \oplus m_{\alpha_1\gamma_2;\beta}R^{\beta}(h) \\ &\quad \oplus \sum_{\gamma\in A_{\rm III}} \oplus \{m_{\alpha_1\gamma_2;\gamma}R^{\gamma}(h) \\ &\quad \oplus m_{\alpha_1\bar\gamma_2;\gamma}Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}, \\ &\qquad \text{for all } h\in H, \end{aligned} \tag{II.4}$$

$$N^\dagger R^{\alpha_1\bar\gamma_2}(h)N = \sum_{\alpha\in A_{\rm I}} \oplus m_{\alpha_1\gamma_2;\alpha}R^{\alpha}(h)$$

$$\oplus \sum_{\beta\in A_{II}} \oplus m_{\alpha_1\gamma_2;\beta} R^{\beta}(h)$$
$$\oplus \sum_{\gamma\in A_{III}} \oplus \{ m_{\alpha_1\bar{\gamma}_2;\gamma} R^{\gamma}(h)$$
$$\oplus m_{\alpha_1\gamma_2;\gamma} Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma} \},$$
$$\text{for all } h\in H, \qquad (\mathrm{II}.5)$$

and where we have already used the symmetry relations (I.8)–(I.10).

A. CG coefficients of type I

Due to our general procedure, we rewrite the defining equations for CG coefficients of type I in the same way as in the previous papers:

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{W}_k^{\alpha w} = \sum_l R^{\alpha}_{lk}(h)\mathbf{W}_l^{\alpha w}, \quad \text{for all } h\in H, \qquad (\mathrm{II}.6)$$

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U^{\alpha}_{lk} \mathbf{W}_l^{\alpha w},$$
$$w = 1,2,\dots,M_{\alpha_1\gamma_2;\alpha}, \quad k = 1,2,\dots,n_{\alpha}. \qquad (\mathrm{II}.7)$$

The notation reads in more detail

$$\{\mathbf{W}_k^{\alpha w}\}_{i,bj} = \{\mathbf{W}_k^{\alpha_1\gamma_2;\alpha w}\}_{i,bj} = W^{\alpha_1\gamma_2}_{i,bj;\alpha w k},$$
$$\alpha\in A_{\mathrm{I}}, \quad w = 1,2,\dots,M_{\alpha_1\gamma_2;\alpha}, \quad k = 1,2,\dots,n_{\alpha},$$
$$i = 1,2,\dots,n_{\alpha_1}, \quad b = 1,2, \text{ and } j = 1,2,\dots,n_{\gamma_2}, \qquad (\mathrm{II}.8)$$

where the double index b,j originates from Eqs. (I.21) and (I.22) of Ref. 2. Because of Eqs. (II.6) and (II.7), it is possible to interpret the columns of the CG matrix W as H-adapted vectors of an $2n_{\alpha_1}n_{\gamma_2}$-dimensional Euclidean space $\mathscr{W}^{\alpha_1\gamma_2}$, which have to satisfy additionally Eq. (II.7).

Since W is assumed to be unitary, the vectors

$$\mathbf{W}_k^{\alpha w}, \quad w = 1,2,\dots,M_{\alpha_1\gamma_2;\alpha}, \quad k = 1,2,\dots,n_{\alpha}, \qquad (\mathrm{II}.9)$$

define an orthonormal basis of

$$\mathscr{W}^{\alpha_1\gamma_2;\alpha} = \sum_i \mathbb{E}^{\alpha}_{ii} \mathscr{W}^{\alpha_1\gamma_2},$$
$$\dim \mathscr{W}^{\alpha_1\gamma_2} = n_{\alpha} M_{\alpha_1\gamma_2;\alpha}, \qquad (\mathrm{II}.10)$$

where the units $\mathbb{E}^{\alpha}_{ij}$ decompose into a direct sum of two submatrices $E^{\alpha(\alpha_1\gamma_2)}_{ij}$ and $E^{\alpha(\alpha_1\bar{\gamma}_2)}_{ij}$ in accord with Eq. (II.1).

$$\mathbb{E}^{\alpha}_{ij} = \mathbb{E}^{\alpha_1\gamma_2;\alpha}_{ij} = E^{\alpha_1\gamma_2;\alpha}_{ij} \oplus E^{\alpha_1\bar{\gamma}_2;\alpha}_{ij}, \qquad (\mathrm{II}.11)$$

$$E^{\alpha_1\gamma_2;\alpha}_{ij} = \frac{n_{\alpha}}{|H|} \sum_h R^{\alpha *}_{ij}(h) R^{\alpha_1\gamma_2}(h), \qquad (\mathrm{II}.12)$$

$$E^{\alpha_1\bar{\gamma}_2;\alpha}_{ij} = \frac{n_{\alpha}}{|H|} \sum_h R^{\alpha *}_{ij}(h) R^{\alpha_1\bar{\gamma}_2}(h). \qquad (\mathrm{II}.13)$$

By means of the following definitions:

$$\{\mathbf{Q}_k^{\alpha v 1}\}_{i,bj} = \delta_{b1} \{\mathbf{M}_k^{\alpha v}\}_{ij}, \qquad (\mathrm{II}.14)$$
$$\{\mathbf{Q}_k^{\alpha v 2}\}_{i,bj} = \delta_{b2} \{\mathbf{N}_k^{\alpha v}\}_{ij}, \qquad (\mathrm{II}.15)$$

it is obvious that the vectors

$$\mathbf{Q}_k^{\alpha v a}, \quad a = 1,2, \quad v = 1,2,\dots,m_{\alpha_1\gamma_2;\alpha}, \quad k = 1,2,\dots,n_{\alpha}, \qquad (\mathrm{II}.16)$$

form an other orthonormal basis of $\mathscr{W}^{\alpha_1\gamma_2;\alpha}$, where the vectors $\mathbf{M}_k^{\alpha v}$ and $\mathbf{N}_k^{\alpha v}$ have to be identified with the corresponding columns of the CG matrices M and N. Although the vectors (II.16) transform according to

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{Q}_k^{\alpha v a} = \sum_l R^{\alpha}_{lk}(h)\mathbf{Q}_l^{\alpha v a}, \quad \text{for all } h\in H, \qquad (\mathrm{II}.17)$$

we cannot expect that they are already solutions of Eq. (II.7).

However, the elements of the bases (II.9) and (II.16) must be linked by unitary transformations which may not depend on the index k:

$$\mathbf{W}_k^{\alpha w} = \sum_{av} B_{av;w} \mathbf{Q}_k^{\alpha v a}, \qquad (\mathrm{II}.18)$$

$$\mathbf{Q}_k^{\alpha v a} = \sum_w B^*_{av;w} \mathbf{W}_k^{\alpha w}, \quad k = 1,2,\dots,n_{\alpha}. \qquad (\mathrm{II}.19)$$

In this connection we realize that the symmetry relation (I.8) for the multiplicities plays an essential role.

In order to be able to determine unitary matrices B so that the corresponding vectors (II.18) satisfy Eq. (II.17), we derive

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{Q}_k^{\alpha v a *} = \sum_l U^{\alpha}_{lk} \sum_{a'v'} F_{a'v';av} \mathbf{Q}_l^{\alpha v' a'}, \qquad (\mathrm{II}.20)$$

where

$$F_{a'v';av} = \{BB^T\}_{av;a'v'} = \sum_w B_{a'v';w} B_{av;w},$$
$$a,a' = 1,2 \quad \text{and} \quad v,v' = 1,2,\dots,m_{\alpha_1\gamma_2;\alpha}. \qquad (\mathrm{II}.21)$$

These equations transform Eq. (II.7) as follows:

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U^{\alpha}_{lk} \sum_{w'} \{B^{\dagger} F B^*\}_{w'w} \mathbf{W}_l^{\alpha w'}, \qquad (\mathrm{II}.22)$$

which leads immediately to the defining equation for B, namely

$$FB^* = B, \quad \text{with} \quad BB^{\dagger} = B^{\dagger}B = \mathbb{1}_M. \qquad (\mathrm{II}.23)$$

Hence, if we can find a unitary $M_{\alpha_1\gamma_2;\alpha}$-dimensional matrix B satisfying Eq. (II.23), the corresponding CG coefficients follow immediately from Eq. (II.18). Before attacking this problem let us remark that F is a symmetric unitary matrix which can be readily verified with the aid of Eq. (II.17) by taking the special group element $h = s^2$ and utilizing Eq. (II.10) of Ref. 2. This leads to

$$FF^* = \mathbb{1}_M = \mathbb{1}_{2m}, \qquad (\mathrm{II}.24)$$

which proves our assertion.

The next problem which has to be solved is to compute the matrix elements of F. This can be done by carrying out the scalar products

$$\left\langle \mathbf{Q}_k^{\alpha v' a'}, \mathbb{R}^{\alpha_1\gamma_2}(s) \left\{ \sum_l U^{\alpha}_{kl} \mathbf{Q}_l^{\alpha v a} \right\}^* \right\rangle = F_{a'v';av}, \qquad (\mathrm{II}.25)$$

which are independent of the free index k. This can be shown by means of

$$\mathbb{R}^{\alpha_1\gamma_2}(s)^{\dagger} \mathbb{E}^{\alpha}_{ij} \mathbb{R}^{\alpha_1\gamma_2}(s) = \sum_{kl} U^{\alpha *}_{ik} U^{\alpha}_{jl} \mathbb{E}^{\alpha *}_{kl}, \qquad (\mathrm{II}.26)$$

together with the well-known properties of units and the transformation law (II.17). The matrix elements (II.25) can be simplified if taking the special structure (II.14) and (II.15) of the vectors $\mathbf{Q}_k^{\alpha v a}$ into account:

$$F_{av';av} = 0, \quad a = 1,2, \tag{II.27}$$

$$F_{1v',2v} = F^{\alpha(\alpha_1\gamma_2)}_{v'v} = \left\langle \mathbf{M}^{\alpha v'}_k, U^{\alpha_1} \otimes R^{\gamma_2}(s^2) \left\{ \sum_l U^{\alpha}_{kl} \mathbf{N}^{\alpha v}_l \right\}^* \right\rangle, \tag{II.28}$$

$$F_{2v',1v} = F^{\alpha(\alpha_1\bar{\gamma}_2)}_{v'v} = \left\langle \mathbf{N}^{\alpha v'}_k, U^{\alpha_1} \otimes \mathbb{1}_{\gamma_2} \left\{ \sum_l U^{\alpha}_{kl} \mathbf{M}^{\alpha v}_l \right\}^* \right\rangle. \tag{II.29}$$

The scalar product occurring in Eqs. (II.28) and (II.29) is analogously defined. Furthermore, we note that the matrix elements (II.28) and (II.29) must also be independent of the index k. This property has to be shown with the aid of the following relations:

$$\{U^{\alpha_1} \otimes R^{\gamma_2}(s^2)\}^\dagger E^{\alpha_1\gamma_2;\alpha}_{ij} \{U^{\alpha_1} \otimes R^{\gamma_2}(s^2)\} = \sum_{kl} U^{\alpha*}_{ik} U^{\alpha}_{jl} \{E^{\alpha_1\bar{\gamma}_2}_{kl}\}^*, \tag{II.30}$$

$$\{U^{\alpha_1} \otimes \mathbb{1}_{\gamma_2}\}^\dagger E^{\alpha_1\bar{\gamma}_2;\alpha}_{ij} \{U^{\alpha_1} \otimes \mathbb{1}_{\gamma_2}\} = \sum_{kl} U^{\alpha*}_{ik} U^{\alpha}_{jl} \{E^{\alpha_1\gamma_2;\alpha}_{kl}\}^*, \tag{II.31}$$

and the corresponding transformation law for the vectors $\mathbf{M}^{\alpha v}_k$ and $\mathbf{N}^{\alpha v}_k$ with respect to H. Now let us consider in more detail the matrix F, which reads in matrix notation

$$F = \begin{bmatrix} 0 & F^{\alpha(\alpha_1\gamma_2)} \\ F^{\alpha(\alpha_1\bar{\gamma}_2)} & 0 \end{bmatrix}, \tag{II.32}$$

where the submatrices $F^{\alpha(\alpha_1\gamma_2)}$ and $F^{\alpha(\alpha_1\bar{\gamma}_2)}$ have the same dimension $m_{\alpha_1\gamma_2;\alpha}$. Furthermore, since F is a symmetric, unitary matrix, these submatrices must satisfy

$$F^{\alpha(\alpha_1\gamma_2)^T} = F^{\alpha(\alpha_1\bar{\gamma}_2)}, \tag{II.33}$$

$$F^{\alpha(\alpha_1\gamma_2)} F^{\alpha(\alpha_1\bar{\gamma}_2)*} = \mathbb{1}_m, \tag{II.34}$$

which implies that both submatrices must be unitary and that F calculates, for example, $F^{\alpha(\alpha_1\bar{\gamma}_2)}$, since the other matrix follows directly from Eq. (II.33).

Provided the corresponding columns of the CG matrix M and N can be computed with the aid of the method given in Ref. 3, their components can be written as

$$\{\mathbf{M}^{\alpha v}_k\}_{ij} = \{\mathbf{M}^{\alpha_1\gamma_2;\alpha(i_v j_v)}_k\}_{ij} = \|\mathbf{B}^{\alpha_1\gamma_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \frac{n_\alpha}{|H|} \sum_h R^{\alpha_1}_{ii_v}(h) \times R^{\gamma_2}_{jj_v}(h) R^{\alpha*}_{ka_0}(h),$$
$$v = 1,2,\ldots,m_{\alpha_1\gamma_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \tag{II.35}$$

$$\{\mathbf{N}^{\alpha v}_k\}_{ij} = \{\mathbf{N}^{\alpha_1\bar{\gamma}_2;\alpha(i_v j_v)}_k\}_{ij} = \|\mathbf{B}^{\alpha_1\bar{\gamma}_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \frac{n_\alpha}{|H|} \sum_h R^{\alpha_1}_{ii_v}(h) \times \{Z^{\gamma_2\dagger} R^{\bar{\gamma}_2}(h) Z^{\gamma_2}\}_{jj_v} R^{\alpha*}_{ka_0}(h),$$
$$v = 1,2,\ldots,m_{\alpha_1\gamma_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha. \tag{II.36}$$

If carrying out the scalar products (II.28) and (II.29) with these special values, we obtain the following formulas:

$$F^{\alpha(\alpha_1\gamma_2)}_{v'v} = \|\mathbf{B}^{\alpha_1\gamma_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \|\mathbf{B}^{\alpha_1\bar{\gamma}_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \times \frac{n_\alpha}{|H|} \sum_l \mathbb{R}^{\alpha_1}_{i_v i_v}(hs) R^{\gamma_2}_{j_v j_v}(hs^2) \mathbb{R}^{\alpha*}_{a_0a_0}(hs), \tag{II.37}$$

$$F^{\alpha(\alpha_1\bar{\gamma}_2)}_{v'v} = \|\mathbf{B}^{\alpha_1\bar{\gamma}_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \|\mathbf{B}^{\alpha_1\gamma_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \times \frac{n_\alpha}{|H|} \sum_h \mathbb{R}^{\alpha_1}_{i_v i_v}(hs) \times \{Z^{\gamma_2\dagger} R^{\bar{\gamma}_2}(h) Z^{\gamma_2}\}_{j_v j_v} \mathbb{R}^{\alpha*}_{a_0a_0}(hs). \tag{II.38}$$

Thereby, we have to note that the index sets $\{(i_v, j_v)\}$ in general are quite different and that Eqs. (II.37) and (II.38) have to satisfy Eq. (II.33), even though they are represented by different expressions.

Now let us return to the problem of determining unitary matrices B which are solutions of Eq. (II.23). For this purpose we make the following ansatz:

$$B = \begin{bmatrix} \mathbf{A} & F^{\alpha(\alpha_1\gamma_2)}\mathbf{B}^* \\ F^{\alpha(\alpha_1\bar{\gamma}_2)}\mathbf{A}^* & \mathbf{B} \end{bmatrix}, \tag{II.39}$$

where $\mathbf{A}$ and $\mathbf{B}$ shall be $m_{\alpha_1\gamma_2;\alpha}$-dimensional matrices being proportional by numerical factors to unitary ones, but otherwise arbitrary. Each matrix of the type (II.39) is a solution of Eq. (II.23) which can be readily shown by means of Eqs. (II.33) and (II.34). Taking into account that B must be unitary, we obtain

$$\mathbf{A}\mathbf{A}^+ + \mathbf{B}\mathbf{B}^+ = \mathbb{1}_m, \tag{II.40}$$

$$-(\mathbf{A}\mathbf{A}^T) F^{\alpha(\alpha_1\bar{\gamma}_2)\dagger} = F^{\alpha(\alpha_1\gamma_2)}(\mathbf{B}\mathbf{B}^T)^* \tag{II.41}$$

as additional conditions, which suggest that one chooses, for example,

$$\mathbf{A} = \frac{i}{\sqrt{2}} \mathbb{1}_m \quad \text{and} \quad \mathbf{B} = \frac{1}{\sqrt{2}} F^{\alpha(\alpha_1\bar{\gamma}_2)}, \tag{II.42}$$

as special solutions of Eqs. (II.40) and (II.41). Hence,

$$B = \frac{1}{\sqrt{2}} \begin{bmatrix} i\mathbb{1}_m & \mathbb{1}_m \\ -iF^{\alpha(\alpha_1\bar{\gamma}_2)} & F^{\alpha(\alpha_1\bar{\gamma}_2)} \end{bmatrix}, \tag{II.43}$$

represents a special solution of Eq. (II.23). Consequently, one can identify the multiplicity index w with the pair (a,v), i.e.,

$$w = (a,v), \quad a = 1,2, \quad \text{and} \quad v = 1,2,\ldots,m_{\alpha_1\gamma_2;\alpha}, \tag{II.44}$$

and the corresponding CG coefficients of type I take the form

$$\mathbf{W}^{\alpha(1v)}_k = \frac{i}{\sqrt{2}} \left\{ \mathbf{Q}^{\alpha v1}_k - \sum_{v'} F^{\alpha(\alpha_1\bar{\gamma}_2)}_{v'v} \mathbf{Q}^{\alpha v'2}_k \right\}, \tag{II.45}$$

$$\mathbf{W}^{\alpha(2v)}_k = \frac{1}{\sqrt{2}} \left\{ \mathbf{Q}^{\alpha v1}_k + \sum_{v'} F^{\alpha(\alpha_1\bar{\gamma}_2)}_{v'v} \mathbf{Q}^{\alpha v'2}_k \right\},$$
$$v = 1,2,\ldots,m_{\alpha_1\gamma_2;\alpha}. \tag{II.46}$$

Thus, we have shown that, also for this case, CG coefficients of type I for corepresentations can be traced back by simple formulas to convenient CG coefficients for the subgroup H. The only problem is to compute, for example, the matrix $F^{\alpha(\alpha_1\bar{\gamma}_2)}$.

B. CG coefficients of type II

As already known we start our considerations from

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{W}_{dk}^{\beta w}=\sum_l R_{lk}^{\beta}(h)\mathbf{W}_{dl}^{\beta w},\quad \text{for all } h\in H, \tag{II.47}$$

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{W}_{dk}^{\beta w *}=(-1)^{\Delta(d+1)}\sum_l U_{lk}^{\beta}\mathbf{W}_{d+1,l}^{\beta w},$$

$$w=1,2,\dots,M_{\alpha_1\gamma_2;\beta},\quad d=1,2,\quad \text{and}\quad k=1,2,\dots,n_\beta, \tag{II.48}$$

which are the defining equations for the CG coefficients of type II for corepresentations, where

$$\{\mathbf{W}_{dk}^{\beta w}\}_{i,bj}=\{\mathbf{W}_{dk}^{\alpha_1\gamma_2;\beta w}\}_{i,bj}=W_{i,bj;\beta w d k}^{\alpha_1\gamma_2}$$

$$\beta\in A_{\rm II},\quad w=1,,2,\dots,M_{\alpha_1\gamma_2;\beta},\quad d=1,2,$$

$$\text{and}\quad k=1,2,\dots,n_\beta,\quad 1,2,\dots,n_{\alpha_1},\quad b=1,2,$$

$$\text{and}\quad j=1,2,\dots,n_{\gamma_2}. \tag{II.49}$$

For fixed $\beta\in A_{\rm II}$ the vectors

$$\mathbf{W}_{dk}^{\beta w},\quad w=1,2,\dots,M_{\alpha_1\gamma_2;\beta},$$

$$d=1,2,\ \text{and}\ k=1,2,\dots,n_\beta \tag{II.50}$$

form an orthonormal basis of

$$\mathscr{W}^{\alpha_1\gamma_2;\beta}=\sum_i \mathbb{E}_{ii}^{\beta}\mathscr{W}^{\alpha_1\gamma_2},$$

$$\dim\mathscr{W}^{\alpha_1\gamma_2;\beta}=2n_\beta M_{\alpha_1\gamma_2;\beta}, \tag{II.51}$$

where the corresponding units decompose in a similar way as in the previous part of this section, namely,

$$\mathbb{E}_{ij}^{\beta}=\mathbb{E}_{ij}^{\alpha_1\gamma_2;\beta}=E_{ij}^{\alpha_1\gamma_2;\beta}\oplus E_{ij}^{\alpha_1\bar\gamma_2;\beta}, \tag{II.52}$$

$$E_{ij}^{\alpha_1\gamma_2;\beta}=\frac{n_\beta}{|H|}\sum_h R_{ij}^{\beta *}(h)R^{\alpha_1\gamma_2}(h), \tag{II.53}$$

$$E_{ij}^{\alpha_1\bar\gamma_2;\beta}=\frac{n_\beta}{|H|}\sum_h R_{ij}^{\beta *}(h)R^{\alpha_1\bar\gamma_2}(h). \tag{II.54}$$

Another H-adapted orthonormalized basis of this subspace can be introduced by means of

$$\{\mathbf{Q}_k^{\beta v 1}\}_{i,bj}=\delta_{b1}\{\mathbf{M}_k^{\beta v}\}_{ij}, \tag{II.55}$$

$$\{\mathbf{Q}_k^{\beta v 2}\}_{i,bj}=\delta_{b2}\{\mathbf{N}_k^{\beta v}\}_{ij}, \tag{II.56}$$

which imply that the vectors

$$\mathbf{Q}_k^{\beta v a},\quad a=1,2,\quad v=1,2,\dots,m_{\alpha_1\gamma_2;\beta},\quad k=1,2,\dots,n_\beta, \tag{II.57}$$

are H adapted but in general not a solution of Eq. (II.48), i.e.,

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{Q}_k^{\beta v a}=\sum_l R_{lk}^{\beta}(h)\mathbf{Q}_l^{\beta v a},\quad \text{for all } h\in H. \tag{II.58}$$

In order to be able to satisfy Eq. (II.48), we remember that the elements of the bases (II.50) and (II.57) must be linked by unitary transformations which are independent of the free index k:

$$\mathbf{W}_{dk}^{\beta w}=\sum_{av} B_{av;dw}\mathbf{Q}_k^{\beta v a}, \tag{II.59}$$

$$\mathbf{Q}_k^{\beta v a}=\sum_{dw} B_{av;dw}^{*}\mathbf{W}_{dk}^{\beta w},\quad k=1,2,\dots,n_\beta. \tag{II.60}$$

By similar arguments as in the previous papers we derive

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{Q}_k^{\beta v a *}=\sum_l U_{lk}^{\beta}\sum_{a'v'} F_{a'v';av}\mathbf{Q}_l^{\beta v' a'}, \tag{II.61}$$

where our notations represent

$$G_{dw;d'w'}=(-1)^{\Delta(d+1)}\delta_{d',d+1}\delta_{ww'},$$

$$d,d'=1,2,\quad w,w'=1,2,\dots,M_{\alpha_1\gamma_2;\beta}, \tag{II.62}$$

$$F_{a'v';av}=\{BGB^{T}\}_{av;a'v'}$$

$$=\sum_{dw} B_{av;dw}(-1)^{\Delta(d+1)}B_{a'v';d+1,w},$$

$$a,a'=1,2,\quad v,v'=1,2,\dots,m_{\alpha_1\gamma_2;\beta}. \tag{II.63}$$

Using Eq. (II.61) we obtain for Eq. (II.48)

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{W}_{dk}^{\beta w *}=\sum_l U_{lk}^{\beta}\sum_{d'w'}\{B^{\dagger}FB^{*}\}_{d'w';dw}\mathbf{W}_{d'l}^{\beta w'}, \tag{II.64}$$

which provides us immediately with the defining equation for B, namely,

$$BG^{T}=FB^{*}. \tag{II.65}$$

Hence, if we can find a $2M_{\alpha_1\gamma_2;\beta}$-dimensional unitary matrix satisfying Eq. (II.65), the corresponding CG coefficients follow from Eq. (II.59). In this connection we remark that the $2M_{\alpha_1\gamma_2;\beta}$-dimensional matrix F is uniquely fixed through Eq. (II.64) as antisymmetric and unitary, i.e.,

$$FF^{*}=-\mathbb{1}_{2M}=-\mathbb{1}_{2m}. \tag{II.66}$$

Equation (II.58) with $h=s^2$ and Eq. (I.14) of Ref. 2 are needed to prove the last assertion.

The next step is to calculate the matrix elements of F. This can be done by carrying out the scalar products

$$\left\langle \mathbf{Q}_k^{\beta v' a'},\mathbb{R}^{\alpha_1\gamma_2}(s)\left\{\sum_l U_{kl}^{\beta}\mathbf{Q}_l^{\beta v a}\right\}^{*}\right\rangle=F_{a'v';av}, \tag{II.67}$$

whose values are independent of the index k. In order to verify this proposition we have to use

$$\mathbb{R}^{\alpha_1\gamma_2}(s)^{\dagger}\mathbb{E}_{ij}^{\beta}\mathbb{R}^{\alpha_1\gamma_2}(s)=\sum U_{ik}^{\beta *}U_{jl}^{\beta}\mathbb{E}_{kl}^{\beta *} \tag{II.68}$$

together with the transformation law (II.58) for the vectors $\mathbf{Q}_k^{\beta v a}$. The structure of these vectors can simplify Eq. (II.67) to

$$F_{av';av}=0,\quad a=1,2, \tag{II.69}$$

$$F_{1v';2v}=F_{v'v}^{\beta(\alpha_1\gamma_2)}$$

$$=\left\langle \mathbf{M}_k^{\beta v'},U^{\alpha_1}\otimes R^{\gamma_2}(s^2)\left\{\sum_l U_{kl}^{\beta}\mathbf{N}_l^{\beta v}\right\}^{*}\right\rangle, \tag{II.70}$$

$$F_{2v';1v}=F_{v'v}^{\beta(\alpha_1\bar\gamma_2)}$$

$$=\left\langle \mathbf{N}_k^{\beta v'},U^{\alpha_1}\otimes \mathbb{1}_{\gamma_2}\left\{\sum_l U_{kl}^{\beta}\mathbf{M}_l^{\beta v}\right\}^{*}\right\rangle, \tag{II.71}$$

where the vectors $\mathbf{M}_k^{\beta v}$ and $\mathbf{N}_k^{\beta v}$ are the corresponding columns of the CG matrices M and N, respectively, and where the scalar products occurring in Eqs. (II.70) and (II.71) are analogously defined. The following matrix identities:

$$\{U^{\alpha_1}\otimes R^{\gamma_2}(s^2)\}^{\dagger}E_{ij}^{\alpha_1\gamma_2;\beta}\{U^{\alpha_1}\otimes R^{\gamma_2}(s^2)\}$$

$$=\sum_{kl} U_{ik}^{\beta *}U_{jl}^{\beta}\{E_{kl}^{\alpha_1\bar\gamma_2;\beta}\}^{*}, \tag{II.72}$$

$$\{U^{\alpha_1}\otimes \mathbb{1}_{\gamma_2}\}^{\dagger}E_{ij}^{\alpha_1\bar\gamma_2;\beta}\{U^{\alpha_1}\otimes \mathbb{1}_{\gamma_2}\}$$

$$=\sum_{kl} U_{ik}^{\beta *}U_{jl}^{\beta}\{E_{kl}^{\alpha_1\gamma_2;\beta}\}^{*}, \tag{II.73}$$

together with the transformation properties of the vectors $\mathbf{M}_k^{\beta v}$ and $\mathbf{N}_k^{\beta v}$ with respect to H allow to show that the matrix elements (II.70) and (II.71) are independent of k. Because of Eqs. (II.69)–(II.71), the matrix F reads as

$$F=\begin{bmatrix} 0 & F^{\beta(\alpha_1\gamma_2)} \\ F^{\beta(\alpha_1\bar{\gamma}_2)} & 0 \end{bmatrix}, \qquad \text{(II.74)}$$

and whose property to be unitary and antisymmetric gives

$$F^{\beta(\alpha_1\gamma_2)^T}=-F^{\beta(\alpha_1\bar{\gamma}_2)}, \qquad \text{(II.75)}$$

$$F^{\beta(\alpha_1\gamma_2)}F^{\beta(\alpha_1\bar{\gamma}_2)*}=-\mathbb{1}_m, \qquad \text{(II.76)}$$

which implies that both $m_{\alpha_1\gamma_2;\beta}$-dimensional submatrices are unitary and that F suffices to calculate one of them.

If the columns of the CG matrices M and N can be calculated by means of the method described in Ref. 3, their components take the special values

$$\begin{aligned}\{\mathbf{M}_k^{\beta v}\}_{ij} &= \{\mathbf{M}_k^{\alpha_1\gamma_2;\beta(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\alpha_1\gamma_2;\beta(i_v j_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h R_{ii_v}^{\alpha_1}(h) \\ &\quad\times R_{jj_v}^{\gamma_2}(h)R_{ka_0}^{\beta*}(h), \\ &\quad v=1,2,\dots,m_{\alpha_1\gamma_2;\beta},\quad k=1,2,\dots,n_\beta, \end{aligned} \qquad \text{(II.77)}$$

$$\begin{aligned}\{\mathbf{N}_k^{\beta v}\}_{ij} &= \{\mathbf{N}_k^{\alpha_1\bar{\gamma}_2;\beta(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\alpha_1\bar{\gamma}_2;\beta(i_v j_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h R_{ii_v}^{\alpha_1}(h) \\ &\quad\times\{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{jj_v}R_{ka_0}^{\beta*}(h), \\ &\quad v=1,2,\dots,m_{\alpha_1\gamma_2;\beta},\quad k=1,2,\dots,n_\beta, \end{aligned} \qquad \text{(II.78)}$$

which give rise to the following expressions:

$$\begin{aligned}F_{v'v}^{\beta(\alpha_1\gamma_2)} &= \|\mathbf{B}_{a_0}^{\alpha_1\gamma_2;\beta(i_{v'}j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\alpha_1\bar{\gamma}_2;\beta(i_v j_v)}\|^{-1} \\ &\quad\times\frac{n_\beta}{|H|}\sum_h \mathbb{R}_{i_{v'}i_v}^{\alpha_1}(hs)R_{j_{v'}j_v}^{\gamma_2}(hs^2)\{R^{\beta}(h)U^{\beta}\}_{a_0a_0}^{*}, \end{aligned} \qquad \text{(II.79)}$$

$$\begin{aligned}F_{v'v}^{\beta(\alpha_1\bar{\gamma}_2)} &= \|\mathbf{B}_{a_0}^{\alpha_1\bar{\gamma}_2;\beta(i_{v'}j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\alpha_1\gamma_2;\beta(i_v j_v)}\|^{-1} \\ &\quad\times\frac{n_\beta}{|H|}\sum_h \mathbb{R}_{i_{v'}i_v}^{\alpha_1}(hs) \\ &\quad\times\{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{j_{v'}j_v}\{R^{\beta}(h)U^{\beta}\}_{a_0a_0}^{*}, \end{aligned} \qquad \text{(II.80)}$$

and where the same arguments concerning the sets $\{(i_v,j_v)\}$ and the symmetry relation (II.75) hold as before.

Now we are in the position to determine unitary matrices B which satisfy Eq. (II.65). For this purpose we define

$$\begin{aligned}\{\mathbf{B}^{d,w}\}_{av} &= B_{av;dw}, \\ &\quad a=1,2,\quad v=1,2,\dots,m_{\alpha_1\gamma_2;\beta}, \\ &\quad d=1,2,\quad w=1,2,\dots,M_{\alpha_1\gamma_2;\beta}, \end{aligned} \qquad \text{(II.81)}$$

so that Eq. (II.65) can be written as

$$\begin{aligned}F\mathbf{B}^{d,w*} &= (-1)^{\Delta(d+1)}\mathbf{B}^{d+1,w}, \\ &\quad d=1,2,\quad w=1,2,\dots,M_{\alpha_1\gamma_2;\beta}. \end{aligned} \qquad \text{(II.82)}$$

The condition that B shall be unitary requires that the vectors $\mathbf{B}^{d,w}$ must be orthonormal with respect to both indices, which presupposes an appropriately defined scalar product for these vectors. Furthermore, fixing the vectors $\mathbf{B}^{1,w}$, $w=1,2,\dots,M_{\alpha_1\gamma_2;\beta}$, the corresponding vectors $\mathbf{B}^{2,w}$, $w=1,2,\dots,M_{\alpha_1\gamma_2;\beta}$ are uniquely determined through Eq. (II.82). In order to obtain orthonormal vectors we choose

$$\{\mathbf{B}^{1,w}\}_{av}=\delta_{a1}\delta_{vw}, \qquad \text{(II.83)}$$

which implies for the remaining vectors

$$[\mathbf{B}^{2,w}]_{av}=\delta_{a2}F_{wv}^{\beta(\alpha_1\gamma_2)}. \qquad \text{(II.84)}$$

These vectors define a unitary matrix

$$B=\begin{bmatrix}\mathbb{1}_M & 0 \\ 0 & F^{\beta(\alpha_1\gamma_2)^T}\end{bmatrix}, \qquad \text{(II.85)}$$

which gives rise to the following CG coefficients of type II:

$$\mathbf{W}_{1k}^{\beta w}=\mathbf{Q}_k^{\beta w1}, \qquad \text{(II.86)}$$

$$\mathbf{W}_{2k}^{\beta w}=\sum_v F_{wv}^{\beta(\alpha_1\gamma_2)}\mathbf{Q}_k^{\beta v2},\quad w=1,2,\dots,m_{\alpha_1\gamma_2;\beta}. \qquad \text{(II.87)}$$

Hence, it is possible to identify the multiplicity index w with the original index v, i.e.,

$$w=v,\quad v=1,2,\dots,m_{\alpha_1\gamma_2;\beta}. \qquad \text{(II.88)}$$

Consequently, we have shown that, also for this case, CG coefficients of type II can be traced back by simple formulas to convenient CG coefficients for H.

C. CG coefficients of type III

The defining equations for CG coefficients of type III can be written in the following form:

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{W}_{1k}^{\gamma w}=\sum_l R_{lk}^{\gamma}(h)\mathbf{W}_{1l}^{\gamma w}, \qquad \text{(II.89)}$$

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{W}_{2k}^{\gamma w}=\sum_l\{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{lk}\mathbf{W}_{2l}^{\gamma w},$$

$$\text{for all } h\in H, \qquad \text{(II.90)}$$

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{W}_{1k}^{\gamma w*}=\mathbf{W}_{2k}^{\gamma w}, \qquad \text{(II.91)}$$

$$\begin{aligned}\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{W}_{2k}^{\gamma w*} &= \sum_l R_{lk}^{\gamma}(s^2)\mathbf{W}_{1l}^{\gamma w}, \\ &\quad w=1,2,\dots,M_{\alpha_1\gamma_2;\gamma},\quad k=1,2,\dots,n_\gamma, \end{aligned} \qquad \text{(II.92)}$$

where we have introduced the notation

$$\begin{aligned}\{\mathbf{W}_{dk}^{\gamma w}\}_{i,bj} &= \{\mathbf{W}_{dk}^{\alpha_1\gamma_2;\gamma w}\}_{i,bj}=W_{i,bj;\gamma w dk}^{\alpha_1\gamma_2}, \\ &\gamma\in A_{\text{III}},\quad w=1,2,\dots,M_{\alpha_1\gamma_2;\gamma}, \\ &d=1,2,\quad\text{and}\quad k=1,2,\dots,n_\gamma, \\ &i=1,2,\dots,n_{\alpha_1},\quad b=1,2,\quad\text{and}\quad j=1,2,\dots,n_{\gamma_2}. \end{aligned} \qquad \text{(II.93)}$$

Since the CG matrix W is assumed to be unitary, the vectors representing columns of W, i.e.,

$$\mathbf{W}_{dk}^{\gamma w},\quad w=1,2,\dots,M_{\alpha_1\gamma_2;\gamma},\quad d=1,2,\quad k=1,2,\dots,n_\gamma, \qquad \text{(II.94)}$$

form an orthonormal basis of

$$\mathscr{W}^{\alpha_1\gamma_2;\gamma}=\sum_i\{\mathbb{E}_{ii}^{\gamma}+\mathbb{E}_{ii}^{\bar{\gamma}}\}\mathscr{W}^{\alpha_1\gamma_2},$$

$$\dim\mathscr{W}^{\alpha_1\gamma_2;\gamma}=2n_\gamma M_{\alpha_1\gamma_2;\gamma}, \qquad \text{(II.95)}$$

where the units $\mathbb{E}_{ij}^{\gamma}$ and $\mathbb{E}_{ij}^{\bar{\gamma}}$ decompose into a direct sum of submatrices, namely,

$$\mathbb{E}_{ij}^{\gamma}=\mathbb{E}_{ij}^{\alpha_1\gamma_2;\gamma}=E_{ij}^{\alpha_1\gamma_2;\gamma}\oplus E_{ij}^{\alpha_1\bar{\gamma}_2;\gamma}, \qquad \text{(II.96)}$$

$$E^{\alpha_1\gamma_2;\gamma}_{ij} = \frac{n_\gamma}{|H|}\sum_h R^{\gamma *}_{ij}(h) R^{\alpha_1\gamma_2}(h), \qquad (II.97)$$

$$E^{\alpha_1\bar{\gamma}_2;\gamma}_{ij} = \frac{n_\gamma}{|H|}\sum_h R^{\gamma *}_{ij}(h) R^{\alpha_1\bar{\gamma}_2}(h), \qquad (II.98)$$

$$\mathbb{E}^{\bar{\gamma}}_{ij} = \mathbb{E}^{\alpha_1\gamma_2;\bar{\gamma}}_{ij} = E^{\alpha_1\gamma_2;\bar{\gamma}}_{ij} \oplus E^{\alpha_1\bar{\gamma}_2;\bar{\gamma}}_{ij}, \qquad (II.99)$$

$$E^{\alpha_1\gamma_2;\bar{\gamma}}_{ij} = \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}^*_{ij} R^{\alpha_1\gamma_2}(h), \qquad (II.100)$$

$$E^{\alpha_1\bar{\gamma}_2;\bar{\gamma}}_{ij} = \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}^*_{ij} R^{\alpha_1\bar{\gamma}_2}(h). \qquad (II.101)$$

By means of the definitions

$$\{\mathbf{Q}^{\gamma v1}_k\}_{i,bj} = \delta_{b1}\{\mathbf{M}^{\gamma v}_k\}_{ij}, \qquad (II.102)$$

$$\{\mathbf{Q}^{\gamma v2}_k\}_{i,bj} = \delta_{b2}\{\mathbf{N}^{\gamma v}_k\}_{ij}, \qquad (II.103)$$

$$\{\mathbf{Q}^{\bar{\gamma} v1}_k\}_{i,bj} = \delta_{b1}\{\mathbf{M}^{\bar{\gamma} v}_k\}_{ij}, \qquad (II.104)$$

$$\{\mathbf{Q}^{\bar{\gamma} v2}_k\}_{i,bj} = \delta_{b2}\{\mathbf{N}^{\bar{\gamma} v}_k\}_{ij}, \qquad (II.105)$$

we define a further orthonormal basis of $\mathscr{W}^{\alpha_1\gamma_2;\gamma}$:

$$\mathbf{Q}^{\gamma va}_k, \quad a = 1,2, \quad v = 1,2,\ldots,m_1(a),$$
$$k = 1,2,\ldots,n_\gamma, \quad m_1(1) = m_{\alpha_1\gamma_2;\gamma}, \quad m_1(2) = m_{\alpha_1\bar{\gamma}_2;\gamma}, \qquad (II.106)$$

$$\mathbf{Q}^{\bar{\gamma} va}_k, \quad a = 1,2, \quad v = 1,2,\ldots,m_2(a),$$
$$k = 1,2,\ldots,n_\gamma, \quad m_2(1) = m_{\alpha_1\bar{\gamma}_2;\gamma}, \quad m_2(2) = m_{\alpha_1\gamma_2;\gamma}. \qquad (II.107)$$

Since the multiplicities $m_{\alpha_1\gamma_2;\gamma}$ and $m_{\alpha_1\bar{\gamma}_2;\gamma}$ are in general different, we are forced to distinguish them by introducing the notation $m_i(a)$. Nevertheless, the transformation properties of the vectors (II.106) and (II.107) with respect to H are given by

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{Q}^{\gamma va}_k = \sum_l R^{\gamma}_{lk}(h)\mathbf{Q}^{\gamma va}_l, \qquad (II.108)$$

$$\mathbb{R}^{\alpha_1\gamma_2}(h)\mathbf{Q}^{\bar{\gamma} va}_k = \sum_l \{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}_{lk}\mathbf{Q}^{\bar{\gamma} va}_l, \quad \text{for all } h\in H. \qquad (II.109)$$

Although Eqs. (II.108) and (II.109) are in accordance with Eqs. (II.89) and (II.90), we cannot expect that they are also solutions of Eqs. (II.91) and (II.92). However, due to Schur's lemma with respect to H, both bases (II.94) and (II.106) and (II.107) must be linked by unitary transformations which are independent of the index k:

$$\mathbf{W}^{\gamma w}_{1k} = \sum_{av} B_{av;w}\mathbf{Q}^{\gamma va}_k, \qquad (II.110)$$

$$\mathbf{Q}^{\gamma va}_k = \sum_{w} B^*_{av;w}\mathbf{W}^{\gamma w}_{1k}, \quad k = 1,2,\ldots,n_\gamma, \qquad (II.111)$$

$$\mathbf{W}^{\gamma w}_{2k} = \sum_{av} C_{av;w}\mathbf{Q}^{\bar{\gamma} va}_k, \qquad (II.112)$$

$$\mathbf{Q}^{\bar{\gamma} va}_k = \sum_{w} C^*_{av;w}\mathbf{W}^{\gamma w}_{2k}, \quad k = 1,2,\ldots,n_\gamma. \qquad (II.113)$$

Concerning the matrix notation for B and C, we have to realize that the index (a,v) differs for B and C, if $m_{\alpha_1\gamma_2;\gamma} \neq m_{\alpha_1\bar{\gamma}_2;\gamma}$. This fact should always be taken into account for the following considerations.

Hence, our problem is now reduced to the task of determining $M_{\alpha_1\gamma_2;\gamma}$-dimensional unitary matrices B and C, so that the corresponding vectors (II.110) and (II.112) satisfy Eqs. (II.91) and (II.92). A simple calculation yields

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{Q}^{\gamma va *}_k = \sum_{a'v'} F_{a'v';av}\mathbf{Q}^{\bar{\gamma} v'a'}_k, \qquad (II.114)$$

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{Q}^{\bar{\gamma} va *}_k = \sum_l R^{\gamma}_{lk}(s^2)\sum_{a'v'} F_{av;a'v'}\mathbf{Q}^{\gamma v'a'}_l, \qquad (II.115)$$

where we have introduced as notation

$$F_{a'v';av} = \{CB^T\}_{a'v';av} = \sum_w C_{a'v';w}B_{av;w},$$
$$a' = 1,2, \quad \text{and} \quad v' = 1,2,\ldots,m_2(a'),$$
$$a = 1,2, \quad \text{and} \quad v = 1,2,\ldots,m_1(a). \qquad (II.116)$$

These equations now allow Eqs. (II.91) and (II.92) to be written as

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{W}^{\gamma w *}_{1k} = \sum_{w'} \{C^\dagger F B^*\}_{w'w}\mathbf{W}^{\gamma w'}_{2k}, \qquad (II.117)$$

$$\mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{W}^{\gamma w *}_{2k} = \sum_l R^{\gamma}_{lk}(s^2)\sum_{w'} \{C^\dagger F B^*\}^T_{w'w}\mathbf{W}^{\gamma w'}_{1l}, \qquad (II.118)$$

respectively, from which follow the defining equation for B and C, namely,

$$C = FB^*. \qquad (II.119)$$

As in the previous papers we can choose

$$B = \mathbb{1}_M \Longleftrightarrow C = F, \qquad (II.120)$$

as a special solution of Eq. (II.119), since F is a unitary matrix. Thus, we arrive at the final formulas

$$\mathbf{W}^{\gamma(av)}_{1k} = \mathbf{Q}^{\gamma va}_k, \qquad (II.121)$$

$$\mathbf{W}^{\gamma(av)}_{2k} = \sum_{a'=1}^{2}\sum_{v'=1}^{m_2(a')} F_{a'v';av}\mathbf{Q}^{\bar{\gamma} v'a'}_k,$$
$$a = 1,2, \text{ and } v = 1,2,\ldots,m_1(a), \quad k = 1,2,\ldots,n_\gamma, \qquad (II.122)$$

which make it obvious that the multiplicity index w can be identified with the double index (a,v), i.e.,

$$w = (a,v), \quad a = 1,2, \quad \text{and} \quad v = 1,2,\ldots,m_1(a). \qquad (II.123)$$

Consequently, the remaining problem is to compute the matrix elements of F. This can be done by means of

$$\langle \mathbf{Q}^{\bar{\gamma} v'a'}_k, \mathbb{R}^{\alpha_1\gamma_2}(s)\mathbf{Q}^{\gamma va *}_k\rangle = F_{a'v';av}, \qquad (II.124)$$

whose values must be independent of the index k. In order to verify this assertion, one has to use among others

$$\mathbb{R}^{\alpha_1\gamma_2}(s)^\dagger \mathbb{E}^{\bar{\gamma}}_{ij}\, \mathbb{R}^{\alpha_1\gamma_2}(s) = \mathbb{E}^{\gamma *}_{ij}. \qquad (II.125)$$

Utilizing the special structure of the vectors (II.106) and (II.107), the matrix elements (II.124) simplify to

$$F_{av';av} = 0, \quad a = 1,2, \qquad (II.126)$$

$$F_{1v';2v} = F^{\bar{\gamma}(\alpha_1\gamma_2)}_{v'v} = \langle \mathbf{M}^{\bar{\gamma} v'}_k, U^{\alpha_1}\otimes R^{\gamma_2}(s^2)\mathbf{N}^{\gamma v *}_k\rangle, \qquad (II.127)$$

$$F_{2v';1v} = F^{\bar{\gamma}(\alpha_1\bar{\gamma}_2)}_{v'v} = \langle \mathbf{N}^{\bar{\gamma} v'}_k, U^{\alpha_1}\otimes \mathbb{1}_{\gamma_2}\mathbf{M}^{\gamma v *}_k\rangle, \qquad (II.128)$$

where the scalar products occurring in Eqs. (II.127) and (II.128) have to be defined analogously. Together with the relations

$$\{U^{\alpha_1} \otimes R^{\gamma_2}(s^2)\}^\dagger E_{ij}^{\alpha_1\gamma_2;\bar\gamma}\{U^{\alpha_1} \otimes R^{\gamma_2}(s^2)\} = \{E_{ij}^{\alpha_1\bar\gamma_2;\gamma}\}^*, \tag{II.129}$$

$$\{U^{\alpha_1} \otimes \mathbb{1}_{\gamma_2}\}^\dagger E_{ij}^{\alpha_1\bar\gamma_2;\bar\gamma}\{U^{\alpha_1} \otimes \mathbb{1}_{\gamma_2}\} = \{E_{ij}^{\alpha_1\gamma_2;\gamma}\}^*, \tag{II.130}$$

it is easy to show that the matrix elements (II.127) and (II.128) are indeed independent of the index k. In this connection we remark that both submatrices $F^{\bar\gamma(\alpha_1\gamma_2)}$ and $F^{\bar\gamma(\alpha_1\bar\gamma_2)}$ are also unitary.

Before summarizing our results let us assume that the corresponding columns of the CG matrices M and N can be calculated with the aid of the method given in Ref. 3:

$$\begin{aligned}\{\mathbf{M}_k^{\gamma v}\}_{ij} &= \{\mathbf{M}_k^{\alpha_1\gamma_2;\gamma(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\alpha_1\gamma_2;\gamma(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|} \\ &\quad\times\sum_h R_{ii_v}^{\alpha_1}(h)R_{jj_v}^{\gamma_2}(h)R_{ka_0}^{\gamma *}(h), \\ &\quad v=1,2,\ldots,m_1(1),\quad k=1,2,\ldots,n_\gamma,\end{aligned} \tag{II.131}$$

$$\begin{aligned}\{\mathbf{N}_k^{\gamma v}\}_{ij} &= \{\mathbf{N}_k^{\alpha_1\bar\gamma_2;\gamma(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a}^{\alpha_1\bar\gamma_2;\gamma(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R_{ii_v}^{\alpha_1}(h) \\ &\quad\times\{Z^{\gamma_2\dagger}R^{\bar\gamma_2}(h)Z^{\gamma_2}\}_{jj_v}R_{ka_0}^{\gamma *}(h), \\ &\quad v=1,2,\ldots,m_1(2),\quad k=1,2,\ldots,n_\gamma,\end{aligned} \tag{II.132}$$

$$\begin{aligned}\{\mathbf{M}_k^{\bar\gamma v}\}_{ij} &= \{\mathbf{M}_k^{\alpha_1\gamma_2;\bar\gamma(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\alpha_1\gamma_2;\bar\gamma(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R_{ii_v}^{\alpha_1}(h) \\ &\quad\times R_{jj_v}^{\gamma_2}(h)\{Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}_{ka_0}^{*}, \\ &\quad v=1,2,\ldots,m_2(1),\quad k=1,2,\ldots,n_\gamma,\end{aligned} \tag{II.133}$$

$$\begin{aligned}\{\mathbf{N}_k^{\bar\gamma v}\}_{ij} &= \{\mathbf{N}_k^{\alpha_1\bar\gamma_2;\bar\gamma(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\alpha_1\bar\gamma_2;\bar\gamma(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R_{ii_v}^{\alpha_1}(h) \\ &\quad\times\{Z^{\gamma_2\dagger}R^{\bar\gamma_2}(h)Z^{\gamma_2}\}_{jj_v}\{Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}_{ka}^{*}, \\ &\quad v=1,2,\ldots,m_2(2),\quad k=1,2,\ldots,n_\gamma.\end{aligned} \tag{II.134}$$

Thereby, we have to note the different sets $\{(i_v, j_v)\}$ occurring in (II.131)–(II.134). Inserting Eqs. (II.131)–(II.134) into Eqs. (II.127) and (II.128), we obtain

$$\begin{aligned}F_{v'v}^{\bar\gamma(\alpha_1\gamma_2)} &= \|\mathbf{B}_{a_0}^{\alpha_1\gamma_2;\bar\gamma(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\alpha_1\bar\gamma_2;\gamma(i_v j_v);}\|^{-1} \\ &\quad\times\frac{n_\gamma}{|H|}\sum_h \mathbb{R}_{i_{v'}i_v}^{\alpha_1}(hs)R_{j_{v'}j_v}^{\gamma_2}(hs^2) \\ &\quad\times\{Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}_{a_0a_0}^{*},\end{aligned} \tag{II.135}$$

$$\begin{aligned}F_{v'v}^{\bar\gamma(\alpha_1\bar\gamma_2)} &= \|\mathbf{B}_{a_0}^{\alpha_1\bar\gamma_2;\bar\gamma(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\alpha_1\gamma_2;\gamma(i_v j_v)}\|^{-1} \\ &\quad\times\frac{n_\gamma}{|H|}\sum_h \mathbb{R}_{i_{v'}i_v}^{\alpha_1}(hs)\{Z^{\gamma_2\dagger}R^{\bar\gamma_2}(h)Z^{\gamma_2}\}_{j_{v'}j_v} \\ &\quad\times\{Z^{\gamma\dagger}R^{\bar\gamma}(h)Z^{\gamma}\}_{a_0a_0}^{*}.\end{aligned} \tag{II.136}$$

Concluding this part we summarize our final results

$$\begin{aligned}&\mathbf{W}_{1k}^{\gamma(av)} = \mathbf{Q}_k^{\gamma va}, \\ &a=1,2 \quad\text{and}\quad v=1,2,\ldots,m_1(a), \\ &k=1,2,\ldots,n_\gamma,\end{aligned} \tag{II.137}$$

$$\begin{aligned}\mathbf{W}_{2k}^{\gamma(1v)} &= \sum_{v'=1}^{m_2(2)} F_{v'v}^{\bar\gamma(\alpha_1\bar\gamma_2)}\mathbf{Q}_k^{\bar\gamma v'2}, \\ &\quad v=1,2,\ldots,m_1(1),\quad k=1,2,\ldots,n_\gamma,\end{aligned} \tag{II.138}$$

$$\begin{aligned}\mathbf{W}_{2k}^{\gamma(2v)} &= \sum_{v'=1}^{m_2(1)} F_{v'v}^{\bar\gamma(\alpha_1\gamma_2)}\mathbf{Q}_k^{\bar\gamma v'1}, \\ &\quad v=1,2,\ldots,m_1(2),\quad k=1,2,\ldots,n_\gamma,\end{aligned} \tag{II.139}$$

which show that CG coefficients of type III are linked by simple unitary transformations with convenient CG coefficients for the subgroup H. The only problem is to compute the unitary submatrices $F^{\bar\gamma(\alpha_1\gamma_2)}$ and $F^{\bar\gamma(\alpha_1\bar\gamma_2)}$, whose dimensions however are in general not equal.

SUMMARY

This paper deals with the computation of CG coefficients for Kronecker products which are composed of counirreps of type I and III. The first step of the present method is to compute convenient CG matrices M and N which are needed to decompose the Kronecker products $R^{\alpha_1\gamma_2}$ and $R^{\alpha_1\bar\gamma_2}$ into a direct sum of their irreducible constituents, since $\mathbb{R}^{\alpha_1\gamma_2}\downarrow H = R^{\alpha_1\gamma_2} \oplus R^{\alpha_1\bar\gamma_2}$. Provided this task has been solved, one has to proceed as follows.

In order to obtain CG coefficients of type I one has to compute the matrix B, i.e.,

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbb{1}_m & \mathbb{1}_m \\ -iF^{\alpha(\alpha_1\bar\gamma_2)} & F^{\alpha(\alpha_1\bar\gamma_2)}\end{bmatrix},$$

which link CG coefficients of type I with convenient ones for H, presupposing the definitions (II.14) and (II.15) are taken into account. For this purpose it suffices to calculate the matrix elements of the $m_{\alpha_1\gamma_2;\alpha}$-dimensional unitary matrix $F^{\alpha(\alpha_1\bar\gamma_2)}$ by means of Eq. (II.29). Thereby, we have to note that the symmetry relation $m_{\alpha_1\gamma_2;\alpha} = m_{\alpha_1\bar\gamma_2;\alpha}$ gives a simple solution for the multiplicity problem.

CG coefficients of type II are given by Eqs. (II.86) and (II.87), where the definitions (II.55) and (II.56) have to be used. Therefore, the only problem is to calculate the matrix elements of the $m_{\alpha_1\gamma_2;\beta}$-dimensional unitary submatrix $F^{\beta(\alpha_1\gamma_2)}$ of B, where

$$B = \begin{bmatrix} \mathbb{1}_m & 0 \\ 0 & F^{\beta(\alpha_1\gamma_2)T}\end{bmatrix},$$

which links the corresponding CG coefficients. The matrix elements of $F^{\beta(\alpha_1\gamma_2)}$ are defined by Eq. (II.70). As in the previous case the symmetry relation $m_{\alpha_1\gamma_2;\beta} = m_{\alpha_1\bar\gamma_2;\beta}$ gives rise to a simple solution for the multiplicity problem.

Because of the special solution (II.120) of Eq. (II.119), i.e.,

$$B = \mathbb{1}_M \quad\text{and}\quad C = \begin{bmatrix} 0 & F^{\bar\gamma(\alpha_1\gamma_2)} \\ F^{\bar\gamma(\alpha_1\bar\gamma_2)} & 0\end{bmatrix},$$

the corresponding CG coefficients of type III are given by Eqs. (II.137)–(II.139), where the definitions (II.102)–(II.105) have to be taken into account. Consequently, it suffices to compute the matrix elements of the unitary submatrices $F^{\bar\gamma(\alpha_1\gamma_2)}$ and $F^{\bar\gamma(\alpha_1\bar\gamma_2)}$ by means of Eqs. (II.127) and (II.128), respectively.

Although the considered case is more complicated than the previous ones, since R^{γ_2} and $R^{\bar{\gamma}_2}$ are inequivalent unirreps of H, we succeeded in solving the multiplicity problem for each type of co-unirreps without reference to a special magnetic group.

[1]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[2]R. Dirl, J. Math. Phys. **21**, $\times\times\times$ (1980).
[3]R. Dirl, J. Math. Phys. **20**, 659 (1979).

Clebsch–Gordan coefficients for corepresentations. II ⊗ II

R. Dirl
Institut für theoretische Physik, TU Wien A-1040 Wien, Karlsplatz 13, Austria

(Received 18 December 1978; accepted for publication 25 May 1979)

A general procedure is used to determine quite general Clebsch–Gordan coefficients for corepresentations in terms of convenient Clebsch–Gordan coefficients for the normal subgroup. The considered Kronecker products are composed of corepresentations of type II only.

INTRODUCTION

In this paper we compute CG coefficients for corepresentations where the considered Kronecker products are composed of co-unirreps of type II only. Like in the previous papers we simplify this problem essential by assuming that convenient CG coefficients for the normal subgroup H of G are known. This gives rise to the much easier task of determining unitary transformations which combine these coefficients.

The material of this paper is organized as follows: In Sec. I we state our problem and summarize the required multiplicities. Due to the present approach we divide the following section into three different parts according to the possible types of co-unirreps. We derive not only the defining equations for the above mentioned unitary transformations but solve them also quite generally without reference to a special group.

I. MULTIPLICITIES FOR COREPRESENTATIONS

Throughout this paper we are confronted with the problem to decompose the following Kronecker product:

$$\mathbb{R}^{\beta_1\beta_2} = \{\mathbb{R}^{\beta_1\beta_2}(g) = \mathbb{R}^{\beta_1}(g) \otimes \mathbb{R}^{\beta_2}(g) : g \in G\}, \tag{I.1}$$

into a direct sum of its irreducible constituents. Since the corepresentation $\mathbb{R}^{\beta_1\beta_2}$ forms a $4n_{\beta_1}n_{\beta_2}$-dimensional representation which is in general reducible, there must exist a unitary matrix $W^{\beta_1\beta_2} = W$ satisfying

$$W^+\mathbb{R}^{\beta_1\beta_2}(g)W^g = \sum_{\alpha\in A_\text{I}} \oplus M_{\beta_1\beta_2;\alpha}\mathbb{R}^\alpha(g) \oplus \sum_{\beta\in A_\text{II}} \oplus M_{\beta_1\beta_2;\beta}\mathbb{R}^\beta(g) \oplus \sum_{\gamma\in A_\text{III}} \oplus M_{\beta_1\beta_2;\gamma}\mathbb{R}^\gamma(g), \quad \text{for all } g\in G, \tag{I.2}$$

where the corresponding multiplicites take the values[1]

$$M_{\beta_1\beta_2;\alpha} = 4m_{\beta_1\beta_2;\alpha}, \tag{I.3}$$

$$M_{\beta_1\beta_2;\beta} = 2m_{\beta_1\beta_2;\beta}, \tag{I.4}$$

$$M_{\beta_1\beta_2;\gamma} = 4m_{\beta_1\beta_2;\gamma} = M_{\beta_1\beta_2;\bar{\gamma}}. \tag{I.5}$$

II. CG COEFFICIENTS FOR COREPRESENTATIONS

Due to the present procedure we assume from the outset that convenient CG coefficients for H are already known. For this reason we consider at first the subduced representation

$$\mathbb{R}^{\beta_1\beta_2}\downarrow H = (\oplus 4)R^{\beta_1\beta_2}, \tag{II.1}$$

which indicates that the Kronecker product

$$R^{\beta_1\beta_2} = \{R^{\beta_1\beta_2}(h) = R^{\beta_1}(h) \otimes R^{\beta_2}(h) : h\in H\} \tag{II.2}$$

is contained four times into the reducible representation $\mathbb{R}^{\beta_1\beta_2}\downarrow H$, where $R^{\beta_1\beta_2}$ is a $n_{\beta_1}n_{\beta_2}$-dimensional representation of H. Since $R^{\beta_1\beta_2}$ is in general reducible, there must exist a unitary matrix $M^{\beta_1\beta_2} = M$ which satisfies

$$M^+R^{\beta_1\beta_2}(h)M = \sum_{\alpha\in A_\text{I}} \oplus m_{\beta_1\beta_2;\alpha}R^\alpha(h) \oplus \sum_{\beta\in A_\text{II}} \oplus m_{\beta_1\beta_2;\beta}R^\beta(h) \oplus \sum_{\gamma\in A_\text{III}} \oplus m_{\beta_1\beta_2;\gamma}\{R^\gamma(h) \oplus Z^{\gamma +}R^{\bar{\gamma}}(h)Z^\gamma\}, \quad \text{for all } h\in H. \tag{II.3}$$

A. CG coefficients of type I

Let us start from the already known defining equations for CG coefficients of type I:

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{W}_k^{\alpha w} = \sum_l R_{lk}^\alpha(h)\mathbf{W}_l^{\alpha w}, \quad \text{for all } h\in H, \tag{II.4}$$

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U_{lk}^\alpha \mathbf{W}_l^{\alpha w}, \quad w = 1,2,\ldots,M_{\beta_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \tag{II.5}$$

where the abbreviated notation has the following meaning:

$$\{\mathbf{W}_k^{\alpha w}\}_{ai,bj} = \{\mathbf{W}_k^{\beta_1\beta_2;\alpha w}\}_{ai,bj} = W_{ai,bj;\alpha wk}^{\beta_1\beta_2},$$
$$\alpha\in A_\text{I}, \quad w = 1,2,\ldots,M_{\beta_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha,$$
$$a = 1,2 \text{ and } i = 1,2,\ldots,n_{\beta_1}, \quad b = 1,2 \text{ and } j = 1,2,\ldots,n_{\beta_2}. \tag{II.6}$$

Both double indices a,i and b,j originate from Eqs. (I.21) and (I.22) of Ref. 2. Similar to the previous cases we interpret the columns of the CG matrix W as H-adapted vectors of an $4n_{\beta_1}n_{\beta_2}$-dimensional Euclidean space $\mathscr{W}^{\beta_1\beta_2}$, which have to satisfy additionally Eq. (II.5).

Since it is assumed that W is unitary, the vectors

$$\mathbf{W}_k^{\alpha w}, \quad w = 1,2,\ldots,M_{\beta_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \tag{II.7}$$

define an orthogonal basis of

$$\mathscr{W}^{\beta_1\beta_2;\alpha} = \sum_i \mathbb{E}_{ii}^\alpha\, \mathscr{W}^{\beta_1\beta_2},$$
$$\dim \mathscr{W}^{\beta_1\beta_2;\alpha} = n_\alpha M_{\beta_1\beta_2;\alpha}, \tag{II.8}$$

where the units $\mathbb{E}_{ij}^\alpha$ decompose in accordance with Eq. (II.1) into a direct sum of four identical submatrices, i.e.,

$$\mathbb{E}_{ij}^\alpha = \mathbb{E}_{ij}^{\beta_1\beta_2;\alpha} = (\oplus 4)E_{ij}^{\beta_1\beta_2;\alpha} = (\oplus 4)E_{ij}^\alpha, \tag{II.9}$$

 0022-2488/80/050983-06$01.00

$$E^{\alpha}_{ij} = \frac{n_\alpha}{|H|} E^{\alpha*}_{ij}(h) R^{\beta_1\beta_2}(h). \tag{II.10}$$

A further orthogonal basis of this subspace can be introduced by means of

$$\mathbf{Q}^{\alpha v a b}_k, \quad a,b = 1,2, \; v = 1,2,\ldots,m_{\beta_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \tag{II.11}$$

whose components are defined by

$$\{\mathbf{Q}^{\alpha v a b}_k\}_{a'i,b'j} = \delta_{aa'}\delta_{bb'}\{\mathbf{M}^{\alpha v}_k\}_{ij}, \tag{II.12}$$

where the vectors $\mathbf{M}^{\beta v}_k$ are the corresponding columns of the unitary CG matrix M. Although these vectors transform according to

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{Q}^{\alpha v a b}_k = \sum_l R^{\alpha}_{lk}(h)\mathbf{Q}^{\alpha v a b}_l, \quad \text{for all } h\in H, \tag{II.13}$$

we cannot expect that they are already a solution of Eq. (II.5).

In order to be able to accomplish this task we remember that, by virtue of Schur's lemma, the elements of the bases (II.7) and (II.11) must be linked by a special unitary transformation, namely,

$$\mathbf{W}^{\alpha w}_k = \sum_{abv} B_{abv;w}\mathbf{Q}^{\alpha v a b}_k, \tag{II.14}$$

$$\mathbf{Q}^{\alpha v a b}_k = \sum_w B^*_{abv;w}\, \mathbf{W}^{\alpha w}_k, \quad k = 1,2,\ldots,n_\alpha. \tag{II.15}$$

Hence, the problem is now to determine unitary $M_{\beta_1\beta_2;\alpha}$-dimensional matrices B, so that the corresponding vector (II.14) satisfy Eq. (II.5). Therefore, we derive

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{Q}^{\alpha v a b*}_k = \sum_l U^{\alpha}_{lk} \sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}^{\alpha v' a' b'}_l, \tag{II.16}$$

where

$$F_{a'b'v';abv} = \{BB^T\}_{abv;a'b'v'} = \sum_w B_{abv;w}B_{a'b'v',w}, \quad a,a',b,b' = 1,2, \quad v,v' = 1,2,\ldots,m_{\beta_1\beta_2;\alpha}. \tag{II.17}$$

These equations together with Eqs. (II.14) and (II.15) transform Eq. (II.5) as follows:

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}^{\alpha w*}_k = \sum_l U^{\alpha}_{lk} \sum_{w'} \{B^+FB^*\}_{w'w}\, \mathbf{W}^{\alpha w'}_l, \tag{II.18}$$

which leads us immediately to

$$FB^* = B, \quad \text{with } BB^+ = B^+B = \mathbb{1}_M. \tag{II.19}$$

By similar arguments as in the foregoing papers we can show that F is a symmetric unitary matrix, i.e.,

$$FF^* = \mathbb{1}_M = \mathbb{1}_{4m}. \tag{II.20}$$

Before attacking the problem of determining unitary matrices B which satisfy Eq. (II.19), it is necessary to compute the matrix F. The matrix elements of F are obtained by calculating the scalar products

$$\left\langle \mathbf{Q}^{v'a'b'}_k, \mathbb{R}^{\beta_1\beta_2}(s)\left\{ \sum_l U^{\alpha}_{kl}\mathbf{Q}^{\alpha v a b}_l\right\}^*\right\rangle = F_{a'b'v';abv}, \tag{II.21}$$

whose values are independent of the index k. This can be verified with the aid of

$$\mathbb{R}^{\beta_1\beta_2}(s)^+\mathbb{E}^{\alpha}_{ij}\mathbb{R}^{\beta_1\beta_2}(s) = \sum_{kl} U^{\alpha*}_{ik}U^{\alpha}_{jl}\mathbb{E}^{\alpha*}_{kl} \tag{II.22}$$

and the transformation law (II.13). A simple calculation yields for Eq. (II.21)

$$F_{a'b'v';abv} = (-1)^{\Delta(a')}\delta_{a',a+1}(-1)^{\Delta(b')}\delta_{b',b+1}F^{\alpha(\beta_1\beta_2)}_{v'v}, \tag{II.23}$$

$$F^{\alpha(\beta_1\beta_2)}_{v'v} = \left\langle \mathbf{M}^{\alpha v'}_k, U^{\beta_1}\otimes U^{\beta_2}\left\{ \sum_l U^{\alpha}_{kl}\mathbf{M}^{\alpha v}_l\right\}^*\right\rangle, \tag{II.24}$$

if taking Eq. (II.12) into account and where the scalar product in Eq. (II.24) is analogously defined. Using the relations

$$\{U^{\beta_1}\otimes U^{\beta_2}\}^+ E^{\alpha}_{ij}\{U^{\beta_1}\otimes U^{\beta_2}\} = \sum_{kl} U^{\alpha*}_{ik}U^{\alpha}_{jl}E^{\alpha*}_{kl} \tag{II.25}$$

together with the transformation properties of the vectors $\mathbf{M}^{\alpha v}_k$ with respect to H, it is easy to prove that the matrix elements (II.24) are independent of k. Now let us consider in more detail the symmetric unitary matrix F, which reads in matrix notation

$$F = \begin{bmatrix} 0 & 0 & 0 & F^{\alpha(\beta_1\beta_2)} \\ 0 & 0 & -F^{\alpha(\beta_1\beta_2)} & 0 \\ 0 & -F^{\alpha(\beta_1\beta_2)} & 0 & 0 \\ F^{\alpha(\beta_1\beta_2)} & 0 & 0 & 0 \end{bmatrix}. \tag{II.26}$$

Since F is symmetric and unitary

$$F^{\alpha(\beta_1\beta_2)'} = F^{\alpha(\beta_1\beta_2)}, \tag{II.27}$$

$$F^{\alpha(\beta_1\beta_2)}F^{\alpha(\beta_1\beta_2)*} = \mathbb{1}_m, \tag{II.28}$$

respectively, it follows that the $m_{\beta_1\beta_2;\alpha}$-dimensional submatrix $F^{\alpha(\beta_1\beta_2)}$ is also symmetric and unitary.

Provided the corresponding CG coefficients for H can be computed by means of the method given in Ref. 3, i.e.,

$$\{\mathbf{M}^{\alpha v}_k\}_{ij} = \{\mathbf{M}^{\beta_1\beta_2;\alpha(i_vj_v)}_k\}_{ij} = \|\mathbf{B}^{\beta_1\beta_2;\alpha(i_vj_v)}_{a_0}\|^{-1}\frac{n_\alpha}{|H|}\sum_h R^{\beta_1}_{ii_v}(h)R^{\beta_2}_{jj_v}(h)R^{\alpha*}_{ka_0}(h), \quad v = 1,2,\ldots,m_{\beta_1\beta_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha, \tag{II.29}$$

the matrix elements of $F^{\alpha(\beta_1\beta_2)}$ take the special values

$$F^{\alpha(\beta_1\beta_2)}_{v'v} = \|\mathbf{B}^{\beta_1\beta_2;\alpha(i_{v'}j_{v'})}_{a_0}\|^{-1}\|\mathbf{B}^{\beta_1\beta_2;\alpha(i_vj_v)}_{a_0}\|^{-1} \times\frac{n_\alpha}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}\{R^{\beta_2}(h)U^{\beta_2}\}_{j_{v'}j_v}\mathbb{R}^{\alpha*}_{a_0a_0}(hs), \tag{II.30}$$

which are of course independent of the free index k.

The next step of the present method is to determine a unitary matrix B which satisfies Eq. (II.19). If taking the special structure (II.26) of F into account, it is obvious to make the following ansatz for B:

$$B = \begin{bmatrix} \mathbf{A} & 0 & 0 & \mathbf{FB}^* \\ 0 & \mathbf{A} & -\mathbf{FB}^* & 0 \\ 0 & -\mathbf{FA}^* & \mathbf{B} & 0 \\ \mathbf{FA}^* & 0 & 0 & \mathbf{B} \end{bmatrix}, \tag{II.31}$$

where we have introduced for the sake of simplicity the abbreviated notation $\mathbf{F} = F^{\alpha(\beta_1\beta_2)}$. The $m_{\beta_1\beta_2;\alpha}$-dimensional matrices **A** and **B** shall be proportional by numerical factors to unitary ones, but otherwise arbitrary. Now it is easily verified that, for every pair **A**, **B**, Eq. (II.19) is automatically satisfied. Furthermore, unitarity of B is achieved if

$$\mathbf{AA}^\dagger + \mathbf{BB}^\dagger = \mathbb{1}_m, \qquad \text{(II.32)}$$

$$F^{\alpha(\beta_1\beta_2)}(\mathbf{AA}^T)^* = -(\mathbf{BB}^T)F^{\alpha(\beta_1\beta_2)\dagger} \qquad \text{(II.33)}$$

is satisfied. Clearly, the matrices

$$\mathbf{A} = \frac{i}{\sqrt{2}}\mathbb{1}_m \quad \text{and} \quad \mathbf{B} = \frac{1}{\sqrt{2}}F^{\alpha(\beta_1\beta_2)} \qquad \text{(II.34)}$$

represent a solution of Eq. (II.33), which lead immediately to

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbb{1}_m & 0 & 0 & \mathbb{1}_m \\ 0 & i\mathbb{1}_m & -\mathbb{1}_m & 0 \\ 0 & i\mathbf{F} & \mathbf{F} & 0 \\ -i\mathbf{F} & 0 & 0 & \mathbf{F} \end{bmatrix}. \qquad \text{(II.35)}$$

Concerning the multiplicity index w, we realize that w can be identified with the triplets (a,b,v), i.e.,

$$w = (a,b,v), \quad a,b = 1,2 \text{ and } v = 1,2,\ldots,m_{\beta_1\beta_2;\alpha}. \qquad \text{(II.36)}$$

Hence, we arrive at the final results

$$\mathbf{W}_k^{\alpha(11v)} = \frac{i}{\sqrt{2}}\left\{\mathbf{Q}_k^{\alpha v11} - \sum_{v'} F_{v'v}^{\alpha(\beta_1\beta_2)}\mathbf{Q}_k^{\alpha v'22}\right\}, \qquad \text{(II.37)}$$

$$\mathbf{W}_k^{\alpha(12v)} = \frac{i}{\sqrt{2}}\left\{\mathbf{Q}_k^{\alpha v12} + \sum_{v'} F_{v'v}^{\alpha(\beta_1\beta_2)}\mathbf{Q}_k^{\alpha v'21}\right\}, \qquad \text{(II.38)}$$

$$\mathbf{W}_k^{\alpha(21v)} = \frac{1}{\sqrt{2}}\left\{-\mathbf{Q}_k^{\alpha v12} + \sum_{v'} F_{v'v}^{\alpha(\beta_1\beta_2)}\mathbf{Q}_k^{\alpha v'21}\right\}, \qquad \text{(II.39)}$$

$$\mathbf{W}_k^{\alpha(22v)} = \frac{1}{\sqrt{2}}\left\{\mathbf{Q}_k^{\alpha v11} + \sum_{v'} F_{v'v}^{\alpha(\beta_1\beta_2)}\mathbf{Q}_k^{\alpha v'22}\right\},$$

$$v = 1,2,\ldots,m_{\beta_1\beta_2;\alpha}, \qquad \text{(II.40)}$$

which show that CG coefficients of type I for corepresentations are connected by simple unitary transformations with convenient CG coefficients for the normal subgroup H. The only problem is to compute the matrix elements of $F^{\alpha(\beta_1\beta_2)}$, where its symmetry should be utilized in any way.

B. CG coefficients of type II

The defining equations for CG coefficients of type II read as already known:

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{W}_{dk}^{\beta w} = \sum_l R_{lk}^{\beta}(h)\mathbf{W}_{dl}^{\beta w}, \quad \text{for all } h\in H, \qquad \text{(II.41)}$$

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}_{dk}^{\beta w *} = (-1)^{\Delta(d+1)}\sum_l U_{lk}^{\beta}\mathbf{W}_{d+1,l}^{\beta w},$$

$$w = 1,2,\ldots,M_{\beta_1\beta_2;\beta}, \quad d = 1,2, \text{ and } k = 1,2,\ldots,n_\beta. \qquad \text{(II.42)}$$

where our notation means

$$\{\mathbf{W}_{dk}^{\beta w}\}_{ai,bj} = \{\mathbf{W}_{dk}^{\beta_1\beta_2;\beta w}\}_{ai,bj} = W_{ai,bj;\beta w dk}^{\beta_1\beta_2}$$

$\beta\in A_{\rm II}$, $w = 1,2,\ldots,M_{\beta_1\beta_2;\beta}$, $d = 1,2$, and $k = 1,2,\ldots,n_\beta$,

$$a = 1,2 \text{ and } i = 1,2,\ldots,n_{\beta_1}, \quad b = 1,2, \text{ and } j = 1,2,\ldots,n_{\beta_2}. \qquad \text{(II.43)}$$

Due to the present procedure we interpret the vectors

$$\mathbf{W}_{dk}^{\beta w}, \quad w = 1,2,\ldots,M_{\beta_1\beta_2;\beta}, \quad d = 1,2, \; k = 1,2,\ldots,n_\beta, \qquad \text{(II.44)}$$

as H-adapted orthonormalized basis of

$$\mathscr{W}^{\beta_1\beta_2;\beta} = \sum_i \mathbb{E}_{ii}^{\beta}\,\mathscr{W}^{\beta_1\beta_2},$$

$$\dim\mathscr{W}^{\beta_1\beta_2;\beta} = 2n_\beta M_{\beta_1\beta_2;\alpha}, \qquad \text{(II.45)}$$

which have to satisfy additionally Eq. (II.42). The units $\mathbb{E}_{ij}^{\beta}$ can be written in the following form:

$$\mathbb{E}_{ij}^{\beta} = \mathbb{E}_{ij}^{\beta_1\beta_2;\beta} = (\oplus 4)E_{ij}^{\beta_1\beta_2;\beta} = (\oplus 4)E_{ij}^{\beta}, \qquad \text{(II.46)}$$

$$E_{ij}^{\beta} = \frac{n_\beta}{|H|}\sum_h R_{ij}^{\beta *}(h)R^{\beta_1\beta_2}(h), \qquad \text{(II.47)}$$

which is in accordance to Eq. (II.1). A further basis of this subspace of $\mathscr{W}^{\beta_1\beta_2;\beta}$ are the vectors

$$\mathbf{Q}_k^{\beta vab}, \quad a,b = 1,2, \quad v = 1,2,\ldots,m_{\beta_1\beta_2;\beta}, \quad k = 1,2,\ldots,n_\beta, \qquad \text{(II.48)}$$

whose components are given by

$$\{\mathbf{Q}_k^{\beta vab}\}_{a'i,b'j} = \delta_{aa'}\delta_{bb'}\{\mathbf{M}_k^{\beta v}\}_{ij} \qquad \text{(II.49)}$$

and where the vectors $\mathbf{M}_k^{\beta v}$ represent corresponding columns of the unitary CG matrix M. These vectors satisfy the first condition, namely,

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{Q}_k^{\beta vab} = \sum_l R_{lk}^{\beta}(h)\mathbf{Q}_l^{\beta vab}, \quad \text{for all } h\in H, \qquad \text{(II.50)}$$

but we cannot expect that they are already a solution of Eq. (II.42).

For this purpose we define unitary transformations which link the elements of the bases (II.44) and (II.48):

$$\mathbf{W}_{dk}^{\beta w} = \sum_{abv} B_{abv,dw}\mathbf{Q}_k^{\beta vab}, \qquad \text{(II.51)}$$

$$\mathbf{Q}_k^{\beta vab} = \sum_{dw} B_{abv,dw}^{*}\mathbf{W}_{dk}^{\beta w}, \quad k = 1,2,\ldots,n_\beta. \qquad \text{(II.52)}$$

In order to determine unitary matrices B, so that the corresponding vectors (II.51) are solutions of Eq. (II.42), we derive

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{Q}_k^{\beta vab *} = \sum_l U_{lk}^{\beta}\sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}_l^{\beta v'a'b'}, \qquad \text{(II.53)}$$

where

$$G_{dw;d'w'} = (-1)^{\Delta(d+1)}\delta_{d',d+1}\delta_{ww'},$$

$$d,d' = 1,2, \quad w,w' = 1,2,\ldots,M_{\beta_1\beta_2;\beta}, \qquad \text{(II.54)}$$

$$F_{a'b'v';abv} = \{BGB^T\}_{abv,a'b'v'}$$

$$= \sum_{dw} B_{abv;dw}(-1)^{\Delta(d+1)}B_{a'b'v',d+1,w},$$

$$a,a',b,b' = 1,2, \quad v,v' = 1,2,\ldots,m_{\beta_1\beta_2;\beta}. \qquad \text{(II.55)}$$

Equation (II.53) transforms Eq. (II.42) as follows:

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}_{dk}^{\beta w *} = \sum_l U_{lk}^{\beta}\sum_{d'w'}\{B^{+}FB^{*}\}_{d'w',dw}\mathbf{W}_{d',l}^{\beta w'}, \qquad \text{(II.56)}$$

which yields immediately the defining equation for B, namely,

$$BG^T = FB^*. \tag{II.57}$$

In this connection we remark that F is a $2M_{\beta_1\beta_2;\beta}$-dimensional antisymmetric unitary matrix which can be shown by similar arguments as in the foregoing papers:

$$FF^* = -\mathbb{1}_{2M} = -\mathbb{1}_{4m}. \tag{II.58}$$

In order to be able to solve Eq. (II.57) it is necessary to calculate the matrix elements of F. This can be done by carrying out

$$\left\langle \mathbf{Q}_k^{\beta v'a'b'}, \mathbb{R}^{\beta_1\beta_2}(s)\left\{\sum_l U^\beta_{kl}\mathbf{Q}_l^{\beta vab}\right\}^*\right\rangle = F_{a'b'v';abv}, \tag{II.59}$$

whose independence of the index k has to be proven by

$$\mathbb{R}^{\beta_1\beta_2}(s)^\dagger \mathbb{E}^\beta_{ij}\mathbb{R}^{\beta_1\beta_2}(s) = \sum_{kl} U^{\beta *}_{ik} U^\beta_{jl}\mathbb{E}^{\beta *}_{kl}. \tag{II.60}$$

If taking the special structure (II.49) of the vectors $\mathbf{Q}_k^{\beta vab}$ into account, the matrix elements (II.59) turn out to be

$$F_{a'b'v';abv} = (-1)^{\Delta(a')}\delta_{a',a+1}(-1)^{\Delta(a')} \times\delta_{b',b+1}F^{\beta(\beta_1\beta_2)}_{v'v}, \tag{II.61}$$

$$F^{\beta(\beta_1\beta_2)}_{v'v} = \left\langle \mathbf{M}_k^{\beta v'}, U^{\beta_1}\otimes U^{\beta_2}\left\{\sum_l U^\beta_{kl}\mathbf{M}_l^{\beta v}\right\}^*\right\rangle, \tag{II.62}$$

where the scalar product in Eq. (II.62) is analogously defined. The following relations:

$$\{U^{\beta_1}\otimes U^{\beta_2}\}^\dagger E^\beta_{ij}\{U^{\beta_1}\otimes U^{\beta_2}\} = \sum_{kl} U^\beta_{ik}U^\beta_{jl}E^{\beta *}_{kl} \tag{II.63}$$

are needed to prove that the matrix elements (II.62) are independent of the index k. The matrix F written down in matrix notation

$$F = \begin{bmatrix} 0 & 0 & 0 & F^{\beta(\beta_1\beta_2)} \\ 0 & 0 & -F^{\beta(\beta_1\beta_2)} & 0 \\ 0 & -F^{\beta(\beta_1\beta_2)} & 0 & 0 \\ F^{\beta(\beta_1\beta_2)} & 0 & 0 & 0 \end{bmatrix} \tag{II.64}$$

makes obvious that the following relations must hold:

$$F^{\beta(\beta_1\beta_2)T} = -F^{\beta(\beta_1\beta_2)}, \tag{II.65}$$

$$F^{\beta(\beta_1\beta_2)}F^{\beta(\beta_1\beta_2)*} = -\mathbb{1}_m, \tag{II.66}$$

since F is itself an antisymmetric unitary matrix. Thus, we obtain the nonobvious by-product

$$m_{\beta_1\beta_2;\beta} > 1, \quad \text{if } m_{\beta_1\beta_2;\beta}\neq 0, \tag{II.67}$$

since the $m_{\beta_1\beta_2;\beta}$-dimensional submatrix $F^{\beta(\beta_1\beta_2)}$ is antisymmetric and unitary.

Now let us assume that the required columns of the CG matrix M can be computed by means of the method described in Ref. 3. Their components therefore take the form

$$\begin{aligned}\{\mathbf{M}_k^{\beta v}\}_{ij} &= \{\mathbf{M}_k^{\beta_1\beta_2;\beta(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\beta_1\beta_2;\beta(i_v j_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h R^{\beta_1}_{ii_v}(h)R^{\beta_2}_{jj_v}(h)R^{\beta *}_{ka_0}(h), \\ & v = 1,2,\ldots,m_{\beta_1\beta_2;\beta}, \quad k = 1,2,\ldots,n_\beta, \end{aligned} \tag{II.68}$$

which lead us to

$$\begin{aligned}F^{\beta(\beta_1\beta_2)}_{v'v} &= \|\mathbf{B}_{a_0}^{\beta_1\beta_2;\beta(i_{v'}j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\beta_1\beta_2;\beta(i_v j_v)}\|^{-1} \\ &\times\frac{n_\beta}{|H|}\sum_h \{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}\{R^{\beta_2}(h)U^{\beta_2}\}_{j_{v'}j_v} \\ &\times\{R^\beta(h)U^\beta\}^*_{a_0a_0}, \end{aligned} \tag{II.69}$$

whose values are independent of k.

Now we are in the position to determine unitary matrices B satisfying Eq. (II.57). For this purpose we proceed as in the foregoing papers and consider the columns of B as vectors of an $2M_{\beta_1\beta_2;\beta}$-dimensional Euclidean space

$$\begin{aligned}&\{\mathbf{B}^{d,w}\}_{abv} = B_{abv;dw}, \\ &d = 1,2, \quad w = 1,2,\ldots,M_{\beta_1\beta_2;\beta}, \\ &a,b = 1,2, \quad v = 1,2,\ldots,m_{\beta_1\beta_2;\beta}. \end{aligned} \tag{II.70}$$

Therefore, Eq. (II.57) can be written as

$$F\mathbf{B}^{d,w*} = (-1)^{\Delta(d+1)}\mathbf{B}^{d+1,w}, \quad d = 1,2, \quad w = 1,2,\ldots,M_{\beta_1\beta_2;\beta}. \tag{II.71}$$

Since B must be unitary, it suffices to determine, for example, the vectors $\mathbf{B}^{1,w}, w = 1,2,\ldots,M_{\beta_1\beta_2;\beta}$ in such a way that B fulfills the required property. By taking this property into account, it is obvious to choose

$$\begin{aligned}&\{\mathbf{B}^{1,w}\}_{a'b'v'} = \delta_{a'1}\delta_{bb'}\delta_{vv'}, \\ &b = 1,2, \quad v = 1,2,\ldots,m_{\beta_1\beta_2;\beta}, \end{aligned} \tag{II.72}$$

which allows one to identify the multiplicity index w with the pair (b,v), i.e.,

$$w = (b,v), \quad b = 1,2, \quad \text{and } v = 1,2,\ldots,m_{\beta_1\beta_2;\beta}. \tag{II.73}$$

The remaining vectors $\mathbf{B}^{2,w}$ follow directly from Eqs. (II.71) and (II.64). Their components take the values

$$\{\mathbf{B}^{2,(bv)}\}_{a'b'v'} = \delta_{a'2}(-1)^{\Delta(b')}\delta_{b',b+1}F^{\beta(\beta_1\beta_2)}_{v'v} \tag{II.74}$$

and show that the corresponding matrix B is indeed unitary:

$$B = \begin{bmatrix} \mathbb{1}_m & 0 & 0 & 0 \\ 0 & \mathbb{1}_m & 0 & 0 \\ 0 & 0 & 0 & F^{\beta(\beta_1\beta_2)} \\ 0 & 0 & -F^{\beta(\beta_1\beta_2)} & 0 \end{bmatrix}. \tag{II.75}$$

Inserting the special values (II.72) and (II.74) into Eq. (II.51), we obtain immediately the corresponding CG coefficients of type II:

$$\mathbf{W}^{\beta(bv)}_{1k} = \mathbf{Q}_k^{\beta v1b}, \tag{II.76}$$

$$\mathbf{W}^{\beta(bv)}_{2k} = (-1)^{\Delta(b+1)}\sum_{v'}F^{\beta(\beta_1\beta_2)}_{v'v}\mathbf{Q}_k^{\beta v'2,b+1}, \quad b = 1,2, \quad \text{and } v = 1,2,\ldots,m_{\beta_1\beta_2;\beta}. \tag{II.77}$$

Concluding this part, we remark that CG coefficients of type II for corepresentations are connected by simple unitary transformations, whose dimensions must be larger that one, with convenient CG coefficients for the subgroup H.

C. CG coefficients of type III

Let us start once more with the defining equations for CG coefficients of type III:

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{W}^{\gamma w}_{1k} = \sum_l R^\gamma_{lk}(h)\mathbf{W}^{\gamma w}_{1l}, \tag{II.78}$$

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{W}_{2k}^{\gamma w} = \sum_l \{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{lk}\mathbf{W}_{2l}^{\gamma w},$$

$$\text{for all } h \in H, \quad \text{(II.79)}$$

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}_{1k}^{\gamma w *} = \mathbf{W}_{2k}^{\gamma w}, \quad \text{(II.80)}$$

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}_{2k}^{\gamma w *} = \sum_l R^{\gamma}_{lk}(s^2)\mathbf{W}_{1l}^{\gamma w},$$

$$w = 1,2,\dots,M_{\beta_1\beta_2;\gamma}, \quad k = 1,2,\dots,n_\gamma, \quad \text{(III.81)}$$

where our notation means

$$\{\mathbf{W}_{dk}^{\gamma w}\}_{ai,bj} = \{\mathbf{W}_{dk}^{\gamma w}\}_{ai,bj} = W^{\beta_1\beta_2}_{ai,bj;\gamma w d k},$$

$$\gamma \in A_{\text{III}}, \quad w = 1,2,\dots,M_{\beta_1\beta_2;\gamma}, \quad d = 1,2, \text{ and } k = 1,2,\dots,n_\gamma,$$

$$a = 1,2 \text{ and } i = 1,2,\dots,n_{\beta_1}, \quad b = 1,2, \text{ and } j = 1,2,\dots,n_{\beta_2}. \quad \text{(II.82)}$$

Since W is supposed to be unitary, the vectors

$$\mathbf{W}_{dk}^{\gamma w}, \quad w = 1,2,\dots,M_{\beta_1\beta_2;\gamma}, \quad d = 1,2,\ k = 1,2,\dots,n_\gamma, \quad \text{(II.83)}$$

representing columns of the CG matrix W form an orthonormal basis of

$$\mathscr{W}^{\beta_1\beta_2;\gamma} = \sum_i \{\mathbb{E}^{\gamma}_{ii} + \mathbb{E}^{\bar{\gamma}}_{ii}\}\, \mathscr{W}^{\beta_1\beta_2};$$

$$\dim \mathscr{W}^{\beta_1\beta_2;\gamma} = 2n_\gamma M_{\beta_1\beta_2;\gamma}, \quad \text{(II.84)}$$

where the units $\mathbb{E}^{\gamma}_{ij}$ and $\mathbb{E}^{\bar{\gamma}}_{ij}$ are defined analogously:

$$\mathbb{E}^{\gamma}_{ij} = \mathbb{E}^{\beta_1\beta_2;\gamma}_{ij} = (\oplus 2)\, E^{\beta_1\beta_2;\gamma}_{ij} = (\oplus 2)E^{\gamma}_{ij}, \quad \text{(II.85)}$$

$$E^{\gamma}_{ij} = \frac{n_\gamma}{|H|}\sum_h R^{\gamma *}_{ij}(h)R^{\beta_1\beta_2}(h), \quad \text{(II.86)}$$

$$\mathbb{E}^{\bar{\gamma}}_{ij} = \mathbb{E}^{\beta_1\beta_2;\bar{\gamma}}_{ij} = (\oplus 2)E^{\beta_1\beta_2;\bar{\gamma}}_{ij} = (\oplus 2)E^{\bar{\gamma}}_{ij}, \quad \text{(II.87)}$$

$$E^{\bar{\gamma}}_{ij} = \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{ij}R^{\beta_1\beta_2}(h), \quad \text{(II.88)}$$

Equations (II.85) and (II.87) are in accord with Eq. (II.1). Obviously, the following definitions:

$$\{\mathbf{Q}_k^{\gamma v ab}\}_{a'i,b'j} = \delta_{aa'}\delta_{bb'}\{\mathbf{M}_k^{\gamma v}\}_{ij}, \quad \text{(II.89)}$$

$$\{\mathbf{Q}_k^{\bar{\gamma} v ab}\}_{a'i,b'j} = \delta_{aa'}\delta_{bb'}\{\mathbf{M}_k^{\bar{\gamma} v}\}_{ij}, \quad \text{(II.90)}$$

allow one to define a further orthonormal basis of $\mathscr{W}^{\beta_1\beta_2;\gamma}$, namely,

$$\mathbf{Q}_k^{\gamma v ab}, \quad a,b = 1,2,\ v = 1,2,\dots,m_{\beta_1\beta_2;\gamma}, \quad k = 1,2,\dots,n_\gamma, \quad \text{(II.91)}$$

$$\mathbf{Q}_k^{\bar{\gamma} v ab}, \quad a,b = 1,2,\ v = 1,2,\dots,m_{\beta_1\beta_2;\gamma}, \quad k = 1,2,\dots,n_\gamma, \quad \text{(II.92)}$$

where the vectors $\mathbf{M}_k^{\gamma v}$ and $\mathbf{M}_k^{\bar{\gamma} v}$ are the corresponding columns of the CG matrix M. In virtue of their definition they transform with respect to H according to

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{Q}_k^{\gamma v ab} = \sum_l R^{\gamma}_{lk}(h)\mathbf{Q}_l^{\gamma v ab}, \quad \text{(II.93)}$$

$$\mathbb{R}^{\beta_1\beta_2}(h)\mathbf{Q}_k^{\bar{\gamma} v ab} = \sum_l \{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{lk}\mathbf{Q}_l^{\bar{\gamma} v ab},$$

$$\text{for all } h \in H, \quad \text{(II.94)}$$

but are in general not vectors which fulfill Eqs. (II.80) and (II.81).

In order to be able to solve this problem we define unitary transformations which link the elements of the bases (II.83) and (II.91) and (II.92):

$$\mathbf{W}_{1k}^{\gamma w} = \sum_{abv} B_{abv;w}\mathbf{Q}_k^{\gamma v ab}, \quad \text{(II.95)}$$

$$\mathbf{Q}_k^{\gamma v ab} = \sum_w B^*_{abv;w}\mathbf{W}_{1k}^{\gamma w}, \quad k = 1,2,\dots,n_\gamma, \quad \text{(II.96)}$$

$$\mathbf{W}_{2k}^{\gamma w} = \sum_{abv} C_{abv;w}\mathbf{Q}_k^{\bar{\gamma} v ab}, \quad \text{(II.97)}$$

$$\mathbf{Q}_k^{\bar{\gamma} v ab} = \sum_w C^*_{abv;w}\mathbf{W}_{2k}^{\gamma w}, \quad k = 1,2,\dots,n_\gamma, \quad \text{(II.98)}$$

and determine them in such a way that Eqs. (II.80) and (II.81) are satisfied.

For this reason we derive

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{Q}_k^{\gamma v ab *} = \sum_{a'b'v';abv} F_{a'b'v';abv}\,\mathbf{Q}_k^{\bar{\gamma} v ab}, \quad \text{(II.99)}$$

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{Q}_k^{\bar{\gamma} v ab *} = \sum_l R^{\gamma}_{lk}(s^2)\sum_{a'b'v'} F^T_{a'b'v';abv}\mathbf{Q}_l^{\gamma v' a' b'}, \quad \text{(II.100)}$$

where

$$F_{a'b'v';abv} = \{CB^T\}_{a'b'v';abv} = \sum_w C_{a'b'v';w}B_{abv;w},$$

$$a,a',b,b' = 1,2, \quad v,v' = 1,2,\dots,m_{\beta_1\beta_2;\gamma}. \quad \text{(II.101)}$$

These equations transform Eqs. (II.80) and (II.81) as follows:

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}_{1k}^{\gamma w *} = \sum_{w'}\{C^+FB^*\}_{w'w}\mathbf{W}_{2k}^{\gamma w'}, \quad \text{(II.102)}$$

$$\mathbb{R}^{\beta_1\beta_2}(s)\mathbf{W}_{2k}^{\gamma w *} = \sum_l R^{\gamma}_{lk}(s^2)\sum_{w'}\{C^+FB^*\}^T_{w'w}\,\mathbf{W}_{1l}^{\gamma w'}. \quad \text{(II.103)}$$

Hence, if we can find $M_{\beta_1\beta_2;\gamma}$-dimensional unitary matrices B and C which satisfy

$$C = FB^*, \quad \text{(II.104)}$$

the corresponding CG coefficients follow immediately from Eqs. (II.95) and (II.97).

Since F is a unitary matrix, we can choose

$$B = \mathbb{1}_M \Longleftrightarrow C = F \quad \text{(II.105)}$$

as a special solution of Eq. (II.104). Hence, we obtain

$$\mathbf{W}_{1k}^{\gamma(abv)} = \mathbf{Q}_k^{\gamma v ab}, \quad \text{(II.106)}$$

$$\mathbf{W}_{2k}^{\gamma(abv)} = \sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}_k^{\bar{\gamma} v' a' b'},$$

$$a,b = 1,2, \quad v = 1,2,\dots,m_{\beta_1\beta_2;\gamma}, \quad \text{(II.107)}$$

which show that we can take the triplets (a,b,v) as the multiplicity index w:

$$w = (a,b,v), \quad a,b = 1,2 \text{ and } v = 1,2,\dots,m_{\beta_1\beta_2;\gamma}. \quad \text{(II.108)}$$

Accordingly, the problem remains of computing the matrix elements of F. This can be done by means of

$$\langle \mathbf{Q}_k^{\bar{\gamma} v' a' b'}, \mathbb{R}^{\beta_1\beta_2}(s)\mathbf{Q}_k^{\gamma v ab *}\rangle = F_{a'b'v';abv}, \quad \text{(II.109)}$$

whose independence of the index k has to be shown with the aid of the following relations:

$$\mathbb{R}^{\beta_1\beta_2}(s)^+\mathbb{E}^{\bar{\gamma}}_{ij}\mathbb{R}^{\beta_1\beta_2}(s) = \mathbb{E}^{\gamma *}_{ij}. \quad \text{(II.110)}$$

By virtue of the special structure of the vectors $\mathbf{Q}_k^{\gamma v a b}$ and $\mathbf{Q}_k^{\bar{\gamma} v a b}$, the matrix elements (II.109) simplify to

$$F_{a'b'v';abv} = (-1)^{\Delta(a')}\delta_{a',a+1}(-1)^{\Delta(b')}\delta_{b',b+1}F_{v'v}^{\bar{\gamma}(\beta_1\beta_2)}, \tag{II.111}$$

$$F_{v'v}^{\bar{\gamma}(\beta_1\beta_2)} = \langle \mathbf{M}_k^{\bar{\gamma}v'}, U^{\beta_1}\otimes U^{\beta_2}\mathbf{M}_k^{\gamma v *}\rangle, \tag{II.112}$$

where the scalar product in Eq. (II.112) is analogulsy defined. In this connection we have to note that the $m_{\beta_1\beta_2;\gamma}$-dimensional submatrix $F^{\bar{\gamma}(\beta_1\beta_2)}$ is also unitary. Furthermore,

$$\{U^{\beta_1}\otimes U^{\beta_2}\}^{+}E_{ij}^{\bar{\gamma}}\{U^{\beta_1}\otimes U^{\beta_2}\} = E_{ij}^{\gamma *} \tag{II.113}$$

can be used to verify that the matrix elements (II.112) are independent of k.

Now let us assume that it is possible to compute the corresponding columns of the CG matrix M with the aid of the method given in Ref. 3, which implies that their components can be written as

$$\begin{aligned}\{\mathbf{M}_k^{\gamma v}\}_{ij} &= \{\mathbf{M}_k^{\beta_1\beta_2;\gamma(i_v j_v)}\}_{ij}\\ &= \|\mathbf{B}_{a_0}^{\beta_1\beta_2;\gamma(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R_{ii_v}^{\beta_1}(h)R_{jj_v}^{\beta_2}(h)R_{ka_0}^{\gamma *}(h),\\ &v = 1,2,\ldots,m_{\beta_1\beta_2;\gamma},\quad k = 1,2,\ldots,n_\gamma,\end{aligned} \tag{II.114}$$

$$\begin{aligned}\{\mathbf{M}_k^{\bar{\gamma} v}\}_{ij} &= \{\mathbf{M}_k^{\beta_1\beta_2;\bar{\gamma}(i_v j_v)}\}_{ij}\\ &= \|\mathbf{B}_{a_0}^{\beta_1\beta_2;\bar{\gamma}(i_v j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R_{ii_v}^{\beta_1}(h)R_{jj_v}^{\beta_2}(h),\\ &\times\{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{ka_0}^{*},\\ &\qquad v = 1,2,\ldots,m_{\beta_1\beta_2;\gamma},\quad k = 1,2,\ldots,n_\gamma.\end{aligned} \tag{II.115}$$

Accordingly, Eq. (II.112) turns out to be

$$\begin{aligned}F_{v'v}^{\bar{\gamma}(\beta_1\beta_2)} &= \|\mathbf{B}_{a_0}^{\beta_1\beta_2;\bar{\gamma}(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\beta_1\beta_2;\gamma(i_v j_v)}\|^{-1}\\ &\times\frac{n_\gamma}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}\{R^{\beta_2}(h)U^{\beta_2}\}_{j_{v'}j_v}\\ &\times\{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{a_0a_0}^{*}\end{aligned} \tag{II.116}$$

and where the index sets $\{(i_v, j_v\}$ occurring in Eqs. (II.114) and (II.115) are however in general not identical.

Inserting Eq. (II.111) into (II.107), we arrive to the final results

$$\mathbf{W}_{1k}^{\gamma(abv)} = \mathbf{Q}_k^{\gamma v a b}, \tag{II.117}$$

$$\begin{aligned}\mathbf{W}_{2k}^{\gamma(abv)} &= (-1)^{\Delta(a+1)+\Delta(b+1)}\sum_{v'}F_{v'v}^{\bar{\gamma}(\beta_1\beta_2)}\mathbf{Q}_k^{\bar{\gamma}v',a+1,b+1},\\ &a,b = 1,2,\quad v = 1,2,\ldots,m_{\beta_1\beta_2;\gamma},\end{aligned} \tag{II.118}$$

which show that CG coefficients of type III are linked by simple unitary transformations with CG coefficients for H and that the multiplicity problem is solved in a very special way.

SUMMARY

In this paper we considered Kronecker products which are composed of co-unirreps of type II only. Because of $\mathbb{R}^{\beta_1\beta_2}\downarrow H = (\oplus 4)R^{\beta_1\beta_2}$, the fist step of the present approach is to determine a suitable CG matrix M which decomposes $R^{\beta_1\beta_2}$ into a direct sum of its irreducible constituents. Provided this has been done, CG coefficients for co-unirreps are obtainable as follows.

CG coefficients of type I are given in this case by Eqs. (II.37)–(II.40) where the definitions (II.12) have to be taken into account. Therefore, the only problem is to calculate the submatrix $F^{\alpha(\beta_1\beta_2)}$ of the special solution (II.35) of Eq. (II.19):

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbb{1}_m & 0 & 0 & \mathbb{1}_m\\ 0 & i\mathbb{1}_m & -\mathbb{1}_m & 0\\ 0 & i\mathbf{F} & \mathbf{F} & 0\\ -i\mathbf{F} & 0 & 0 & \mathbf{F}\end{bmatrix}.$$

The matrix elements of the $m_{\beta_1\beta_2;\alpha}$-dimensional unitary matrix $\mathbf{F} = F^{\alpha(\beta_1\beta_2)}$ are uniquely fixed through Eq. (II.24), at which its property to be symmetric should be utilized in any case.

In the second case, i.e., the computation of CG coefficients of type II, it suffices to compute by means of Eq. (II.62) the matrix elements of the $m_{\beta_1\beta_2;\beta}$-dimensional antisymmetric unitary submatrix $F^{\beta(\beta_1\beta_2)}$ of

$$B = \begin{bmatrix} \mathbb{1}_m & 0 & 0 & 0\\ 0 & \mathbb{1}_m & 0 & 0\\ 0 & 0 & 0 & F^{\beta(\beta_1\beta_2)}\\ 0 & 0 & -F^{\beta(\beta_1\beta_2)} & 0\end{bmatrix},$$

since the corresponding CG coefficients are given by Eqs. (II.76) and (II.77) where the definitions (II.49) have to be used.

Corresponding to the special solution (II.105) of Eq. (II.104), i.e.,

$$B = \mathbb{1}_M$$

and

$$C = \begin{bmatrix} 0 & 0 & 0 & F^{\bar{\gamma}(\beta_1\beta_2)}\\ 0 & 0 & -F^{\bar{\gamma}(\beta_1\beta_2)} & 0\\ 0 & -F^{\bar{\gamma}(\beta_1\beta_2)} & 0 & 0\\ F^{\bar{\gamma}(\beta_1\beta_2)} & 0 & 0 & 0\end{bmatrix},$$

CG coefficients of type III are given by Eqs. (II.117) and (II.118), where of course the definitions (II.89) and (II.90) have to be used. Consequently, the only problem is to compute the $m_{\beta_1\beta_2;\gamma}$-dimensional unitary matrix $F^{\bar{\gamma}(\beta_1\beta_2)}$, whose matrix elements are uniquely defined by Eq. (II.112).

Summarizing our results, we succeeded in solving the multiplicity problem for each type of CG coefficients without reference to a special magnetic group.

[1] C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[2] R. Dirl, J. Math. Phys. **21**, xxx (1980).
[3] R. Dirl, J. Math. Phys. **20**, 659 (1979).

Clebsch–Gordan coefficients for corepresentations. II ⊗ III

R. Dirl
Institut für theoretische Physik, TU Wien A-1040 Wien, Karlsplatz 13, Austria

(Received 18 December 1978; accepted for publication 25 May 1979)

Clebsch–Gordan coefficients for corepresentations are determined quite generally in terms of such coefficients for the normal subgroup, at which the Kronecker products are composed of corepresentations of type II and III.

INTRODUCTION

This paper continues a series of papers which deal with the problem of computing CG coefficients for corepresentations in terms of appropriately determined CG coefficients for the normal subgroup. Within the present paper we restrict our considerations to Kronecker products which are composed of co-unirreps of type II and III. Like in the previous cases we utilize the representation theory of the normal subgroup by taking convenient CG coefficients for this group for granted.

We organize the material of this paper as follows: The first section is devoted to stating the considered problem and to derive useful symmetry relations for the required multiplicities. In accordance to the possible types of co-unirreps we divide the second section into three parts. For each case we find simple defining equations for the unitary transformations which link CG coefficients for the underlying corepresentation with convenient ones for the normal subgroup. These equations are solved quite generally without reference to a special group.

I. MULTIPLICITIES FOR COREPRESENTATIONS

Throughout this paper we consider the Kronecker product

$$\mathbb{R}^{\beta_1\gamma_2} = \{\mathbb{R}^{\beta_1\gamma_2}(g) = \mathbb{R}^{\beta_1}(g) \otimes \mathbb{R}^{\gamma_2}(g): g\in G\}, \tag{I.1}$$

which defines a $4n_{\beta_1}n_{\gamma_2}$-dimensional unitary corepresentation. Since this representation is in general reducible, there must exist a unitary matrix $W^{\beta_1\gamma_2} = W$ which decomposes $\mathbb{R}^{\beta_1\gamma_2}$ into a direct sum of its irreducible constituents:

$$W^{+}\mathbb{R}^{\beta_1\gamma_2}(g)W^{g} = \sum_{\alpha\in A_{\rm I}} \oplus M_{\beta_1\gamma_2;\alpha}\mathbb{R}^{\alpha}(g) \oplus \sum_{\beta\in A_{\rm II}} \oplus M_{\beta_1\gamma_2;\beta}\mathbb{R}^{\beta}(g) \oplus \sum_{\gamma\in A_{\rm III}} \oplus M_{\beta_1\gamma_2;\gamma}\mathbb{R}^{\gamma}(g), \quad \text{for all } g\in G. \tag{I.2}$$

The multiplicities occuring in Eq. (I.2) are well known[1] and take the following values:

$$M_{\beta_1\gamma_2;\alpha} = 2\{m_{\beta_1\gamma_2;\alpha} + m_{\beta_1\bar{\gamma}_2;\alpha}\}, \tag{I.3}$$

$$M_{\beta_1\gamma_2;\beta} = m_{\beta_1\gamma_2;\beta} + m_{\beta_1\bar{\gamma}_2;\beta}, \tag{I.4}$$

$$M_{\beta_1\gamma_2;\gamma} = 2\{m_{\beta_1\gamma_2;\gamma} + m_{\beta_1\bar{\gamma}_2;\gamma}\} = M_{\beta_1\gamma_2;\bar{\gamma}}. \tag{I.5}$$

By similar arguments as in Ref. 2 we derive the symmetry relations

$$m_{\beta_1\gamma_2;\mu} = m_{\beta_1\bar{\gamma}_2;\bar{\mu}}, \quad \text{for all } \mu\in A_H, \tag{I.6}$$

which allow one to simplify the multiplicity formulas as follows:

$$M_{\beta_1\gamma_2;\alpha} = 4m_{\beta_1\gamma_2;\alpha}, \tag{I.7}$$

$$M_{\beta_1\gamma_2;\beta} = 2m_{\beta_1\gamma_2;\beta}, \tag{I.8}$$

$$M_{\beta_1\gamma_2;\gamma} = M_{\beta_1\bar{\gamma}_2;\bar{\gamma}} = M_{\beta_1\gamma_2;\bar{\gamma}}, \tag{I.9}$$

but where we cannot conclude from the last equation that the multiplicities $m_{\beta_1\gamma_2;\gamma}$ and $m_{\beta_1\bar{\gamma}_2;\gamma}$ are equal.

II. CG COEFFICIENTS FOR COREPRESENTATIONS

Due to the present approach the first task is to find out what kind of Kronecker products of unirreps with respect to the normal subgroup H are contained in the subduced representation

$$\mathbb{R}^{\beta_1\gamma_2}\downarrow H = (\oplus 2)\{R^{\beta_1\gamma_2} \oplus R^{\beta_1\bar{\gamma}_2}\}. \tag{II.1}$$

Thereby, we have introduced the notation

$$R^{\beta_1\gamma_2} = \{R^{\beta_1\gamma_2}(h) = R^{\beta_1}(h) \otimes R^{\gamma_2}(h): h\in H\}, \tag{II.2}$$

$$R^{\beta_1\bar{\gamma}_2} = \{R^{\beta_1\bar{\gamma}_2}(h) = R^{\beta_1}(h) \otimes Z^{\gamma_2+}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}: h\in H\}, \tag{II.3}$$

which distinguishes the different Kronecker products occuring twice into $\mathbb{R}^{\beta_1\gamma_2}\downarrow H$. Since both $n_{\beta_1}n_{\gamma_2}$-dimensional unitary representations $R^{\beta_1\gamma_2}$ and $R^{\beta_1\bar{\gamma}_2}$ are in general reducible, there must exist two unitary matrices $M^{\beta_1\gamma_2} = M$ and $N^{\beta_1\gamma_2} = N$ satisfying

$$M^{\dagger}R^{\beta_1\gamma_2}(h)M = \sum_{\alpha\in A_{\rm I}} \oplus m_{\beta_1\gamma_2;\alpha}R^{\alpha}(h) \oplus \sum_{\beta\in A_{\rm II}} \oplus m_{\beta_1\gamma_2;\beta}R^{\beta}(h) \oplus \sum_{\beta\in A_{\rm III}} \oplus \{m_{\beta_1\gamma_2;\gamma}R^{\gamma}(h) \oplus m_{\beta_1\bar{\gamma}_2;\gamma}Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}, \tag{II.4}$$

$$N^{\dagger}R^{\beta_1\bar{\gamma}_2}(h)N = \sum_{\alpha\in A_{\rm I}} \oplus m_{\beta_1\gamma_2;\alpha}R^{\alpha}(h) \oplus \sum_{\beta\in A_{\rm II}} \oplus m_{\beta_1\gamma_2;\beta}R^{\beta}(h) \oplus \sum_{\gamma\in A_{\rm III}} \oplus \{m_{\beta_1\bar{\gamma}_2;\gamma}R^{\gamma}(h) \oplus m_{\beta_1\gamma_2;\gamma}Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}, \quad \text{for all } h\in H, \tag{II.5}$$

respectively, whose matrix elements are taken for granted.

A. CG coefficients of type I

Let us start from the defining equations for CG coefficients of type I:

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{W}_l^{\alpha w} = \sum_l R_{lk}^{\alpha}(h)\mathbf{W}_l^{\alpha w}, \quad \text{for all } h\in H, \tag{II.6}$$

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U_{lk}^{\alpha}\mathbf{W}_l^{\alpha w};$$
$$w = 1,2,...,M_{\beta_1\gamma_2;\alpha}, \quad k = 1,2,...,n_\alpha, \tag{II.7}$$

where the components of the column vectors $\mathbf{W}_k^{\alpha w}$ are given by

$$\{\mathbf{W}_k^{\alpha w}\}_{ai,bj} = \{\mathbf{W}_k^{\beta_1\gamma_2;\alpha w}\}_{ai,bj} = W^{\beta_1\gamma_2}_{ai,bj;\alpha wk}$$
$$\alpha\in A_1, \quad w = 1,2,...,M_{\beta_1\gamma_2;\alpha}, \quad k = 1,2,...,n_\alpha,$$
$$a = 1,2 \text{ and } i = 1,2,...,n_{\beta_1}, \quad b = 1,2 \text{ and } j = 1,2,...,n_{\gamma_2}. \tag{II.8}$$

The double indices (a,i) and (b,j) originate from Eqs. (I.21) and (I.22) and (I.23) and (I.24), respectively, of Ref. 3. Like in the previous cases we interpret the columns of the unitary $4n_{\beta_1}n_{\gamma_2}$-dimensional CG matrix W as H-adapted vectors of a corresponding Euclidean space $\mathscr{W}^{\beta_1\gamma_2}$, which have to satisfy additionally Eq. (II.7).

Consequently, the vectors

$$\mathbf{W}_k^{\alpha w}, \quad w = 1,2,...,M_{\beta_1\gamma_2;\alpha}, \quad k = 1,2,...,n_\alpha, \tag{II.9}$$

form an orthonormal basis of

$$\mathscr{W}^{\beta_1\gamma_2;\alpha} = \sum_i \mathbb{E}_{ii}^{\alpha}\mathscr{W}^{\beta_1\gamma_2},$$
$$\dim \mathscr{W}^{\beta_1\gamma_2;\alpha} = n_\alpha M_{\beta_1\gamma_2;\alpha}, \tag{II.10}$$

where the corresponding units $\mathbb{E}_{ij}^{\alpha}$ decompose in accordance with Eq. (II.1) into a direct sum of the submatrices $E_{ij}^{\beta_1\gamma_2;\alpha}$ and $E_{ij}^{\beta_1\bar{\gamma}_2;\alpha}$:

$$\mathbb{E}_{ij}^{\alpha} = \mathbb{E}_{ij}^{\beta_1\gamma_2;\alpha} = (\oplus 2)\{E_{ij}^{\beta_1\gamma_2;\alpha} \oplus E_{ij}^{\beta_1\bar{\gamma}_2;\alpha}\}, \tag{II.11}$$

$$E_{ij}^{\beta_1\gamma_2;\alpha} = \frac{n_\alpha}{|H|}\sum_h R_{ij}^{\alpha *}(h)R^{\beta_1\gamma_2}(h), \tag{II.12}$$

$$E_{ij}^{\beta_1\bar{\gamma}_2;\alpha} = \frac{n_\alpha}{|H|}\sum_h R_{ij}^{\alpha *}(h)R^{\beta_1\bar{\gamma}_2}(h). \tag{II.13}$$

The structure of $\mathbb{R}^{\beta_1\gamma_2}\downarrow H$ suggests a definition by means of

$$\{\mathbf{Q}_k^{\alpha va1}\}_{a'i,b'j} = \delta_{aa'}\delta_{b'1}\{\mathbf{M}_k^{\alpha v}\}_{ij}, \quad a = 1,2, \tag{II.14}$$

$$\{\mathbf{Q}_k^{\alpha va2}\}_{a'i,b'j} = \delta_{aa'}\delta_{b'2}\{\mathbf{N}_k^{\alpha v}\}_{ij}, \quad a = 1,2, \tag{II.15}$$

a further orthonormal basis of $\mathscr{W}^{\beta_1\gamma_2;\alpha}$, namely,

$$\mathbf{Q}_k^{\alpha vab}, \quad a,b = 1,2, \ v = 1,2,...,m_{\beta_1\gamma_2;\alpha},$$
$$k = 1,2,...,n_\alpha, \tag{II.16}$$

which are especially suited to simplifying the following considerations, since they are already H adapted, i.e.,

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{Q}_k^{\alpha vab} = \sum_l R_{lk}^{\alpha}(h)\mathbf{Q}_l^{\alpha vab}, \quad \text{for all } h\in H. \tag{II.17}$$

Despite this transformation law we cannot expect that these vectors transform also according to Eq. (II.7). Nevertheless, the elements of the bases (II.9) and (II.16) must be linked by unitary transformations which are independent of the index k:

$$\mathbf{W}_k^{\alpha w} = \sum_{abv} B_{abv;w}\mathbf{Q}_k^{\alpha vab}, \tag{II.18}$$

$$\mathbf{Q}_k^{\alpha vab} = \sum_w B^*_{abv;w}\mathbf{W}_k^{\alpha w}, \quad k = 1,2,...,n_\alpha. \tag{II.19}$$

In order to be able to determine unitary matrices B, so that the corresponding vectors (II.18) are satisfying Eq. (II.7), we derive

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{Q}_k^{\alpha vab *} = \sum_l U_{lk}^{\alpha}\sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}_l^{\alpha v'a'b'}, \tag{II.20}$$

where

$$F_{a'b'v';abv} = \{BB^T\}_{abv;a'b'v'} = \sum_w B_{abv;w}B_{a'b'v';w},$$
$$a,a',b,b' = 1,2, \quad v,v' = 1,2,...,m_{\beta_1\gamma_2;\alpha}. \tag{II.21}$$

By utilizing these relations we obtain

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U_{lk}^{\alpha}\sum_{w'}\{B^\dagger FB^*\}_{w'w}\mathbf{W}_l^{\alpha w'}. \tag{II.22}$$

Hence, any $M_{\beta_1\gamma_2;\alpha}$-dimensional matrix B satisfying

$$FB^* = B, \quad \text{with } BB^+ = B^+B = \mathbb{1}_M, \tag{II.23}$$

allows one to write down the corresponding CG coefficients of type I. By similar arguments we can show that F is a symmetric unitary matrix, i.e.,

$$FF^* = \mathbb{1}_M = \mathbb{1}_{4m}. \tag{II.24}$$

The next problem is to compute the matrix elements of F. This can be done by means of

$$\left\langle \mathbf{Q}_k^{\alpha v'a'b'}, \mathbb{R}^{\beta_1\gamma_2}(s)\left\{\sum_l U_{kl}^{\alpha}\mathbf{Q}_l^{\alpha vab}\right\}^*\right\rangle = F_{a'b'v';abv}, \tag{II.25}$$

whose values are independent of the index k. This property can be readily verified with the aid of

$$\mathbb{R}^{\beta_1\gamma_2}(s)^\dagger\mathbb{E}_{ij}^{\alpha}\mathbb{R}^{\beta_1\gamma_2}(s) = \sum_{kl} U_{ik}^{\alpha *}U_{jl}^{\alpha}\mathbb{E}_{kl}^{\alpha *}. \tag{II.26}$$

If we take the special structure of the vectors $\mathbf{Q}_k^{\alpha vab}$ into account, the matrix elements (II.25) simplify to

$$F_{a'bv';abv} = 0, \quad b = 1,2, \tag{II.27}$$

$$F_{a'1v';a2v} = (-1)^{\Delta(\alpha')}\delta_{a',a+1}F_{v'v}^{\alpha(\beta_1\gamma_2)}, \tag{II.28}$$

$$F_{a'2v';a1v} = (-1)^{\Delta(\alpha')}\delta_{a',a+1}F_{v'v}^{\alpha(\beta_1\bar{\gamma}_2)}, \tag{II.29}$$

$$F_{v'v}^{\alpha(\beta_1\gamma_2)} = \left\langle \mathbf{M}_k^{\alpha v'}, U^{\beta_1}\oplus R^{\gamma_2}(s^2)\left\{\sum_l U_{kl}^{\alpha}\mathbf{N}_l^{\alpha v}\right\}^*\right\rangle, \tag{II.30}$$

$$F_{v'v}^{\alpha(\beta_1\bar{\gamma}_2)} = \left\langle \mathbf{N}_k^{\alpha v'}, U^{\beta_1}\oplus \mathbb{1}_{\gamma_2}\left\{\sum_l U_{kl}^{\alpha}\mathbf{M}_l^{\alpha v}\right\}^*\right\rangle, \tag{II.31}$$

where the scalar product on the right hand side of Eqs. (II.30) and (II.31) is analogously defined. The following relations:

$$\{U^{\beta_1}\otimes R^{\gamma_2}(s^2)\}^\dagger E_{ij}^{\beta_1\gamma_2;\alpha}\{U^{\beta_1}\otimes R^{\gamma_2}(s^2)\}$$
$$= \sum_{kl} U_{ik}^{\alpha *}U_{jl}^{\alpha}\{E_{kl}^{\beta_1\bar{\gamma}_2;\alpha}\}^*, \tag{II.32}$$

$$\{U^{\beta_1}\otimes \mathbb{1}_{\gamma_2}\}^\dagger E_{ij}^{\beta_1\bar{\gamma}_2;\alpha}\{U^{\beta_1}\otimes \mathbb{1}_{\gamma_2}\}$$
$$= \sum_{kl} U_{ik}^{\alpha *}U_{jl}^{\alpha}\{E_{kl}^{\beta_1\gamma_2;\alpha}\}^*, \tag{II.33}$$

which connect the units (II.12) and (II.13), have to be used among others to prove that Eqs. (II.30) and (II.31) are independent of k. Now let us consider in more detail the matrix

F:

$$F = \begin{bmatrix} 0 & 0 & 0 & F^{\alpha(\beta_1\gamma_2)} \\ 0 & 0 & F^{\alpha(\beta_1\bar{\gamma}_2)} & 0 \\ 0 & -F^{\alpha(\beta_1\gamma_2)} & 0 & 0 \\ -F^{\alpha(\beta_1\bar{\gamma}_2)} & 0 & 0 & 0 \end{bmatrix}. \tag{II.34}$$

Since F is a symmetric unitary matrix, it follows for the $m_{\beta_1\gamma_2;\alpha}$-dimensional submatrices $F^{\alpha(\beta_1\gamma_2)}$ and $F^{\alpha(\beta_1\bar{\gamma}_2)}$ that they are linked by transposition and that both are unitary, i.e.,

$$F^{\alpha(\beta_1\gamma_2)T} = -F^{\alpha(\beta_1\bar{\gamma}_2)}, \tag{II.35}$$

$$F^{\alpha(\beta_1\gamma_2)}F^{\alpha(\beta_1\bar{\gamma}_2)*} = -1_m . \tag{II.36}$$

Consequently, it suffices to calculate one of them by virtue of Eq. (II.35).

Like in the previous papers we assume that it is possible to calculate the matrix elements of the CG matrices M and N with the aid of the method given in Ref. 4. Therefore, their matrix elements can be written as

$$\begin{aligned}\{\mathbf{M}_k^{\alpha v}\}_{ij} &= \{\mathbf{M}_k^{\beta_1\gamma_2;\alpha(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\beta_1\gamma_2;\alpha(i_v j_v)}\|^{-1}\frac{n_\alpha}{|H|}\sum_h R^{\beta_1}_{ii_v}(h)R^{\gamma_2}_{jj_v}(h)R^{\alpha *}_{ka_0}(h), \\ &\qquad v = 1,2,\ldots,m_{\beta_1\gamma_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha ,\end{aligned} \tag{II.37}$$

$$\begin{aligned}\{\mathbf{N}_k^{\alpha v}\}_{ij} &= \{\mathbf{N}_k^{\beta_1\bar{\gamma}_2;\alpha(i_v j_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\beta_1\bar{\gamma}_2;\alpha(i_v j_v)}\|^{-1}\frac{n_\alpha}{|H|}\sum_h R^{\beta_1}_{ii_v}(h) \\ &\quad\times\{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{jj_v}R^{\alpha *}_{ka_0}(h), \\ &\qquad v = 1,2,\ldots,m_{\beta_1\gamma_2;\alpha}, \quad k = 1,2,\ldots,n_\alpha ,\end{aligned} \tag{II.38}$$

with in general different sets $\{(i_v, j_v)\}$. Equations (II.30) and (II.38) turn out to be

$$\begin{aligned}F^{\alpha(\beta_1\gamma_2)}_{v'v} &= \|\mathbf{B}_{a_0}^{\beta_1\gamma_2;\alpha(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\beta_1\bar{\gamma}_2;\alpha(i_v j_v)}\|^{-1} \\ &\quad\times\frac{n_\alpha}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}R^{\gamma_2}_{j_{v'}j_v}(hs^2)\mathbb{R}^{\alpha *}_{a_0a_0}(hs),\end{aligned} \tag{II.39}$$

$$\begin{aligned}F^{\alpha(\beta_1\bar{\gamma}_2)}_{v'v} &= \|\mathbf{B}_{a_0}^{\beta_1\bar{\gamma}_2;\alpha(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\beta_1\gamma_2;\alpha(i_v j_v)}\|^{-1} \\ &\quad\times\frac{n_\alpha}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}\{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{j_{v'}j_v} \\ &\quad\times\mathbb{R}^{\alpha *}_{a_0a_0}(h),\end{aligned} \tag{II.40}$$

but must satisfy in any way the relation (II.35).

Apart from these special values for the matrix elements of F, we are now in the position to solve Eq. (II.23). Because of the relatively complicated structure (II.34) of F, we are forced to make a more general ansatz for B:

$$B = \begin{bmatrix} \mathbf{A} & 0 & 0 & \mathbf{F}\mathbf{D}^* \\ 0 & \mathbf{B} & \overline{\mathbf{F}}\mathbf{C}^* & 0 \\ 0 & -\mathbf{F}\mathbf{B}^* & \mathbf{C} & 0 \\ -\overline{\mathbf{F}}\mathbf{A}^* & 0 & 0 & \mathbf{D} \end{bmatrix}. \tag{II.41}$$

For the sake of simplicity we have introduced the abbreviated notation $\mathbf{F} = F^{\alpha(\beta_1\gamma_2)}$ and $\overline{\mathbf{F}} = F^{\alpha(\beta_1\bar{\gamma}_2)}$. Furthermore, the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ shall be proportional by numerical factors to $m_{\beta_1\gamma_2;\alpha}$-dimensional unitary ones, but otherwise arbitrary. Now it is easy to verify that any matrix B of the type (II.41) is a solution of Eq. (II.23). Since B is required to be unitary, we obtain the following restricting conditions:

$$\mathbf{A}\mathbf{A}^\dagger + \mathbf{D}\mathbf{D}^\dagger = \mathbf{1}_m \quad\text{and}\quad \mathbf{B}\mathbf{B}^\dagger + \mathbf{C}\mathbf{C}^\dagger = \mathbf{1}_m , \tag{II.42}$$

$$F^{\alpha(\beta_1\gamma_2)}(\mathbf{B}\mathbf{B}^T)^* = (CC^T)F^{\alpha(\beta_1\bar{\gamma}_2)\dagger}, \tag{II.43}$$

$$F^{\alpha(\beta_1\bar{\gamma}_2)}(\mathbf{A}\mathbf{A}^T)^* = (\mathbf{D}\mathbf{D}^T)F^{\alpha(\beta_1\gamma_2)\dagger}. \tag{II.44}$$

Consequently, it is obvious to choose

$$\mathbf{A} = \mathbf{B} = \frac{i}{\sqrt{2}}\mathbf{1}_m , \tag{II.45}$$

which implies

$$\mathbf{C} = -\frac{1}{\sqrt{2}}F^{\alpha(\beta_1\gamma_2)} \quad\text{and}\quad \mathbf{D} = -\frac{1}{\sqrt{2}}F^{\alpha(\beta_1\bar{\gamma}_2)} \tag{II.46}$$

so that the corresponding matrix B reads as

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbf{1}_m & 0 & 0 & \mathbf{1}_m \\ 0 & i\mathbf{1}_m & \mathbf{1}_m & 0 \\ 0 & i\mathbf{F} & -\mathbf{F} & 0 \\ i\overline{\mathbf{F}} & 0 & 0 & -\overline{\mathbf{F}} \end{bmatrix}. \tag{II.47}$$

Obviously, this special solution allows one to identify the multiplicity index w with the triplet (a,b,v), i.e.,

$$w = (a,b,v); \quad a,b = 1,2 \ \text{and}\ v = 1,2,\ldots,m_{\beta_1\gamma_2;\alpha} . \tag{II.48}$$

The corresponding CO coefficients of type I follow immediately from Eq. (II.18):

$$\mathbf{W}_k^{\alpha(11v)} = \frac{i}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v11} - \sum_{v'}F^{\alpha(\beta_1\gamma_2)}_{vv'}\mathbf{Q}_k^{\alpha v'22}\Big\}, \tag{II.49}$$

$$\mathbf{W}_k^{\alpha(12v)} = \frac{i}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v12} + \sum_{v'}F^{\alpha(\beta_1\gamma_2)}_{v'v}\mathbf{Q}_k^{\alpha v'21}\Big\}, \tag{II.50}$$

$$\mathbf{W}_k^{\alpha(21v)} = \frac{1}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v12} - \sum_{v'}F^{\alpha(\beta_1\gamma_2)}_{v'v}\mathbf{Q}_k^{\alpha v'21}\Big\}, \tag{II.51}$$

$$\begin{aligned}\mathbf{W}_k^{\alpha(22v)} &= \frac{1}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v11} + \sum_{v'}F^{\alpha(\beta_1\gamma_2)}_{vv'}\mathbf{Q}_k^{\alpha v'22}\Big\}, \\ &\qquad v = 1,2,\ldots,m_{\beta_1\gamma_2;\alpha} .\end{aligned} \tag{II.52}$$

To summarize our results, we have shown that CG coefficients of type I for corepresentations are linked by simple unitary transformations with convenients CG coefficients for H and that it suffices to calculate for example, $F^{\alpha(\beta_1\gamma_2)}$.

B. CG coefficients of type II

As already known the defining equations for CG coefficients of type II read as

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{W}_{dk}^{\beta w} = \sum_l R^{\beta}_{lk}(h)\mathbf{W}_{dl}^{\beta w}, \quad \text{for all } h\in H, \tag{II.53}$$

$$\begin{aligned}&\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}_{dk}^{\beta w *} = (-1)^{\Delta(d+1)}\sum_l U^{\beta}_{lk}\mathbf{W}_{d+1,l}^{\beta w}, \\ &w = 1,2,\ldots,M_{\beta_1\gamma_2;\beta}, \ d = 1,2, \ \text{and}\ k = 1,2,\ldots,n_\beta ,\end{aligned} \tag{II.54}$$

where

$$\{\mathbf{W}_{dk}^{\beta w}\}_{ai,bj} = \{\mathbf{W}_{dk}^{\beta_1\gamma_2\beta w}\}_{ai,bj} = W^{\beta_1\gamma_2}_{ai,bj,\beta w dk} ,$$

$\beta \in A_{11}$, $w = 1,2,...,M_{\beta_1\gamma_2;\beta}$, $d = 1,2$ and $k = 1,2,...,n_\beta$, $a = 1,2$ and $i = 1,2,...,n_{\beta_1}$, $b = 1,2$ and $j = 1,2,...,n_{\gamma_2}$. (II.55)

Since the CG matrix W is assumed to be unitary, the following vectors:

$$\mathbf{W}^{\beta w}_{dk}, \quad w = 1,2,...,M_{\beta_1\gamma_2;\beta}, \quad d = 1,2, \text{ and } k = 1,2,...,n_\beta, \tag{II.56}$$

form an orthonormal basis of

$$\mathscr{W}^{\beta_1\gamma_2;\beta} = \sum_i \mathbb{E}^\beta_{ii} \mathscr{W}^{\beta_1\gamma_2},$$

$$\dim \mathscr{W}^{\beta_1\gamma_2;\beta} = 2n_\beta M_{\beta_1\gamma_2;\beta}, \tag{II.57}$$

where the units $\mathbb{E}^\beta_{ij}$ can be written as

$$\mathbb{E}^\beta_{ij} = \mathbb{E}^{\beta_1\gamma_2;\beta}_{ij} = (\oplus 2)\{E^{\beta_1\gamma_2;\beta}_{ij} \oplus E^{\beta_1\bar{\gamma}_2;\beta}_{ij}\}, \tag{II.58}$$

$$E^{\beta_1\gamma_2;\beta}_{ij} = \frac{n_\beta}{|H|} \sum_h R^{\beta *}_{ij}(h) R^{\beta_1\gamma_2}(h), \tag{II.59}$$

$$E^{\beta_1\bar{\gamma}_2;\beta}_{ij} = \frac{n_\beta}{|H|} \sum_h R^{\beta *}_{ij}(h) R^{\beta_1\bar{\gamma}_2}(h), \tag{II.60}$$

which is in accordance with Eq. (II.1). Similar to the previous case, the structure of $\mathbb{R}^{\beta_1\gamma_2}\downarrow H$ suggests that one can define by means of

$$\{\mathbf{Q}^{\beta v a 1}_k\}_{a'i,b'j} = \delta_{aa'}\delta_{b'1}\{\mathbf{M}^{\beta v}_k\}_{ij}, \quad a = 1,2, \tag{II.61}$$

$$\{\mathbf{Q}^{\beta v a 2}_k\}_{a'i,b'j} = \delta_{aa'}\delta_{b'2}\{\mathbf{N}^{\beta v}_k\}_{ij}, \quad a = 1,2, \tag{II.62}$$

a further orthonormal basis of $\mathscr{W}^{\beta_1\gamma_2;\beta}$, namely,

$$\mathbf{Q}^{\beta vab}_k, \quad a,b = 1,2, \quad v = 1,2,...,m_{\beta_1\gamma_2;\beta}, \quad k = 1,2,...,n_\beta. \tag{II.63}$$

Although these vectors transform already according to

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{Q}^{\beta vab}_k = \sum_l R^\beta_{lk}(h)\mathbf{Q}^{\beta vab}_l, \quad \text{for all } h \in H, \tag{II.64}$$

we cannot assume that they are also solutions of Eq. (II.54) which have to be satisfied in any way.

Therefore, we consider

$$\mathbf{W}^{\beta w}_{dk} = \sum_{abv} B_{abv;dw}\mathbf{Q}^{\beta vab}_k, \tag{II.65}$$

$$\mathbf{Q}^{\beta vab}_k = \sum_{dw} B^*_{abv;dw}\mathbf{W}^{\beta w}_{dk}, \quad k = 1,2,...,n_\beta, \tag{II.66}$$

and determine the $2M_{\beta_1\gamma_2;\beta}$-dimensional unitary matrices B in such way that the corresponding vectors (II.65) satisfy Eq. (II.54).

The following relations:

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{Q}^{\beta vab *}_k = \sum_l U^\beta_{lk} \sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}^{\beta v'a'b'}_l, \tag{II.67}$$

containing the definitions

$$G_{dw;d'w'} = (-1)^{\Delta(d+1)}\delta_{d',d+1}\delta_{ww'}, \quad d,d' = 1,2, \quad w,w' = 1,2,...,M_{\beta_1\gamma_2;\beta}, \tag{II.68}$$

$$F_{a'b'v';abv} = \{BGB^T\}_{abv;a'b'v'} = \sum_{dw} B_{abv;dw}(-1)^{\Delta(d+1)}B_{a'b'v';d+1,w},$$

$$a,a',b,b' = 1,2, \quad v,v' = 1,2,...,m_{\beta_1\gamma_2;\beta}, \tag{II.69}$$

have to be used to transform Eq. (II.54) as

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}^{\beta w *}_{dk} = \sum_l U^\beta_{lk} \sum_{d'w'} \{B^\dagger F B^*\}_{d'w';dw}\mathbf{W}^{\beta w'}_{d'l}. \tag{II.70}$$

Hence, if we can find a unitary matrix B satisfying

$$BG^T = FB^*, \tag{II.71}$$

the corresponding CG coefficients of type II are immediately obtained from Eq. (II.65), at which F represents a antisymmetric unitary matrix, i.e.,

$$FF^* = -\mathbb{1}_{2M} = -\mathbb{1}_{4m}. \tag{II.72}$$

The matrix elements of F are uniquely fixed through Eq. (II.67) by calculating the scalar products

$$\left\langle \mathbf{Q}^{\beta v'a'b'}_k, \mathbb{R}^{\beta_1\gamma_2}(s)\left\{\sum_l U^\beta_{kl}\mathbf{Q}^{\beta vab}_l\right\}^*\right\rangle = F_{a'b'v';abv}, \tag{II.73}$$

whose values must be independent of k. This assertion can be proven by means of

$$\mathbb{R}^{\beta_1\gamma_2}(s)^+\mathbb{E}^\beta_{ij}\mathbb{R}^{\beta_1\gamma_2}(s) = \sum_{kl} U^{\beta *}_{ik}U^\beta_{jl}\mathbb{E}^{\beta *}_{kl} \tag{II.74}$$

and the transformation properties (II.64). The matrix elements (II.73) simplify to

$$F_{a'bv';abv} = 0, \quad b = 1,2, \tag{II.75}$$

$$F_{a'1v';a2v} = (-1)^{\Delta(a')}\delta_{a',a+1}F^{\beta(\beta_1\gamma_2)}_{v'v}, \tag{II.76}$$

$$F_{a'2v';a1v} = (-1)^{\Delta(a')}\delta_{a',a+1}F^{\beta(\beta_1\bar{\gamma}_2)}_{v'v}, \tag{II.77}$$

$$F^{\beta(\beta_1\gamma_2)}_{v'v} = \left\langle \mathbf{M}^{\beta v'}_k, U^{\beta_1}\otimes R^{\gamma_2}(s^2)\left\{\sum_l U^\beta_{kl}\mathbf{N}^{\beta v}_l\right\}^*\right\rangle, \tag{II.78}$$

$$F^{\beta(\beta_1\bar{\gamma}_2)}_{v'v} = \left\langle \mathbf{N}^{\beta v'}_k, U^{\beta_1}\otimes \mathbb{1}_{\gamma_2}\left\{\sum_l U^\beta_{kl}\mathbf{M}^{\beta v}_l\right\}^*\right\rangle, \tag{II.79}$$

if inserting Eqs. (II.61) and (II.62) into Eq. (II.73) and where the scalar product occuring in Eqs. (II.78) and (II.79) is analogously defined. The values (II.78) and (II.79) are independent of k. The last proposition can be shown by using

$$\{U^{\beta_1}\otimes R^{\gamma_2}(s^2)\}^+ E^{\beta_1\gamma_2;\beta}_{ij}\{U^{\beta_1}\otimes R^{\gamma_2}(s^2)\} = \sum_{kl} U^{\beta *}_{ik}U^\beta_{jl}\{E^{\beta_1\bar{\gamma}_2;\beta}_{kl}\}^*, \tag{II.80}$$

$$\{U^{\beta_1}\otimes \mathbb{1}_{\gamma_2}\}^+ E^{\beta_1\bar{\gamma}_2;\beta}_{ij}\{U^{\beta_1}\otimes \mathbb{1}_{\gamma_2}\} = \sum_{kl} U^{\beta *}_{ik}U^\beta_{jl}\{E^{\beta_1\gamma_2;\beta}_{kl}\}^*, \tag{II.81}$$

and the fact that the vectors $\mathbf{M}^{\beta v}_k$ and $\mathbf{N}^{\beta v}_k$ transform according to the unirrep R^β. Summarizing our results, the matrix F can be written as

$$F = \begin{bmatrix} 0 & 0 & 0 & F^{\beta(\beta_1\gamma_2)} \\ 0 & 0 & F^{\beta(\beta_1\bar{\gamma}_2)} & 0 \\ 0 & -F^{\beta(\beta_1\gamma_2)} & 0 & 0 \\ -F^{\beta(\beta_1\bar{\gamma}_2)} & 0 & 0 & 0 \end{bmatrix} \tag{II.82}$$

and makes obvious that the following relations must hold:

$$F^{\beta(\beta_1\gamma_2)T} = F^{\beta(\beta_1\bar{\gamma}_2)}, \tag{II.83}$$

$$F^{\beta(\beta_1\gamma_2)}F^{\beta(\beta_1\bar{\gamma}_2)*} = \mathbb{1}_m, \tag{II.84}$$

since F is a antisymmetric unitary matrix. Consequently, it suffices to determine the $m_{\beta_1\gamma_2;\beta}$-dimensional submatrix $F^{\beta(\beta_1\gamma_2)}$, since the other matrix follows from Eq. (II.83).

Provided the corresponding columns of the CG matrices M and N can be computed by means of the method discussed in Ref. 4, their components take the form

$$\{\mathbf{M}_k^{\beta v}\}_{ij} = \{\mathbf{M}_k^{\beta_1\gamma_2;\beta(i_v j_v)}\}_{ij}$$
$$= \|\mathbf{B}_{a_0}^{\beta_1\gamma_2;\beta(i_v j_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h R^{\beta_1}_{ii_v}(h)R^{\gamma_2}_{jj_v}(h)R^{\beta *}_{ka_0}(h),$$
$$v=1,2,\ldots,m_{\beta_1\gamma_2;\beta}, \quad k=1,2,\ldots,n_\beta, \qquad \text{(II.85)}$$

$$\{\mathbf{N}_k^{\beta v}\}_{ij} = \{\mathbf{N}_k^{\beta_1\bar{\gamma}_2;\beta(i_v j_v)}\}_{ij}$$
$$= \|\mathbf{B}_{a_0}^{\beta_1\bar{\gamma}_2;\beta(i_v j_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h R^{\beta_1}_{ii_v}(h)$$
$$\times\{Z^{\gamma_2 +}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{jj_v}R^{\beta *}_{ka_0}(h),$$
$$v=1,2,\ldots,m_{\beta_1\gamma_2;\beta}, \quad k=1,2,\ldots,n_\beta, \qquad \text{(II.86)}$$

and lead to the following values for Eqs. (II.78) and (II.79):

$$F^{\beta(\beta_1\gamma_2)}_{v'v} = \|\mathbf{B}_{a_0}^{\beta_1\gamma_2;\beta(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\beta_1\bar{\gamma}_2;\beta(i_v j_v)}\|^{-1}$$
$$\times\frac{n_\beta}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}$$
$$\times R^{\gamma_2}_{j_{v'}j_v}(hs^2)\{R^\beta(h)U^\beta\}^*_{a_0a_0}, \qquad \text{(II.87)}$$

$$F^{\beta(\beta_1\bar{\gamma}_2)}_{v'v}$$
$$= \|\mathbf{B}_{a_0}^{\beta_1\bar{\gamma}_2;\beta(i_{v'} j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\beta_1\gamma_2;\beta(i_v j_v)}\|^{-1}$$
$$\times\frac{n_\beta}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}\{Z^{\gamma_2 +}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{j_{v'}j_v}$$
$$\times\{R^\beta(h)U^\beta\}^*_{a_0a_0}, \qquad \text{(II.88)}$$

respectively, at which the same argumentation must hold as before.

Now let us return to the problem of determining unitary matrices B which satisfy Eq. (II.71). For this purpose we introduce once more the vector notation

$$\{\mathbf{B}^{d,w}\}_{abv} = B_{abv,dw},$$
$$d=1,2 \text{ and } w=1,2,\ldots,M_{\beta_1\gamma_2;\beta},$$
$$a,b,=1,2 \text{ and } v=1,2,\ldots,m_{\beta_1\gamma_2;\beta}, \qquad \text{(II.89)}$$

which allows one to rewrite Eq. (II.71) as follows:

$$F\mathbf{B}^{d,w*} = (-1)^{\Delta(d+1)}\mathbf{B}^{d+1,w},$$
$$d=1,2 \text{ and } w=1,2,\ldots,M_{\beta_1\gamma_2;\beta}. \qquad \text{(II.90)}$$

In this connection we have to note that, fixing the vectors $\mathbf{B}^{1,w}$, $w=1,2,\ldots,M_{\beta_1\gamma_2;\beta}$, the remaining vectors $\mathbf{B}^{2,w}$, $w=1,2,\ldots,M_{\beta_1\gamma_2;\beta}$, are uniquely determined through Eq. (II.90). Furthermore, if we can choose the vectors $\mathbf{B}^{1,w}$ in such a way that the corresponding matrix B is unitary, we have solved our problem. This property of B can be achieved, choosing, for example,

$$\{\mathbf{B}^{1,w=(bv)}\}_{ab'v'} = \delta_{a1}\delta_{bb'}\delta_{vv'}, \qquad \text{(II.91)}$$

which allows one to identify the multiplicity index w with the pair (b,v), i.e.,

$$w=(b,v), \quad b=1,2 \text{ and } v=1,2,\ldots,m_{\beta_1\gamma_2;\beta}. \qquad \text{(II.92)}$$

Hence, it follows from Eq. (II.90)

$$\{\mathbf{B}^{2,(1v)}\}_{abv'} = \delta_{a2}\delta_{b2}F^{\beta(\beta_1\bar{\gamma}_2)}_{v'v}, \qquad \text{(II.93)}$$
$$\{\mathbf{B}^{2,(2v)}\}_{abv'} = \delta_{a2}\delta_{b1}F^{\beta(\beta_1\gamma_2)}_{v'v}, \qquad \text{(II.94)}$$

in matrix notation

$$B = \begin{bmatrix} \mathbb{1}_m & 0 & 0 & 0 \\ 0 & \mathbb{1}_m & 0 & 0 \\ 0 & 0 & 0 & F^{\beta(\beta_1\gamma_2)} \\ 0 & 0 & F^{\beta(\beta_1\gamma_2)T} & 0 \end{bmatrix}, \qquad \text{(II.95)}$$

where we have already taken into account the symmetry relation (II.83). Consequently, the corresponding CG coefficients are given by

$$\mathbf{W}_{1k}^{\beta(bv)} = \mathbf{Q}_k^{\beta v 1 b}, \qquad \text{(II.96)}$$
$$\mathbf{W}_{2k}^{\beta(1v)} = \sum_{v'}F^{\beta(\beta_1\gamma_2)}_{vv'}\mathbf{Q}_k^{\beta v' 21}, \qquad \text{(II.97)}$$
$$\mathbf{W}_{2k}^{\beta(2v)} = \sum_{v'}F^{\beta(\beta_1\gamma_2)}_{v'v}\mathbf{Q}_k^{\beta v' 22}, \quad v=1,2,\ldots,m_{\beta_1\gamma_2;\beta}. \qquad \text{(II.98)}$$

To summarize the results, we have shown that CG coefficients of type II can be traced back by simple unitary transformations to convenient CG coefficients for the normal subgroup H, where the only problem is to compute the $m_{\beta_1\gamma_2;\beta}$-dimensional unitary matrix $F^{\beta(\beta_1\gamma_2)}$.

C. CG coefficients of type III

We start our considerations by summarizing the defining equations for CG coefficients of type III for the considered Kronecker product $\mathbb{R}^{\beta_1\gamma_2}$:

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{W}_{1k}^{\gamma w} = \sum_l R^\gamma_{lk}(h)\mathbf{W}_{1l}^{\gamma w}, \qquad \text{(II.99)}$$

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{W}_{2k}^{\gamma w} = \sum_l\{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^\gamma\}_{lk}\mathbf{W}_{2l}^{\gamma w}, \quad \text{for all } h\in H, \qquad \text{(II.100)}$$

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}_{1k}^{\gamma w*} = \mathbf{W}_{2k}^{\gamma w}, \qquad \text{(II.101)}$$

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}_{2k}^{\gamma w*} = \sum_l R^\gamma_{lk}(s^2)\mathbf{W}_{1l}^{\gamma w},$$
$$w=1,2,\ldots,M_{\beta_1\gamma_2;\gamma}, \quad k=1,2,\ldots,n_\gamma. \qquad \text{(II.102)}$$

The vectors $\mathbf{W}_{dk}^{\gamma w}$ whose components are defined by

$$\{\mathbf{W}_{dk}^{\gamma w}\}_{ai,bj} = \{\mathbf{W}_{dk}^{\beta_1\gamma_2;\gamma w}\}_{ai,bj} = W^{\beta_1\gamma_2}_{ai,bj,\gamma wdk},$$
$$\gamma\in A_{\mathrm{III}}, \quad w=1,2,\ldots,M_{\beta_1\gamma_2;\gamma}, \quad d=1,2, \text{ and } k=1,2,\ldots,n_\gamma,$$
$$a=1,2 \text{ and } i=1,2,\ldots,n_{\beta_1}, \quad b=1,2 \text{ and } j=1,2,\ldots,n_{\gamma_2} \qquad \text{(II.103)}$$

represent columns of the unitary CG matrix W.

Consequently, the vectors

$$\mathbf{W}_{dk}^{\gamma w}, \quad w=1,2,\ldots,M_{\beta_1\gamma_2;\gamma}, \quad d=1,2, \text{ and } k=1,2,\ldots,n_\gamma, \qquad \text{(II.104)}$$

define an orthonormal basis of

$$\mathscr{W}^{\beta_1\gamma_2;\gamma} = \sum_i\{\mathbb{E}_{ii}^\gamma + \mathbb{E}_{ii}^{\bar{\gamma}}\}\mathscr{W}^{\beta_1\gamma_2;\gamma},$$
$$\dim\mathscr{W}^{\beta_1\gamma_2;\gamma} = 2n_\gamma M_{\beta_1\gamma_2;\gamma}, \qquad \text{(II.105)}$$

where the corresponding units $\mathbb{E}_{ij}^\gamma$ and $\mathbb{E}_{ij}^{\bar{\gamma}}$ can be written in accordance with Eq. (II.1) as

$$\mathbb{E}_{ij}^{\gamma} = \mathbb{E}_{ij}^{\beta_1\gamma_2;\gamma} = (\oplus 2)\{E_{ij}^{\beta_1\gamma_2;\gamma} \oplus E_{ij}^{\beta_1\bar{\gamma}_2;\gamma}\}, \quad \text{(II.106)}$$

$$E_{ij}^{\beta_1\gamma_2;\gamma} = \frac{n_\gamma}{|H|}\sum_h R_{ij}^{\gamma *}(h) R^{\beta_1\gamma_2}(h), \quad \text{(II.107)}$$

$$E_{ij}^{\beta_1\bar{\gamma}_2;\gamma} = \frac{n_\gamma}{|H|}\sum_h R_{ij}^{\gamma *}(h) R^{\beta_1\bar{\gamma}_2}(h), \quad \text{(II.108)}$$

and

$$\mathbb{E}_{ij}^{\bar{\gamma}} = \mathbb{E}_{ij}^{\beta_1\gamma_2;\bar{\gamma}} = (\oplus 2)\{E_{ij}^{\beta_1\gamma_2;\bar{\gamma}} \oplus E_{ij}^{\beta_1\bar{\gamma}_2;\bar{\gamma}}\}, \quad \text{(II.109)}$$

$$E_{ij}^{\beta_1\gamma_2;\bar{\gamma}} = \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma +} R^{\bar{\gamma}}(h) Z^{\gamma}\}_{ij}^{*} R^{\beta_1\gamma_2}(h), \quad \text{(II.110)}$$

$$E_{ij}^{\beta_1\bar{\gamma}_2;\bar{\gamma}} = \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma +} R^{\bar{\gamma}}(h) Z^{\gamma}\}_{ij}^{*} R^{\beta_1\bar{\gamma}_2}(h), \quad \text{(II.111)}$$

Since the unitary CG matrices M and N are assumed to be known, it is suggestive by virtue of the structure of $\mathbb{R}^{\beta_1\gamma_2}\downarrow H$ to define, by means of

$$\{\mathbf{Q}_k^{\gamma v a 1}\}_{a'i,b'j} = \delta_{aa'}\delta_{b'1}\{\mathbf{M}_k^{\gamma v}\}_{ij}, \quad a = 1,2, \quad \text{(II.112)}$$

$$\{\mathbf{Q}_k^{\gamma v a 2}\}_{a'i,b'j} = \delta_{aa'}\delta_{b'2}\{\mathbf{N}_k^{\gamma v}\}_{ij}, \quad a = 1,2, \quad \text{(II.113)}$$

$$\{\mathbf{Q}_k^{\bar{\gamma} v a 1}\}_{a'i,b'j} = \delta_{aa'}\delta_{b'1}\{\mathbf{M}_k^{\bar{\gamma} v}\}_{ij}, \quad a = 1,2, \quad \text{(II.114)}$$

$$\{\mathbf{Q}_k^{\bar{\gamma} v a 2}\}_{a'i,b'j} = \delta_{aa'}\delta_{b'2}\{\mathbf{N}_k^{\bar{\gamma} v}\}_{ij}, \quad a = 1,2, \quad \text{(II.115)}$$

a further orthonormal basis of $\mathscr{W}^{\beta_1\gamma_2;\gamma}$, namely,

$$\mathbf{Q}_k^{\gamma v a b}, \quad a,b = 1,2, \ v = 1,2,\ldots,m_1(b),$$
$$m_1(1) = m_{\beta_1\gamma_2;\gamma}, \ m_1(2) = m_{\beta_1\bar{\gamma}_2;\gamma}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(II.116)}$$

$$\mathbf{Q}_k^{\bar{\gamma} v a b}, \quad a,b = 1,2, \ v = 1,2,\ldots,m_2(b),$$
$$m_2(1) = m_{\beta_1\bar{\gamma}_2;\gamma}, \ m_2(2) = m_{\beta_1\gamma_2;\gamma}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(II.117)}$$

where we cannot assume from the outset that the multiplicities $m_i(1)$ and $m_i(2)$ are equal. This fact forces us to proceed more carefully than in the foregoing cases. Nevertheless, because of their transformation properties with respect to H, i.e.,

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{Q}_k^{\gamma v a b} = \sum_l R_{lk}^{\gamma}(h)\mathbf{Q}_l^{\gamma v a b}, \quad \text{(II.118)}$$

$$\mathbb{R}^{\beta_1\gamma_2}(h)\mathbf{Q}_k^{\bar{\gamma} v a b} = \sum_l \{Z^{\gamma +} R^{\bar{\gamma}}(h) Z^{\gamma}\}_{lk}\mathbf{Q}_l^{\bar{\gamma} v a b},$$
$$\text{for all } h \in H, \quad \text{(II.119)}$$

the vectors (II.116) and (II.117) are especially suited to simplify the following considerations.

However, since in general the vectors (II.116) and (II.117) do not satisfy Eqs. [II.101) and (II.102), we define the following unitary transformations:

$$\mathbf{W}_{1k}^{\gamma w} = \sum_{abv} B_{abv;w}\mathbf{Q}_k^{\gamma v a b}, \quad \text{(II.120)}$$

$$\mathbf{Q}_k^{\gamma v a b} = \sum_w B_{abv;w}^{*}\mathbf{W}_{1k}^{\gamma w}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(II.121)}$$

$$\mathbf{W}_{2k}^{\gamma w} = \sum_{abv} C_{abv;w}\mathbf{Q}_k^{\bar{\gamma} v a b}, \quad \text{(II.122)}$$

$$\mathbf{Q}_k^{\bar{\gamma} v a b} = \sum_w C_{abv;w}^{*}\mathbf{W}_{2k}^{\gamma w}, \quad k = 1,2,\ldots,n_\gamma, \quad \text{(II.123)}$$

where we have already taken into account Schur's lemma with respect to H and that the unirreps R^{γ} and $R^{\bar{\gamma}}$ are inequivalent. Thereby, we have to note that the indices (a,b,v) of B and C should not be confused, if $m_1(1) \neq m_1(2)$.

In order to be able to determine B and C we consider

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{Q}_k^{\gamma v a b *} = \sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}_k^{\bar{\gamma} v' a' b'}, \quad \text{(II.124)}$$

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{Q}_k^{\bar{\gamma} v a b *} = \sum_l R_{lk}^{\gamma}(s^2) \sum_{a'b'v'} F_{abv;a'b'v'}\mathbf{Q}_l^{\gamma v' a' b'}, \quad \text{(II.125)}$$

where

$$F_{a'b'v';abv} = \{CB^{T}\}_{a'b'v';abv} = \sum_w C_{a'b'v';w}B_{abv;w},$$
$$a'b' = 1,2, \quad v' = 1,2,\ldots,m_2(b'),$$
$$a,b = 1,2, \quad v = 1,2\cdots m_1(b). \quad \text{(II.126)}$$

A simple inspection of

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}_{1k}^{\gamma w *} = \sum_{w'} \{C^{+}FB^{*}\}_{w'w}\mathbf{W}_{2k}^{\gamma w'}, \quad \text{(II.127)}$$

$$\mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{W}_{2k}^{\gamma w *} = \sum_l R_{lk}^{\gamma}(s^2)\sum_{w'}\{C^{+}FB^{*}\}_{w'w}^{T}\mathbf{W}_{1l}^{\gamma w'}, \quad \text{(II.128)}$$

yields

$$C = FB^{*}, \quad \text{(II.129)}$$

representing the defining equation for the $M_{\beta_1\gamma_2;\gamma}$-dimensional unitary matrices B and C.

Since F is a unitary matrix, we can take

$$B = \mathbb{1}_M \Leftrightarrow C = F \quad \text{(II.130)}$$

as a special solution of Eq. (II.129), Equation (II.130) shows that the multiplicity index w can be identified with the triplets (a,b,v), i.e.,

$$w = (a,b,v), \quad a,b = 1,2, \ \text{and} \ v = 1,2,\ldots,m_1(b). \quad \text{(II.131)}$$

Thus we arrive at the final formulas

$$\mathbf{W}_{1k}^{\gamma(abv)} = \mathbf{Q}_k^{\gamma v a b},$$
$$a,b = 1,2 \ \text{and} \ v = 1,2,\ldots,m_1(b), \quad k = 1,2,\ldots,n_\gamma, \quad \text{(II.132)}$$

$$\mathbf{W}_{2k}^{\gamma(abv)} = \sum_{a'b'}\sum_{v'=1}^{m_2(b')} F_{a'b'v';abv}\mathbf{Q}_k^{\bar{\gamma} v' a' b'},$$
$$a,b = 1,2 \ \text{and} \ v = 1,2,\ldots,m_1(b), \ k = 1,2,\ldots,n_\gamma, \quad \text{(II.133)}$$

where the possibility that the multiplicities $m_1(b)$ can take different values should always be taken into account.

The last problem is now to compute the matrix elements of F. This can be done by means of

$$\langle \mathbf{Q}_k^{\bar{\gamma} v' a' b'}, \mathbb{R}^{\beta_1\gamma_2}(s)\mathbf{Q}_k^{\gamma v a b *}\rangle = F_{a'b'v';abv}. \quad \text{(II.134)}$$

The matrix identities

$$\mathbb{R}^{\beta_1\gamma_2}(s)^{+}\mathbb{E}_{ij}^{\bar{\gamma}}\mathbb{R}^{\beta_1\gamma_2}(s) = \mathbb{E}_{ij}^{\gamma *} \quad \text{(II.135)}$$

have to be used when verifying that the matrix elements are independent of the free index k. The matrix elements (II.134) simplify to

$$F_{a'bv';abv} = 0, \quad b = 1,2, \quad \text{(II.136)}$$

$$F_{a'1v';a2v} = (-1)^{\Delta(a')}\delta_{a',a+1} F^{\bar{\gamma}(\beta_1\gamma_2)}_{v'v}, \tag{II.137}$$

$$F_{a'2v';a1v} = (-1)^{\Delta(a')}\delta_{a',a+1} F^{\bar{\gamma}(\beta_1\bar{\gamma}_2)}_{v'v}, \tag{II.138}$$

$$F^{\bar{\gamma}(\beta_1\gamma_2)}_{v'v} = \langle \mathbf{M}^{\bar{\gamma}v'}_k, U^{\beta_1} \otimes R^{\gamma_2}(s^2)\mathbf{N}^{\gamma v *}_k \rangle, \tag{II.139}$$

$$F^{\bar{\gamma}(\beta_1\bar{\gamma}_2)}_{v'v} = \langle \mathbf{N}^{\bar{\gamma}v'}_k, U^{\beta_1} \otimes \mathbf{1}_{\gamma_2}\mathbf{M}^{\gamma v *}_k \rangle, \tag{II.140}$$

where the scalar product in Eqs. (II.139) and (II.140) is analoguously defined. The matrix elements (II.139) and (II.140) are independent of k, which can be proven by means of the relations

$$\{U^{\beta_1} \otimes R^{\gamma_2}(s^2)\}^\dagger E^{\beta_1\gamma_2;\bar{\gamma}}_{ij}\{U^{\beta_1} \otimes R^{\gamma_2}(s^2)\} = \{E^{\beta_1\bar{\gamma}_2;\gamma}_{ij}\}^*, \tag{II.141}$$

$$\{U^{\beta_1} \otimes \mathbf{1}_{\gamma_2}\}^\dagger E^{\beta_1\bar{\gamma}_2;\bar{\gamma}}_{ij}\{U^{\beta_1} \otimes \mathbf{1}_{\gamma_2}\} = \{E^{\beta_1\gamma_2;\gamma}_{ij}\}^*, \tag{II.142}$$

and the fact that the vectors $\mathbf{M}^{\gamma v}_k(\mathbf{N}^{\bar{\gamma}v}_k)$ transform according to the unirrep $R^{\gamma}(R^{\bar{\gamma}})$.

Before summarizing the results, let us assume that the corresponding columns of the CG matrices M and N can be computed with the aid of the method given in Ref. 4. Their components take the following form:

$$\begin{aligned}\{\mathbf{M}^{\gamma v}_k\}_{ij} &= \{\mathbf{M}^{\beta_1\gamma_2;\gamma(i_v,j_v)}_k\}_{ij} \\ &= \|\mathbf{B}^{\beta_1\gamma_2;\gamma(i_v,j_v)}_{a_0}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R^{\beta_1}_{ii_v}(h)R^{\gamma_2}_{jj_v}(h)R^{\gamma *}_{ka_0}(h), \\ &\qquad v = 1,2,\dots,m_1(1), \quad k = 1,2,\dots,n_\gamma,\end{aligned} \tag{II.143}$$

$$\begin{aligned}\{\mathbf{N}^{\gamma v}_k\}_{ij} &= \{\mathbf{N}^{\beta_1\bar{\gamma}_2;\gamma(i_v,j_v)}_k\}_{ij} \\ &= \|\mathbf{B}^{\beta_1\bar{\gamma}_2;\gamma(i_v,j_v)}_{a_0}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R^{\beta_1}_{ii_v}(h) \\ &\quad\times\{Z^{\gamma_2 +}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{jj_v}R^{\gamma *}_{ka_0}(h), \\ &\qquad v = 1,2,\dots,m_1(2), \quad k = 1,2,\dots,n_\gamma,\end{aligned} \tag{II.144}$$

and

$$\begin{aligned}\{\mathbf{M}^{\bar{\gamma}v}_k\}_{ij} &= \{\mathbf{M}^{\beta_1\gamma_2;\bar{\gamma}(i_v,j_v)}_k\}_{ij} \\ &= \|\mathbf{B}^{\beta_1\gamma_2;\bar{\gamma}(i_v,j_v)}_{a_0}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R^{\beta_1}_{ii_v}(h) \\ &\quad\times R^{\gamma_2}_{jj_v}(h)\{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{ka_0}, \\ &\qquad v = 1,2,\dots,m_2(1), \quad k = 1,2,\dots,n_\gamma,\end{aligned} \tag{II.145}$$

$$\begin{aligned}\{\mathbf{N}^{\bar{\gamma}v}_k\}_{ij} &= \{\mathbf{N}^{\beta_1\bar{\gamma}_2;\bar{\gamma}(i_v,j_v)}_k\}_{ij} \\ &= \|\mathbf{B}^{\beta_1\bar{\gamma}_2;\bar{\gamma}(i_v,j_v)}_{a_0}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R^{\beta_1}_{ii_v}(h) \\ &\quad\times\{Z^{\gamma_2 +}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{jj_v}\{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{ka_0}, \\ &\qquad v = 1,2,\dots,m_2(2), \quad k = 1,2,\dots,n_\gamma.\end{aligned} \tag{II.146}$$

Inserting these special values into Eqs. (II.139) and (II.140), we obtain after straightforward calculations

$$\begin{aligned}F^{\bar{\gamma}(\beta_1\gamma_2)}_{v'v} &= \|\mathbf{B}^{\beta_1\gamma_2;\bar{\gamma}(i_{v'}j_{v'})}_{a_0}\|^{-1}\|\mathbf{B}^{\beta_1\bar{\gamma}_2;\gamma(i_vj_v)}_{a_0}\|^{-1} \\ &\quad\times\frac{n_\gamma}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}R^{\gamma_2}_{j_{v'}j_v}(hs^2) \\ &\quad\times\{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{a_0a_0},\end{aligned} \tag{II.147}$$

$$\begin{aligned}F^{\bar{\gamma}(\beta_1\bar{\gamma}_2)}_{v'v} &= \|\mathbf{B}^{\beta_1\bar{\gamma}_2;\bar{\gamma}(i_{v'}j_{v'})}_{a_0}\|^{-1}\|\mathbf{B}^{\beta_1\gamma_2;\gamma(i_vj_v)}_{a_0}\|^{-1} \\ &\quad\times\frac{n_\gamma}{|H|}\sum_h\{R^{\beta_1}(h)U^{\beta_1}\}_{i_{v'}i_v}\{Z^{\gamma_2 +}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{j_{v'}j_v} \\ &\quad\times\{Z^{\gamma +}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{a_0a_0},\end{aligned} \tag{II.148}$$

where the indices (i_v, j_v) appearing in Eqs. (II.143)–(II.146) originate in general from quite different sets.

Hence, our final formulas read as

$$\begin{aligned}\mathbf{W}^{\gamma(abv)}_{1k} &= \mathbf{Q}^{\gamma vab}_k, \\ &\quad a,b = 1,2 \text{ and } v = 1,2,\dots,m_1(b), \quad k = 1,2,\dots,n_\gamma,\end{aligned} \tag{II.149}$$

$$\begin{aligned}\mathbf{W}^{\gamma(a1v)}_{2k} &= (-1)^{\Delta(a+1)}\sum_{v'=1}^{m_2(2)}F^{\bar{\gamma}(\beta_1\bar{\gamma}_2)}_{v'v}\mathbf{Q}^{\bar{\gamma}v',a+1,2}_k, \\ &\quad a = 1,2 \text{ and } v = 1,2,\dots,m_1(1), \quad k = 1,2,\dots,n_\gamma,\end{aligned} \tag{II.150}$$

$$\begin{aligned}\mathbf{W}^{\gamma(a2v)}_{2k} &= (-1)^{\Delta(a+1)}\sum_{v'=1}^{m_2(1)}F^{\bar{\gamma}(\beta_1\gamma_2)}_{v'v}\mathbf{Q}^{\bar{\gamma}v',a+1,1}_k, \\ &\quad a = 1,2 \text{ and } v = 1,2,\dots,m_1(2), \quad k = 1,2,\dots,n_\gamma,\end{aligned} \tag{II.151}$$

and show that CG coefficients of type III for corepresentations are linked by simple unitary transformations with convenient CG coefficients for the normal subgroup H. The only problem thereby is to compute the submatrices $F^{\bar{\gamma}(\beta_1\gamma_2)}$ and $F^{\bar{\gamma}(\beta_1\bar{\gamma}_2)}$ which are unitary but whose dimensions are not necessarily equal.

SUMMARY

This paper deals with the problem of determining CG coefficients for Kronecker products which are composed of co-unirreps of type II and III. Due to the present method the first step must be the computation of convenient CG matrices for H. Because of $\mathbb{R}^{\beta_1\gamma_2}\downarrow H = (\oplus 2)(R^{\beta_1\gamma_2} \oplus R^{\beta_1\bar{\gamma}_2})$, it is necessary to calculate two CG matrices M and N, which yield a decomposition of $R^{\beta_1\gamma_2}$, and $R^{\beta_1\bar{\gamma}_2}$, respectively, into a direct sum of their irreducible constituents. Provided this has been carried out, CG coefficients for corepresentations have to be calculated as follows.

The problem of calculating CG coefficients of type I is reduced to the task of determining the $m_{\beta_1\gamma_2;\alpha}$-dimensional unitary submatrix $F^{\alpha(\beta_1\gamma_2)}$ being contained in the special solution (II.47) for B:

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbf{1}_m & 0 & 0 & \mathbf{1}_m \\ 0 & i\mathbf{1}_m & \mathbf{1}_m & 0 \\ 0 & i\mathbf{F} & -\mathbf{F} & 0 \\ -\mathbf{F}^T & 0 & 0 & \mathbf{F}^T \end{bmatrix}.$$

The matrix elements of $\mathbf{F} = F^{\alpha(\beta_1\gamma_2)}$ have to be computed by means of Eq. (II.35). Thus, using the definitions (II.14) and (II.15), the corresponding CG coefficients are given by Eqs. (II.49)–(II.52), where the additional symmetry relation $m_{\beta_1\gamma_2;\alpha} = m_{\beta_1\bar{\gamma}_2;\alpha}$ gives rise to the simple solution for the multiplicity problem.

CG coefficients of type II are readily obtained from Eqs. (II.96)–(II.98), if one takes the definitions (II.61) and (II.62) into account. This implies that the $m_{\beta_1\gamma_2;\beta}$-dimensional unitary submatrix $F^{\beta(\beta_1\gamma_2)}$ of B, i.e.,

$$B = \begin{bmatrix} \mathbf{1}_m & 0 & 0 & 0 \\ 0 & \mathbf{1}_m & 0 & 0 \\ 0 & 0 & 0 & F^{\beta(\beta_1\gamma_2)} \\ 0 & 0 & F^{\beta(\beta_1\gamma_2)^T} & 0 \end{bmatrix},$$

has to be computed by means of Eq. (II.78). Furthermore, due to the additional symmetry relation $m_{\beta_1\gamma_2;\beta} = m_{\beta_1\bar{\gamma}_2;\beta}$, we arrive at a simple solution for the multiplicity problem.

Because of the special solution (II.130) of Eq. (II.129), i.e.,

$B = \mathbb{1}_M$ and

$$C = \begin{bmatrix} 0 & 0 & 0 & F^{\bar{\gamma}(\beta_1\gamma_2)} \\ 0 & 0 & F^{\bar{\gamma}(\beta_1\gamma_2)} & 0 \\ 0 & -F^{\bar{\gamma}(\beta_1\gamma_2)} & 0 & 0 \\ -F^{\bar{\gamma}(\beta_1\gamma_2)} & 0 & 0 & 0 \end{bmatrix},$$

the corresponding CG coefficients of type III are given by Eqs. (II.149)–(II.151), where the definitions (II.112)–(II.115) have to be used. Hence, it suffices to compute the unitary submatrices $F^{\bar{\gamma}(\beta_1\gamma_2)}$ and $F^{\bar{\gamma}(\beta_1\bar{\gamma}_2)}$ by means of Eqs. (II.139) and (II.140), respectively, whose dimensions are however not necessarily equal.

Although the present case is more complicated than the previous one (due to the inequivalence of R^{γ_2} and $R^{\bar{\gamma}_2}$), we were able to solve the multiplicity problem without reference to a special magnetic group, where especially for the first two cases additional symmetry relations for multiplicities (referring to subductions with respect to the normal subgroup H) play an essential role.

[1]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 972).

[2]R. Dirl, J. Math. Phys. **21**, xxx (1980).

[3]R. Dirl, J. Math. Phys. **21**, xxx (1980).

[4]R. Dirl, J. Math. Phys. **20**, 659 (1979).

Clebsch–Gordan coefficients for corepresentations. III ⊗ III

R. Dirl
Institut für theoretische Physik, TU Wien A-1040 Wien, Karlsplatz 13, Austria

(Received 18 December 1978; accepted for publication 25 May 1979)

By means of a general method, Clebsch–Gordan coefficients for corepresentations are traced back by simple unitary transformations to convenient Clebsch–Gordan coefficients for the normal subgroup. The considered Kronecker products are composed of corepresentations of type III only.

INTRODUCTION

The present paper concludes a series of papers which deal with the computation of CG coefficients for corepresentations. For the last case we are confronted with the most complicated situation, since the considered Kronecker products are composed of co-unirreps of type III only. We proceed in the same way as in the previous papers by assuming that convenient CG coefficients for the normal subgroup H of G are known. This assumption leads to the much easier task of determining unitary transformations which link CG coefficients for corepresentations with those of the normal subgroup.

The material is organized as follows: In Sec. I we state the problem and derive useful symmetry relations for the required multiplicities. Section II is devided into three parts due to the different types of co-unirreps. For each case we derive not only simple defining equations for the above mentioned unitary transformations but also solve them quite generally.

I. MULTIPLICITIES FOR COREPRESENTATIONS

Within the present paper we consider Kronecker products of the kind

$$\mathbb{R}^{\gamma_1\gamma_2} = \{\mathbb{R}^{\gamma_1\gamma_2}(g) = \mathbb{R}^{\gamma_1}(g) \otimes \mathbb{R}^{\gamma_2}(g) : g \in G\}. \tag{I.1}$$

$\mathbb{R}^{\gamma_1\gamma_2}$ forms a $4n_{\gamma_1}n_{\gamma_2}$-dimensional corepresentation which is in general reducible. Hence, there must exist a unitary matrix $W^{\gamma_1\gamma_2} = W$ which engenders the desired decomposition of $\mathbb{R}^{\gamma_1\gamma_2}$ into a direct sum of its irreducible consituents:

$$W^\dagger \mathbb{R}^{\gamma_1\gamma_2}(g) W^g = \sum_{\alpha\in A_{\rm I}} \oplus M_{\gamma_1\gamma_2;\alpha} \mathbb{R}^\alpha(g) \oplus \sum_{\beta\in A_{\rm II}} \oplus M_{\gamma_1\gamma_2;\beta} \mathbb{R}^\beta(g) \oplus \sum_{\gamma\in A_{\rm III}} \oplus M_{\gamma_1\gamma_2;\gamma} \mathbb{R}^\gamma(g), \quad \text{for all } g\in G. \tag{I.2}$$

The multiplicities occuring in Eq. (I.2) are given by[1]

$$M_{\gamma_1\gamma_2;\alpha} = m_{\gamma_1\gamma_2;\alpha} + m_{\gamma_1\bar\gamma_2;\alpha} + m_{\bar\gamma_1\gamma_2;\alpha} + m_{\bar\gamma_1\bar\gamma_2;\alpha}, \tag{I.3}$$

$$M_{\gamma_1\gamma_2;\beta} = \tfrac{1}{2}\{m_{\gamma_1\gamma_2;\beta} + m_{\gamma_1\bar\gamma_2;\beta} + m_{\bar\gamma_1\gamma_2;\beta} + m_{\bar\gamma_1\bar\gamma_2;\beta}\}, \tag{I.4}$$

$$M_{\gamma_1\gamma_2;\gamma} = m_{\gamma_1\gamma_2;\gamma} + m_{\gamma_1\bar\gamma_2;\gamma} + m_{\bar\gamma_1\gamma_2;\gamma} + m_{\bar\gamma_1\bar\gamma_2;\gamma}, \tag{I.5}$$

which can be simplified by means of the symmetry relations

$$m_{\gamma_1\gamma_2;\mu} = m_{\bar\gamma_1\bar\gamma_2;\bar\mu}, \quad \text{for all } \mu\in A_H, \tag{I.6}$$

$$m_{\gamma_1\bar\gamma_2;\mu} = m_{\bar\gamma_1\gamma_2;\bar\mu}, \quad \text{for all } \mu\in A_H, \tag{I.7}$$

to the following expressions:

$$M_{\gamma_1\gamma_2;\alpha} = 2\{m_{\gamma_1\gamma_2;\alpha} + m_{\gamma_1\bar\gamma_2;\alpha}\}, \tag{I.8}$$

$$M_{\gamma_1\gamma_2;\beta} = m_{\gamma_1\gamma_2;\beta} + m_{\gamma_1\bar\gamma_2;\beta}, \tag{I.9}$$

$$M_{\gamma_1\gamma_2;\gamma} = M_{\bar\gamma_1\bar\gamma_2;\bar\gamma} = M_{\gamma_1\gamma_2;\bar\gamma}. \tag{I.10}$$

Thereby, we cannot assume in general that the multiplicities occuring in Eqs. (I.8)–(I.10) are equal.

II. CG COEFFICIENTS FOR COREPRESENTATIONS

The first step of our procedure is to investigate the subduced representation

$$\mathbb{R}^{\gamma_1\gamma_2}\downarrow H = R^{\gamma_1\gamma_2} \oplus R^{\gamma_1\bar\gamma_2} \oplus R^{\bar\gamma_1\gamma_2} \oplus R^{\bar\gamma_1\bar\gamma_2}, \tag{II.1}$$

where the different Kronecker products referring to representations of H are distinguished by

$$R^{\gamma_1\gamma_2} = \{R^{\gamma_1\gamma_2}(h) = R^{\gamma_1}(h) \otimes R^{\gamma_2}(h) : h\in H\}, \tag{II.2}$$

$$R^{\gamma_1\bar\gamma_2} = \{R^{\gamma_1\bar\gamma_2}(h) = R^{\gamma_1}(h) \otimes Z^{\gamma_2\dagger} R^{\bar\gamma_2}(h) Z^{\gamma_2} : h\in H\}, \tag{II.3}$$

$$R^{\bar\gamma_1\gamma_2} = \{R^{\bar\gamma_1\gamma_2}(h) = Z^{\gamma_1\dagger} R^{\bar\gamma_1}(h) Z^{\gamma_1} \otimes R^{\gamma_2}(h) : h\in H\}, \tag{II.4}$$

$$R^{\bar\gamma_1\bar\gamma_2} = \{R^{\bar\gamma_1\bar\gamma_2}(h) = Z^{\gamma_1\dagger} R^{\bar\gamma_1}(h) Z^{\gamma_1} \otimes Z^{\gamma_2\dagger} R^{\bar\gamma_2}(h) Z^{\gamma_2} : h\in G\}. \tag{II.5}$$

Since the $n_{\gamma_1}n_{\gamma_2}$-dimensional representations (II.2)–(II.5) are in general reducible, there must exist four unitary CG matrices $K^{\gamma_1\gamma_2} = K$, $L^{\gamma_1\bar\gamma_2} = L$, $M^{\bar\gamma_1\gamma_2} = M$, and $N^{\bar\gamma_1\bar\gamma_1} = N$ which provides the desired decompositions

$$K^\dagger R^{\gamma_1\gamma_2}(h) K = \sum_{\alpha\in A_{\rm I}} \oplus m_{\gamma_1\gamma_2;\alpha} R^\alpha(h) \oplus \sum_{\beta\in A_{\rm II}} \oplus m_{\gamma_1\gamma_2;\beta} R^\beta(h) \oplus \sum_{\gamma\in A_{\rm III}} \oplus \{m_{\gamma_1\gamma_2;\gamma} R^\gamma(h) \oplus m_{\bar\gamma_1\bar\gamma_2;\gamma} Z^{\gamma\dagger} R^{\bar\gamma}(h) Z^\gamma\}, \tag{II.6}$$

$$L^\dagger R^{\gamma_1\bar\gamma_2}(h) L = \sum_{\alpha\in A_{\rm I}} \oplus m_{\gamma_1\bar\gamma_2;\alpha} R^\alpha(h) \oplus \sum_{\beta\in A_{\rm II}} \oplus m_{\gamma_1\bar\gamma_2;\beta} R^\beta(h) \oplus \sum_{\gamma\in A_{\rm III}} \oplus \{m_{\gamma_1\bar\gamma_2;\gamma} R^\gamma(h) \oplus m_{\bar\gamma_1\gamma_2;\gamma} Z^{\gamma\dagger} R^{\bar\gamma}(h) Z^\gamma\}, \tag{II.7}$$

$$M^\dagger R^{\bar\gamma_1\gamma_2}(h) M$$

 0022-2488/80/050997-09$01.00

$$= \sum_{\alpha\in A_{\rm I}} \oplus m_{\gamma_1\bar\gamma_2;\alpha} R^{\alpha}(h) \oplus \sum_{\beta\in A_{\rm II}} \oplus m_{\gamma_1\bar\gamma_2;\beta} R^{\beta}(h)$$

$$\oplus \sum_{\gamma\in A_{\rm III}} \oplus \{ m_{\bar\gamma_1\gamma_2;\gamma} R^{\gamma}(h) \oplus m_{\gamma_1\bar\gamma_2;\gamma} Z^{\gamma\dagger} R^{\bar\gamma}(h) Z^{\gamma} \}, \tag{II.8}$$

$$N^{\dagger} R^{\bar\gamma_1\bar\gamma_2}(h) N$$

$$= \sum_{\alpha\in A_{\rm I}} \oplus m_{\gamma_1\gamma_2;\alpha} R^{\alpha}(h) \oplus \sum_{\beta\in A_{\rm II}} \oplus m_{\gamma_1\gamma_2;\beta} R^{\beta}(h)$$

$$\oplus \sum_{\gamma\in A_{\rm III}} \oplus \{ m_{\bar\gamma_1\bar\gamma_2;\gamma} R^{\gamma}(h) \oplus m_{\gamma_1\gamma_2;\gamma} Z^{\gamma\dagger} R^{\bar\gamma}(h) Z^{\gamma} \},$$

$$\text{for all } h\in H, \tag{II.9}$$

respectively. Due to the present approach the CG matrices K,L,M, and N are taken for granted. Concerning Eqs. (II.6)–(II.9), we have to note that the symmetry relations (I.6) and (I.7) have already been taken into account.

A. CG coefficients of type I

The defining equations for CG coefficients of type I read as

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{W}_k^{\alpha w} = \sum_l R^{\alpha}_{lk}(h)\mathbf{W}_l^{\alpha w}, \quad \text{for all } h\in H, \tag{II.10}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U^{\alpha}_{lk}\mathbf{W}_l^{\alpha w}, \quad w = 1,2,...,M_{\gamma_1\gamma_2;\alpha},$$

$$k = 1,2,...,n_{\alpha}, \tag{II.11}$$

where th vector notation

$$\{\mathbf{W}_k^{\alpha w}\}_{ai,bj} = \{\mathbf{W}_k^{\gamma_1\gamma_2;\alpha w}\}_{ai,bj} = W^{\gamma_1\gamma_2}_{ai,bj;\alpha wk},$$

$$\alpha\in A_{\rm I}, w = 1,2,...,M_{\gamma_1\gamma_2;\alpha}, k = 1,2,...,n_{\alpha}, a = 1,2 \text{ and}$$

$$i = 1,2,...,n_{\gamma_1}, \quad b = 1,2, \text{ and } j = 1,2,...,n_{\gamma_2} \tag{II.12}$$

allows one to interpret the columns of the CG matrix W as H-adapted vectors of a $4n_{\gamma_1}n_{\gamma_2}$-dimensional Euclidean space $\mathscr{W}^{\gamma_1\gamma_2}$, which must satisfy additionally Eq. (II.11). Hence, the vectors

$$\mathbf{W}_k^{\alpha w}, \quad w = 1,2,...,M_{\gamma_1\gamma_2;\alpha}, \quad k = 1,2,...,n_{\alpha}, \tag{II.13}$$

define an orthonormal basis of

$$\mathscr{W}^{\gamma_1\gamma_2;\alpha} = \sum_i \mathbb{E}^{\alpha}_{ii}\mathscr{W}^{\gamma_1\gamma_2},$$

$$\dim \mathscr{W}^{\gamma_1\gamma_2;\alpha} = n_{\alpha}M_{\gamma_1\gamma_2;\alpha}, \tag{II.14}$$

where the units $\mathbb{E}^{\alpha}_{ij}$ can be written as

$$\mathbb{E}^{\alpha}_{ij} = E^{\gamma_1\gamma_2;\alpha}_{ij} \oplus E^{\gamma_1\bar\gamma_2;\alpha}_{ij} \oplus E^{\bar\gamma_1\gamma_2;\alpha}_{ij} \oplus E^{\bar\gamma_1\bar\gamma_2;\alpha}_{ij}, \tag{II.15}$$

$$E^{\mu_1\mu_2;\alpha}_{ij} = \frac{n_{\alpha}}{|H|}\sum_h R^{\alpha *}_{ij}(h) R^{\mu_1\mu_2}(h),$$

$$\mu_1 = \gamma_1\bar\gamma_1, \quad \mu_2 = \gamma_2,\bar\gamma_2. \tag{II.16}$$

On the other hand, the structure of $\mathbb{R}^{\gamma_1\gamma_2}\downarrow H$ suggests that one can define, by means of

$$\{\mathbf{Q}_k^{\alpha v 11}\}_{ai,bj} = \delta_{a1}\delta_{b1}\{\mathbf{K}_k^{\alpha v}\}_{ij}, \tag{II.17}$$

$$\{\mathbf{Q}_k^{\alpha v 12}\}_{ai,bj} = \delta_{a1}\delta_{b2}\{\mathbf{L}_k^{\alpha v}\}_{ij}, \tag{II.18}$$

$$\{\mathbf{Q}_k^{\alpha v 21}\}_{ai,bj} = \delta_{a2}\delta_{b1}\{\mathbf{M}_k^{\alpha v}\}_{ij}, \tag{II.19}$$

$$\{\mathbf{Q}_k^{\alpha v 22}\}_{ai,bj} = \delta_{a2}\delta_{b2}\{\mathbf{N}_k^{\alpha v}\}_{ij}, \tag{II.20}$$

a further orthonormal basis of $\mathscr{W}^{\gamma_1\gamma_2;\alpha}$, namely,

$$\mathbf{Q}_k^{\alpha vab}, \quad a,b = 1,2, \quad v = 1,2,...,m(a,b),$$

$$m(1,1) = m(2,2) = m_{\gamma_1\gamma_2;\alpha},$$

$$m(1,2) = m(2,1) = m_{\gamma_1\bar\gamma_2;\alpha}, \quad k = 1,2,...,n_{\alpha}, \tag{II.21}$$

where the vectors $\mathbf{K}_k^{\alpha v}$, $\mathbf{L}_k^{\alpha v}$, $\mathbf{M}_k^{\alpha v}$, and $\mathbf{N}_k^{\alpha v}$ are the corresponding columns of the CG matrices K, L, M, and N. Hence, these vectors transform according to

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{Q}_k^{\alpha vab} = \sum_l R^{\alpha}_{lk}(h)\mathbf{Q}_l^{\alpha vab}, \quad \text{for all } h\in H, \tag{II.22}$$

but are in general not a solution of Eq. (II.11).

Since the transformation laws (II.10) and (II.22) coincide, the elements of the bases (II.13) and (II.21) are linked by special unitary transformations

$$\mathbf{W}_k^{\alpha w} = \sum_{abv} B_{abv;w}\mathbf{Q}_k^{\alpha vab}, \tag{II.23}$$

$$\mathbf{Q}_k^{\alpha vab} = \sum_w B^{*}_{abv;w}\mathbf{W}_k^{\alpha w}, \quad k = 1,2,...,n_{\alpha}. \tag{II.24}$$

Thereby, we have to note the definition of the row index (a,b,v) of the $M_{\gamma_1\gamma_2;\alpha}$-dimensional matrix B.

In order to be able to determine unitary matrices B, so that the corresponding vectors (II.23) satisfy Eq. (II.11), we derive

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{Q}_k^{\alpha vab *} = \sum_l U^{\alpha}_{lk}\sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}_l^{\alpha v'a'b'}, \tag{II.25}$$

where

$$F_{a'b'v';abv} = \{BB^{T}\}_{abv;a'b'v'},$$

$$a,b = 1,2, \quad \text{and} \quad v = 1,2,...,m(a,b),$$

$$a',b' = 1,2, \quad \text{and} \quad v' = 1,2,...,m(a',b'). \tag{II.26}$$

This leads to

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}_k^{\alpha w *} = \sum_l U^{\alpha}_{lk}\sum_{w'}\{B^{\dagger}FB^{*}\}_{w'w}\mathbf{W}_l^{\alpha w'} \tag{II.27}$$

and

$$FB^{*} = B, \quad \text{with} \quad BB^{\dagger} = B^{\dagger}B = \mathbb{1}_M. \tag{II.28}$$

Hence, if we can find a matrix B satisfying Eq. (II.28), the corresponding CG coefficients of type I follow immediately from Eq. (II.23). Before attacking this problem, let us mention that F must be a symmetric unitary matrix, i.e.,

$$FF^{*} = \mathbb{1}_M. \tag{II.29}$$

The next step is to compute the matrix elements of F. This can be done by means of

$$\left\langle \mathbf{Q}_k^{\alpha v'a'b'}, \mathbb{R}^{\gamma_1\gamma_2}(s)\left\{\sum_l U^{\alpha}_{kl}\mathbf{Q}_l^{\alpha vab}\right\}^{*}\right\rangle = F_{a'b'v';abv}, \tag{II.30}$$

whose values must be independent of k. This assertion can be proven with the aid of

$$\mathbb{R}^{\gamma_1\gamma_2}(s)^{\dagger}\mathbb{E}^{\alpha}_{ij}\mathbb{R}^{\gamma_1\gamma_2}(s) = \sum_{kl} U^{\alpha *}_{ik}U^{\alpha}_{jl}\mathbb{E}^{\alpha *}_{kl} \tag{II.31}$$

and Eq. (II.22). Inserting Eqs. (II.17)–(II.20) into Eq. (II.30), we obtain

$$F_{11v';22v} = F^{\alpha(\gamma_1\gamma_2)}_{v'v} = \left\langle \mathbf{K}^{\alpha v'}_k, R^{\gamma_1}(s^2)\otimes R^{\gamma_2}(s^2)\left\{ \sum_l U^{\alpha}_{kl}\mathbf{N}^{\alpha v}_l\right\}^*\right\rangle, \tag{II.32}$$

$$F_{12v';21v} = F^{\alpha(\gamma_1\bar{\gamma}_2)}_{v'v} = \left\langle \mathbf{L}^{\alpha v'}_k, R^{\gamma_1}(s^2)\otimes \mathbb{1}_{\gamma_2}\left\{ \sum_l U^{\alpha}_{kl}\mathbf{M}^{\alpha v}_l\right\}^*\right\rangle, \tag{II.33}$$

$$F_{21v';12v} = F^{\alpha(\bar{\gamma}_1\gamma_2)}_{v'v} = \left\langle \mathbf{M}^{\alpha v'}_k, \mathbb{1}_{\gamma_1}\otimes R^{\gamma_2}(s^2)\left\{ \sum_l U^{\alpha}_{kl}\mathbf{L}^{\alpha v}_l\right\}^*\right\rangle, \tag{II.34}$$

$$F_{22v';11v} = F^{\alpha(\bar{\gamma}_1\bar{\gamma}_2)}_{v'v} = \left\langle \mathbf{N}^{\alpha v'}_k, \left\{ \sum_l U^{\alpha}_{kl}\mathbf{K}^{\alpha v}_l\right\}^*\right\rangle, \tag{II.35}$$

$$F_{a'b'v';abv} = 0, \quad \text{otherwise}, \tag{II.36}$$

where the scalar product in Eqs. (II.32)–(II.35) is analogously defined. Obviously, the matrix elements (II.32)–(II.35) must also be independent of k. This can be shown by means of the identities

$$\{R^{\gamma_1}(s^2)\otimes R^{\gamma_2}(s^2)\}^\dagger E^{\gamma_1\gamma_2;\alpha}_{ij}\{R^{\gamma_1}(s^2)\otimes R^{\gamma_2}(s^2)\} = \sum_{kl} U^{\alpha*}_{ik}U^{\alpha}_{jl}\{E^{\bar{\gamma}_1\bar{\gamma}_2;\alpha}_{kl}\}^*, \tag{II.37}$$

$$\{R^{\gamma_1}(s^2)\otimes \mathbb{1}_{\gamma_2}\}^\dagger E^{\gamma_1\bar{\gamma}_2;\alpha}_{ij}\{R^{\gamma_1}(s^2)\otimes \mathbb{1}_{\gamma_2}\} = \sum_{kl} U^{\alpha*}_{ik}U^{\alpha}_{jl}\{E^{\bar{\gamma}_1\gamma_2;\alpha}_{kl}\}^*, \tag{II.38}$$

$$\{\mathbb{1}_{\gamma_1}\otimes R^{\gamma_2}(s^2)\}^\dagger E^{\bar{\gamma}_1\gamma_2;\alpha}_{ij}\{\mathbb{1}_{\gamma_1}\otimes R^{\gamma_2}(s^2)\} = \sum_{kl} U^{\alpha*}_{ik}U^{\alpha}_{jl}\{E^{\gamma_1\bar{\gamma}_2;\alpha}_{kl}\}^*, \tag{II.39}$$

$$E^{\bar{\gamma}_1\bar{\gamma}_2;\alpha}_{ij} = \sum_{kl} U^{\alpha*}_{ik}U^{\alpha}_{jl}\{E^{\gamma_1\gamma_2;\alpha}_{kl}\}^*, \tag{II.40}$$

and the transformation properties of the vectors $\mathbf{K}^{\alpha v}_k$, $\mathbf{L}^{\alpha v}_k$, $\mathbf{M}^{\alpha v}_k$, and $\mathbf{N}^{\alpha v}_k$ with respect to H. Consequently, F reads as

$$F = \begin{bmatrix} 0 & 0 & 0 & F^{\alpha(\gamma_1\gamma_2)} \\ 0 & 0 & F^{\alpha(\gamma_1\bar{\gamma}_2)} & 0 \\ 0 & F^{\alpha(\bar{\gamma}_1\gamma_2)} & 0 & 0 \\ F^{\alpha(\bar{\gamma}_1\bar{\gamma}_2)} & 0 & 0 & 0 \end{bmatrix}. \tag{II.41}$$

Furthermore, since F is symmetric and unitary, it follows

$$F^{\alpha(\gamma_1\gamma_2)^T} = F^{\alpha(\bar{\gamma}_1\bar{\gamma}_2)}, \tag{II.42}$$

$$F^{\alpha(\gamma_1\bar{\gamma}_2)^T} = F^{\alpha(\bar{\gamma}_1\gamma_2)}, \tag{II.43}$$

$$F^{\alpha(\gamma_1\gamma_2)}F^{\alpha(\bar{\gamma}_1\bar{\gamma}_2)*} = \mathbb{1}_{m(1,1)}, \tag{II.44}$$

$$F^{\alpha(\gamma_1\bar{\gamma}_2)}F^{\alpha(\bar{\gamma}_1\gamma_2)*} = \mathbb{1}_{m(1,2)}. \tag{II.45}$$

Consequently, it suffices to calculate, for example, $F^{\alpha(\gamma_1\gamma_2)}$ and $F^{\alpha(\gamma_1\bar{\gamma}_2)}$ since the remaining matrices follow from Eqs. (II.42) and (II.43). Besides this the submatrices are unitary, but their dimensions will in general not be equal.

Before solving Eq. (II.28), let us assume that the corresponding columns of the CG matrices K, L, M, and N are calculable by means of the method presented in Ref. 2. Their components take the form

$$\{\mathbf{K}^{\alpha v}_k\}_{ij} = \{\mathbf{K}^{\gamma_1\gamma_2;\alpha(i_v j_v)}_k\}_{ij} = \|\mathbf{B}^{\gamma_1\gamma_2;\alpha(i_v j_v)}_{a_0}\|^{-1}\frac{n_\alpha}{|H|}\sum_h R^{\gamma_1}_{ii_v}(h)R^{\gamma_2}_{jj_v}(h)R^{\alpha*}_{ka_0}(h),$$

$$v = 1,2,\ldots,m(1,1), \quad k = 1,2,\ldots,n_\alpha, \tag{II.46}$$

$$\{\mathbf{L}^{\alpha v}_k\}_{ij} = \{\mathbf{L}^{\gamma_1\bar{\gamma}_2;\alpha(i_v j_v)}_k\}_{ij} = \|\mathbf{B}^{\gamma_1\bar{\gamma}_2;\alpha(i_v j_v)}_{a_0}\|^{-1}\frac{n_\alpha}{|H|}\sum_h R^{\gamma_1}_{ii_v}(h) \times \{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{jj_v}R^{\alpha*}_{ka_0}(h),$$

$$v = 1,2,..,m(1,2), \quad k = 1,2,\ldots,n_\alpha, \tag{II.47}$$

$$\{\mathbf{M}^{\alpha v}_k\}_{ij} = \{\mathbf{M}^{\bar{\gamma}_1\gamma_2;\alpha(i_v j_v)}_k\}_{ij} \times \|\mathbf{B}^{\bar{\gamma}_1\gamma_2;\alpha(i_v j_v)}_{a_0}\|^{-1}\frac{n_\alpha}{|H|}\sum_h \{Z^{\gamma_1\dagger}R^{\bar{\gamma}_1}(h)Z^{\gamma_1}\}_{ii_v} \times R^{\gamma_2}_{jj_v}(h)R^{\alpha*}_{ka_0}(h),$$

$$v = 1,2,\ldots,m(1,2), \quad k = 1,2,\ldots,n_\alpha, \tag{II.48}$$

$$\{\mathbf{N}^{\alpha v}_k\}_{ij} = \{\mathbf{N}^{\bar{\gamma}_1\bar{\gamma}_2;\alpha(i_v j_v)}_k\}_{ij} = \|\mathbf{B}^{\bar{\gamma}_1\bar{\gamma}_2;\alpha(i_v j_v)}_{a_0}\|^{-1}\frac{n_\alpha}{|H|}\sum_h \{Z^{\gamma_1\dagger}R^{\bar{\gamma}_1}(h)Z^{\gamma_1}\}_{ii_v} \times \{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{jj_v}R^{\alpha*}_{ka_0}(h),$$

$$v = 1,2,\ldots,m(1,1), \quad k = 1,2,\ldots,n_\alpha, \tag{II.49}$$

where however the index sets $\{(i_v, j_v)\}$ occuring in Eqs. (II.46) and (II.49) or (II.47) and (II.48) are in general different. Inserting these special values into Eqs. (II.32) and (II.33), we obtain

$$F^{\alpha(\gamma_1\gamma_2)}_{v'v} = \|\mathbf{B}^{\gamma_1\gamma_2;\alpha(i_{v'} j_{v'})}_{a_0}\|^{-1}\|\mathbf{B}^{\bar{\gamma}_1\bar{\gamma}_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \times \frac{n_\alpha}{|H|}\sum_h R^{\gamma_1}_{i_{v'}i_v}(hs^2)R^{\gamma_2}_{j_{v'}j_v}(hs^2)\mathbb{R}^{\alpha*}_{a_0a_0}(hs), \tag{II.50}$$

$$F^{\alpha(\gamma_1\bar{\gamma}_2)}_{v'v} = \|\mathbf{B}^{\gamma_1\bar{\gamma}_2;\alpha(i_{v'} j_{v'})}_{a_0}\|^{-1}\|\mathbf{B}^{\bar{\gamma}_1\gamma_2;\alpha(i_v j_v)}_{a_0}\|^{-1} \times \frac{n_\alpha}{|H|}\sum_h R^{\gamma_1}_{i_{v'}i_v}(hs^2) \times \{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{j_{v'}j_v}\mathbb{R}^{\alpha*}_{a_0a_0}(hs), \tag{II.51}$$

whereas the remaining matrix elements follow from Eqs. (II.42) and (II.43), since these relations are always valid.

Besides this we are now in the position to determine unitary matrices B which satisfy Eq. (II.28). For this purpose we make a generalized ansatz for B:

$$B = \begin{bmatrix} \mathbf{A} & 0 & 0 & \mathbf{F}\mathbf{D}^* \\ 0 & \mathbf{B} & \mathbf{G}\mathbf{C}^* & 0 \\ 0 & \mathbf{G}^T\mathbf{B}^* & \mathbf{C} & 0 \\ \mathbf{F}^T\mathbf{A}^* & 0 & 0 & \mathbf{D} \end{bmatrix}, \tag{II.52}$$

where $\mathbf{F} = F^{\alpha(\gamma_1\gamma_2)}$ and $\mathbf{G} = F^{\alpha(\gamma_1\bar{\gamma}_2)}$. The symbols $\mathbf{A}$ and $\mathbf{D}$, and $\mathbf{B}$ and $\mathbf{C}$, denote $m(1,1)$-dimensional and $m(1,2)$-dimensional matrices, respectively which shall be proportional by numerical factors to unitary ones, but otherwise arbitrary. Now it is easy to verify that any matrix B of the type (II.52) satisfies Eq. (II.28). Since B is required to be unitary, we obtain

$$\mathbf{A}\mathbf{A}^\dagger + \mathbf{D}\mathbf{D}^\dagger = \mathbb{1}_{m(1,1)} \quad \text{and} \quad \mathbf{B}\mathbf{B}^\dagger + \mathbf{C}\mathbf{C}^\dagger = \mathbb{1}_{m(1,2)}, \tag{II.53}$$

$$\mathbf{F}^T(\mathbf{A}\mathbf{A}^T)^* + (\mathbf{D}\mathbf{D}^T)\mathbf{F}^\dagger = 0, \tag{II.54}$$

$$\mathbf{G}^T(\mathbf{B}\mathbf{B}^T)^* + (\mathbf{C}\mathbf{C}^T)\mathbf{G}^\dagger = 0, \tag{II.55}$$

as restricting conditions for **A** and **D** and **B** and **C** respectively. If we choose

$$\mathbf{A} = \frac{i}{\sqrt{2}}\mathbb{1}_{m(1,1)} \text{ and } \mathbf{B} = \frac{i}{\sqrt{2}}\mathbb{1}_{m(1,2)}, \tag{II.56}$$

we obtain

$$\mathbf{D} = \frac{1}{\sqrt{2}}\mathbf{F}^T \text{ and } \mathbf{C} = \frac{1}{\sqrt{2}}\mathbf{G}^T, \tag{II.57}$$

respectively, in matrix notation

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbb{1}_{m(1,1)} & 0 & 0 & \mathbb{1}_{m(1,1)} \\ 0 & i\mathbb{1}_{m(1,2)} & \mathbb{1}_{m(1,2)} & 0 \\ 0 & -i\mathbf{G}^T & \mathbf{G}^T & 0 \\ -i\mathbf{F}^T & 0 & 0 & \mathbf{F}^T \end{bmatrix}. \tag{II.58}$$

In this connection we remark that Eq. (II.58) satisfies Eq. (II.26). Besides this the special solution (II.58) allows one to identify the multiplicity index w with the triplet (a,b,v), i.e.,

$$w = (a,b,v), \quad a,b = 1,2 \text{ and } v = 1,2,...,m(a,b). \tag{II.59}$$

The corresponding CG coefficients of type I are readily obtained from Eq. (II.23) by inserting the matrix elements of Eq. (II.58):

$$\mathbf{W}_k^{\alpha(11v)} = \frac{i}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v11} - \sum_{v'} F_{vv'}^{\alpha(\gamma_1\gamma_2)}\mathbf{Q}_k^{\alpha v'22}\Big\},$$
$$v = 1,2,..,m(1,1), \tag{II.60}$$

$$\mathbf{W}_k^{\alpha(12v)} = \frac{i}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v12} - \sum_{v'} F_{vv'}^{\alpha(\gamma_1\bar{\gamma}_2)}\mathbf{Q}_k^{\alpha v'21}\Big\},$$
$$v = 1,2,...,m(1,2), \tag{II.61}$$

$$\mathbf{W}_k^{\alpha(21v)} = \frac{1}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v12} + \sum_{v'} F_{vv'}^{\alpha(\gamma_1\bar{\gamma}_2)}\mathbf{Q}_k^{\alpha v'21}\Big\},$$
$$v = 1,2,...,m(1,2), \tag{II.62}$$

$$\mathbf{W}_k^{\alpha(22v)} = \frac{1}{\sqrt{2}}\Big\{\mathbf{Q}_k^{\alpha v11} + \sum_{v'} F_{vv'}^{\alpha(\gamma_1\gamma_2)}\mathbf{Q}_k^{\alpha v'22}\Big\},$$
$$v = 1,2,...,m(1,1). \tag{II.63}$$

These formulas show that CG coefficients of type I can be expressed by simple unitary transformations in terms of convenient CG coefficients for the subgroup H, where it suffices to compute the unitary submatrices $F^{\alpha(\gamma_1\gamma_2)}$ and $F^{\alpha(\gamma_1\bar{\gamma}_2)}$.

B. CG coefficients of type II

The defining equations for CG coefficients of type II take for this kind of Kronecker product the form

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{W}_{dk}^{\beta w} = \sum_l R_{lk}^{\beta}(h)\mathbf{W}_{dl}^{\beta w}, \quad \text{for all } h\in H, \tag{II.64}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}_{dk}^{\beta w*} = (-1)^{\Delta(d+1)}\sum_l U_{lk}^{\beta}\mathbf{W}_{d+1,l}^{\beta w},$$

$$w = 1,2,...,M_{\gamma_1\gamma_2;\beta}, \quad d = 1,2, \text{ and } k = 1,2,..., n_\beta, \tag{II.65}$$

where the vector notation has to be understood as

$$\{\mathbf{W}_{dk}^{\beta w}\}_{ai,bj} = \{\mathbf{W}_{dk}^{\gamma_1\gamma_2;\beta w}\}_{ai,bj} = W_{ai,bj;\beta wdk}^{\gamma_1\gamma_2},$$
$$\beta\in A_{11}, \quad w = 1,2...,M_{\gamma_1\gamma_2;\beta}, \quad d = 1,2, \text{ and } k = 1,2,..n_\beta,$$
$$a = 1,2, \text{ and } i = 1,2,...,n_{\gamma_1}, \quad b = 1,2, \text{ and } j = 1,2,...,n_{\gamma_2}. \tag{II.66}$$

The following vectors:

$$\mathbf{W}_{dk}^{\beta w}, \quad w = 1,2,...,M_{\gamma_1\gamma_2;\beta}, \quad d = 1,2, \text{ and } k = 1,2,...,n_\beta, \tag{II.67}$$

representing columns of the unitary CG matrix W, form an orthonormal basis of

$$\mathscr{W}^{\gamma_1\gamma_2;\beta} = \sum_i \mathbb{E}_{ii}^{\beta}\mathscr{W}^{\gamma_1\gamma_2}, \quad \dim\mathscr{W}^{\gamma_1\gamma_2;\beta} = 2n_\beta M_{\gamma_1\gamma_2;\beta}, \tag{II.68}$$

where the units $\mathbb{E}_{ij}^{\beta}$ can be written as

$$\mathbb{E}_{ij}^{\beta} = E_{ij}^{\gamma_1\gamma_2;\beta}\oplus E_{ij}^{\gamma_1\bar{\gamma}_2;\beta}\oplus E_{ij}^{\bar{\gamma}_1\gamma_2;\beta}\oplus E_{ij}^{\bar{\gamma}_1\bar{\gamma}_2;\beta}, \tag{II.69}$$

$$E_{ij}^{\mu_1\mu_2;\beta} = \frac{n_\beta}{|H|}\sum_h R_{ij}^{\beta*}(h)R^{\mu_1\mu_2}(h),$$
$$\mu_1 = \gamma_1,\bar{\gamma}_1, \quad \mu_2 = \gamma_2,\bar{\gamma}_2. \tag{II.70}$$

Like in the previous case we introduce, by means of

$$\{\mathbf{Q}_k^{\beta v11}\}_{ai,bj} = \delta_{a1}\delta_{b1}\{\mathbf{K}_k^{\beta v}\}_{ij}, \tag{II.71}$$
$$\{\mathbf{Q}_k^{\beta v12}\}_{ai,bj} = \delta_{a1}\delta_{b2}\{\mathbf{L}_k^{\beta v}\}_{ij}, \tag{II.72}$$
$$\{\mathbf{Q}_k^{\beta v21}\}_{ai,bj} = \delta_{a2}\delta_{b1}\{\mathbf{M}_k^{\beta v}\}_{ij}, \tag{II.73}$$
$$\{\mathbf{Q}_k^{\beta v22}\}_{ai,bj} = \delta_{a2}\delta_{b2}\{\mathbf{N}_k^{\beta v}\}_{ij}, \tag{II.74}$$

a further orthonormal basis of $\mathscr{W}^{\gamma_1\gamma_2;\beta}$, namely,

$$\mathbf{Q}_k^{\beta vab}, \quad a,b = 1,2, \; v = 1,2,...,m(a,b),$$
$$m(1,1) = m(2,2) = m_{\gamma_1\gamma_2;\beta},$$
$$m(1,2) = m(2,1) = m_{\gamma_1\bar{\gamma}_2;\beta}, \quad k = 1,2,...,n_\beta, \tag{II.75}$$

which is especially suited to simplifying the following considerations, since they transform with respect to H according to

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{Q}_k^{\beta vab} = \sum_l R_{lk}^{\beta}(h)\mathbf{Q}_l^{\beta vab}, \quad \text{for all } h\in H. \tag{II.76}$$

By similar arguments as before we define unitary transformations

$$\mathbf{W}_{dk}^{\beta w} = \sum_{abv} B_{abv;dw}\mathbf{Q}_k^{\beta vab}, \tag{II.77}$$

$$\mathbf{Q}_k^{\beta vab} = \sum_{dw} B_{abv;dw}^{*}\mathbf{W}_{dk}^{\beta w}, \quad k = 1,2,...,n_\beta, \tag{II.78}$$

and try to determine them in such a way that the corresponding vectors (II.77) satisfy Eq. (II.65), since we cannot expect that the vectors (II.75) are already a solution of Eq. (II.65).

For this reason we derive

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{Q}_k^{\beta vab*} = \sum_l U_{lk}^{\beta}\sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}_l^{\beta v'a'b'}, \tag{II.79}$$

where

$$G_{dw;d'w'} = (-1)^{\Delta(d+1)}\delta_{d',d+1}\delta_{ww'},$$
$$d,d' = 1,2, \quad w,w' = 1,2,...,M_{\gamma_1\gamma_2;\beta}, \tag{II.80}$$

$$F_{a'b'v';abv} = \{BGB^T\}_{abv;a'b'v'},$$
$$a,b = 1,2, \text{ and } v = 1,2,...,m(a,b),$$
$$a',b' = 1,2, \text{ and } v' = 1,2,...,m(a',b'). \tag{II.81}$$

Utilizing these relations, we obtain

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}_{dk}^{\beta w*} = \sum_l U_{lk}^{\beta}\sum_{d'w'}\{B^\dagger FB^*\}_{d'w';dw}\mathbf{W}_{d'l}^{\beta w'}, \tag{II.82}$$

which have as a consequence

$$BG^T = FB^*. \tag{II.83}$$

Hence, any unitary $2M_{\gamma_1\gamma_2;\beta}$-dimensional matrix B satisfying Eq. (II.83) solve our problem. Before solving this problem, let us mention that F is antisymmetric and unitary, i.e.,

$$FF^* = -\mathbb{1}_{2M}, \tag{II.84}$$

and which is a useful relation for the following considerations.

As next step we compute the matrix elements of F by means of

$$\left\langle \mathbf{Q}_k^{\beta v'a'b'}, \mathbb{R}^{\gamma_1\gamma_2}(s)\left\{\sum_l U^\beta_{kl}\mathbf{Q}_l^{\beta vab}\right\}^*\right\rangle = F_{a'b'v';abv}. \tag{II.85}$$

These matrix elements are independent of k, which follows from

$$\mathbb{R}^{\gamma_1\gamma_2}(s)^\dagger \mathbb{E}^\beta_{ij}\mathbb{R}^{\gamma_1\gamma_2}(s) = \sum_{kl} U^{\beta*}_{ik}U^\beta_{jl}\mathbb{E}^{\beta*}_{kl} \tag{II.86}$$

and Eq. (II.76). Inserting Eqs. (II.71)–(II.74) into Eq. (II.85), we obtain

$$\begin{aligned} F_{11v';22v} &= F^{\beta(\gamma_1\gamma_2)}_{v'v} \\ &= \left\langle \mathbf{K}_k^{\beta v'}, R^{\gamma_1}(s^2)\otimes R^{\gamma_2}(s^2)\left\{\sum_l U^\beta_{kl}\mathbf{N}_l^{\beta v}\right\}^*\right\rangle, \end{aligned} \tag{II.87}$$

$$\begin{aligned} F_{12v';21v} &= F^{\beta(\gamma_1\bar\gamma_2)}_{v'v} \\ &= \left\langle \mathbf{L}_k^{\beta v'}, R^{\gamma_1}(s^2)\otimes \mathbb{1}_{\gamma_2}\left\{\sum_l U^\beta_{kl}\mathbf{M}_l^{\beta v}\right\}^*\right\rangle, \end{aligned} \tag{II.88}$$

$$\begin{aligned} F_{21v';12v} &= F^{\beta(\bar\gamma_1\gamma_2)}_{v'v} \\ &= \left\langle \mathbf{M}_k^{\beta v'}, \mathbb{1}_{\gamma_1}\otimes R^{\gamma_2}(s^2)\left\{\sum_l U^\beta_{kl}\mathbf{L}_l^{\beta v}\right\}^*\right\rangle, \end{aligned} \tag{II.89}$$

$$F_{22v';11v} = F^{\beta(\bar\gamma_1\bar\gamma_2)}_{v'v} = \left\langle \mathbf{N}_k^{\beta v'}, \left\{\sum_l U^\beta_{kl}\mathbf{K}_l^{\beta v}\right\}^*\right\rangle, \tag{II.90}$$

$$F_{a'b'v';abv} = 0, \quad \text{otherwise}, \tag{II.91}$$

where the relations

$$\begin{aligned} &\{R^{\gamma_1}(s^2)\otimes R^{\gamma_2}(s^2)\}^\dagger E^{\gamma_1\gamma_2;\beta}_{ij}\{R^{\gamma_1}(s^2)\otimes R^{\gamma_2}(s^2)\} \\ &\quad= \sum_{kl} U^{\beta*}_{ik}U^\beta_{jl}\{E^{\bar\gamma_1\bar\gamma_2;\beta}_{kl}\}^*, \end{aligned} \tag{II.92}$$

$$\begin{aligned} &\{R^{\gamma_1}(s^2)\otimes \mathbb{1}_{\gamma_2}\}^\dagger E^{\gamma_1\bar\gamma_2;\beta}_{ij}\{R^{\gamma_1}(s^2)\otimes \mathbb{1}_{\gamma_2}\} \\ &\quad= \sum_{kl} U^{\beta*}_{ik}U^\beta_{jl}\{E^{\bar\gamma_1\gamma_2;\beta}_{kl}\}^*, \end{aligned} \tag{II.93}$$

$$\begin{aligned} &\{\mathbb{1}_{\gamma_1}\otimes R^{\gamma_2}(s^2)\}^\dagger E^{\bar\gamma_1\gamma_2;\beta}_{ij}\{\mathbb{1}_{\gamma_1}\otimes R^{\gamma_2}(s^2)\} \\ &\quad= \sum_{kl} U^{\beta*}_{ik}U^\beta_{jl}\{E^{\gamma_1\bar\gamma_2;\beta}_{kl}\}^*, \end{aligned} \tag{II.94}$$

$$E^{\bar\gamma_1\bar\gamma_2;\beta}_{ij} = \sum_{kl} U^{\beta*}_{ik}U^\beta_{jl}\{E^{\gamma_1\gamma_2;\beta}_{kl}\}^*, \tag{II.95}$$

have to be used together with the transformations properties of the vectors $\mathbf{K}_k^{\beta v}$, $\mathbf{L}_k^{\beta v}$, $\mathbf{M}_k^{\beta v}$, and $\mathbf{N}_k^{\beta v}$ in order to show that the matrix elements (II.93)–(II.96) are also independent of k. Hence, F takes the special form

$$F = \begin{bmatrix} 0 & 0 & 0 & F^{\beta(\gamma_1\gamma_2)} \\ 0 & 0 & F^{\beta(\gamma_1\bar\gamma_2)} & 0 \\ 0 & F^{\beta(\bar\gamma_1\gamma_2)} & 0 & 0 \\ F^{\beta(\bar\gamma_1\bar\gamma_2)} & 0 & 0 & 0 \end{bmatrix}. \tag{II.96}$$

Therefrom follow for the submatrices $F^{\beta(\dots)}$ the conditions

$$F^{\beta(\gamma_1\gamma_2)^T} = -F^{\beta(\bar\gamma_1\bar\gamma_2)}, \tag{II.97}$$

$$F^{\beta(\gamma_1\bar\gamma_2)^T} = -F^{\beta(\bar\gamma_1\gamma_2)}, \tag{II.98}$$

$$F^{\beta(\gamma_1\gamma_2)}F^{\beta(\bar\gamma_1\bar\gamma_2)*} = -\mathbb{1}_{m(1,1)}, \tag{II.99}$$

$$F^{\beta(\gamma_1\bar\gamma_2)}F^{\beta(\bar\gamma_1\gamma_2)*} = -\mathbb{1}_{m(1,2)}, \tag{II.100}$$

which show that it suffices to compute the matrices $F^{\beta(\gamma_1\gamma_2)}$ and $F^{\beta(\gamma_1\bar\gamma_2)}$ and that both submatrices are also unitary.

Before solving Eq. (II.83), we assume that the corresponding columns of the unitary CG matrices K, L, M, and N can be computed with the method given in Ref. 2, i.e.,

$$\begin{aligned} \{\mathbf{K}_k^{\beta v}\}_{ij} &= \{\mathbf{K}_k^{\gamma_1\gamma_2;\beta(i_vj_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\gamma_1\gamma_2;\beta(i_vj_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h R^{\gamma_1}_{ii_v}(h) \\ &\quad\times R^{\gamma_2}_{jj_v}(h)R^{\beta*}_{ka_0}(h), \\ &\quad v = 1,2,\dots,m(1,1), \quad k = 1,2,\dots,n_\beta, \end{aligned} \tag{II.101}$$

$$\begin{aligned} \{\mathbf{L}_k^{\beta v}\}_{ij} &= \{\mathbf{L}_k^{\gamma_1\bar\gamma_2;\beta(i_vj_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\gamma_1\bar\gamma_2;\beta(i_vj_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h R^{\gamma_1}_{ii_v}(h) \\ &\quad\times\{Z^{\gamma_2\dagger}R^{\bar\gamma_2}(h)Z^{\gamma_2}\}_{jj_v}R^{\beta*}_{ka_0}(h), \\ &\quad v = 1,2,\dots,m(1,2), \quad k = 1,2,\dots,n_\beta, \end{aligned} \tag{II.102}$$

$$\begin{aligned} \{\mathbf{M}_k^{\beta v}\}_{ij} &= \{\mathbf{M}_k^{\bar\gamma_1\gamma_2;\beta(i_vj_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\bar\gamma_1\gamma_2;\beta(i_vj_v)}\|^{-1}\frac{n_\beta}{|H|} \\ &\quad\times\sum_h\{Z^{\gamma_1\dagger}R^{\bar\gamma_1}(h)Z^{\gamma_1}\}_{ii_v}R^{\gamma_2}_{jj_v}(h)R^{\beta*}_{ka_0}(h), \\ &\quad v = 1,2,\dots,m(1,2), \quad k = 1,2,\dots,n_\beta, \end{aligned} \tag{II.103}$$

$$\begin{aligned} \{\mathbf{N}_k^{\beta v}\}_{ij} &= \{\mathbf{N}_k^{\bar\gamma_1\bar\gamma_2;\beta(i_vj_v)}\}_{ij} \\ &= \|\mathbf{B}_{a_0}^{\bar\gamma_1\bar\gamma_2;\beta(i_vj_v)}\|^{-1}\frac{n_\beta}{|H|}\sum_h\{Z^{\gamma_1\dagger}R^{\bar\gamma_1}(h)Z^{\gamma_1}\}_{ii_v} \\ &\quad\times\{Z^{\gamma_2\dagger}R^{\bar\gamma_2}(h)Z^{\gamma_2}\}_{jj_v}R^{\beta*}_{ka_0}(h), \\ &\quad v = 1,2,\dots,m(1,1), \quad k = 1,2,\dots,n_\beta. \end{aligned} \tag{II.104}$$

Concerning the index sets $\{(i_v,j_v)\}$ which occur in Eqs. (II.101) and (II.104) and Eqs. (II.102) and (II.103) we recall that they are not in general equal. Because of Eqs. (II.97) and (II.98), it suffices to calculate Eqs. (II.87) and (II.88):

$$\begin{aligned} F^{\beta(\gamma_1\gamma_2)}_{v'v} &= \|\mathbf{B}_{a_0}^{\gamma_1\gamma_2;\beta(i_{v'}j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\bar\gamma_1\bar\gamma_2;\beta(i_vj_v)}\|^{-1} \\ &\quad\times\frac{n_\beta}{|H|}\sum_h R^{\gamma_1}_{i_{v'}i_v}(hs^2)R^{\gamma_2}_{j_{v'}j_v}(hs^2) \\ &\quad\{R^\beta(h)U^\beta\}^*_{a_0a_0}, \end{aligned} \tag{II.105}$$

$$\begin{aligned} F^{\beta(\gamma_1\bar\gamma_2)}_{v'v} &= \|\mathbf{B}_{a_0}^{\gamma_1\bar\gamma_2;\beta(i_{v'}j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\bar\gamma_1\gamma_2;\beta(i_vj_v)}\|^{-1} \\ &\quad\times\frac{n_\beta}{|H|}\sum_h R^{\gamma_1}_{i_{v'}i_v}(hs^2) \end{aligned}$$

$$\times\{Z^{\gamma_2\dagger}R^{\bar{\gamma}_2}(h)Z^{\gamma_2}\}_{j_{v'}j_v}\{R^{\beta}(h)U^{\beta}\}^*_{a_0a}, \tag{II.106}$$

respectively.

In order to determine unitary matrices B satisfying Eq. (II.89), we proceed in the same way as in the foregoing papers. We define

$$\{\mathbf{B}^{d,w}\}_{abv} = B_{abv;dw},$$
$$d = 1,2, \text{ and } w = 1,2,...,M_{\gamma_1\gamma_2;\beta},$$
$$a,b = 1,2, \text{ and } v = 1,2,...,m(a,b), \tag{II.107}$$

which allows one to rewrite Eq. (II.89) as

$$F\mathbf{B}^{d,w*} = (-1)^{\Delta(d+1)}\mathbf{B}^{d+1,w},$$
$$d = 1,2, \text{ and } w = 1,2,...,M_{\gamma_1\gamma_2;\beta}. \tag{II.108}$$

By similar arguments we have to fix, for example, the vectors $\mathbf{B}^{1,w}$, $w = 1,2,...,M_{\gamma_1\gamma_2;\beta}$ in such a way that the corresponding matrix B is unitary. This can be achieved, for example, by means of

$$\{\mathbf{B}^{1,w=(bv)}\}_{a'b'v'} = \delta_{a'1}\delta_{b'b}\delta_{v'v}, \tag{II.109}$$

which implies

$$\{\mathbf{B}^{2,(bv)}\}_{a'b'v'} = -F_{a'b'v';1bv} \tag{II.110}$$

and shows that we can identify the multiplicity index w with the pair (b,v), i.e.,

$$w = (b,v), \quad b = 1,2, \text{ and } v = 1,2,...,m(1,b), \tag{II.111}$$

since the corresponding matrix B is indeed unitary. Equations (II.110) written down in more detail leads to

$$\{\mathbf{B}^{2,(1v)}\}_{abv'} = \delta_{a2}\delta_{b2}F^{\beta(\gamma_1\gamma_2)}_{vv'}, \tag{II.112}$$

$$\{\mathbf{B}^{2,(2v)}\}_{abv'} = \delta_{a2}\delta_{b1}F^{\beta(\gamma_1\bar{\gamma}_2)}_{vv'}, \tag{II.113}$$

so that the corresponding matrix B reads as

$$B = \begin{bmatrix} \mathbb{1}_{m(1,1)} & 0 & 0 & 0 \\ 0 & \mathbb{1}_{m(1,2)} & 0 & 0 \\ 0 & 0 & 0 & F^{\beta(\gamma_1\bar{\gamma}_2)^T} \\ 0 & 0 & F^{\beta(\gamma_1\gamma_2)^T} & 0 \end{bmatrix}, \tag{II.114}$$

where the different definition of row and column indices of B should not be confused, if $m(1,1)\neq m(1,2)$. Besides this the corresponding CG coefficients of type II are given by

$$\mathbf{W}^{\beta(bv)}_{1k} = \mathbf{Q}^{\beta v1b}_k, \quad b = 1,2, \text{ and } v = 1,2,...,m(1,b), \tag{II.115}$$

$$\mathbf{W}^{\beta(1v)}_{2k} = \sum_{v'} F^{\beta(\gamma_1\gamma_2)}_{vv'}\mathbf{Q}^{\beta v'22}_k, \quad v = 1,2,...,m(1,1), \tag{II.116}$$

$$\mathbf{W}^{\beta(2v)}_{2k} = \sum_{v'} F^{\beta(\gamma_1\bar{\gamma}_2)}_{vv'}\mathbf{Q}^{\beta v'21}_k, \quad v = 1,2,...,m(1,2). \tag{II.117}$$

To conclude this part we realize that CG coefficients of type II can also for this case be linked by simple unitary transformations with convenient ones for the normal subgroup. Thereby, the only problem is to compute the submatrices $F^{\beta(\gamma_1\gamma_2)}$ and $F^{\beta(\gamma_1\bar{\gamma}_2)}$, whose dimensions however are not in general equal.

C. CG coefficients of type III

As already known the defining equations for CG coefficients of type III are of the form

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{W}^{\gamma w}_{1k} = \sum_l R^{\gamma}_{lk}(h)\mathbf{W}^{\gamma w}_{1l}, \tag{II.118}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{W}^{\gamma w}_{2k} = \sum_l \{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}_{lk}\mathbf{W}^{\gamma w}_{2l},$$

$$\text{for all } h\in H, \tag{II.119}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}^{\gamma w*}_{1k} = \mathbf{W}^{\gamma w}_{2k}, \tag{II.120}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}^{\gamma w*}_{2k} = \sum_l R^{\gamma}_{lk}(s^2)\mathbf{W}^{\gamma w}_{1l},$$

$$w = 1,2,...,M_{\gamma_1\gamma_2;\gamma}, \quad k = 1,2,...,n_{\gamma}, \tag{II.121}$$

where the components of the columns of the unitary CG matrix W are denoted by

$$\{\mathbf{W}^{\gamma w}_{dk}\}_{ai,bj} = \{\mathbf{W}^{\gamma_1\gamma_2;\gamma w}_{dk}\}_{ai,bj} = W^{\gamma_1\gamma_2}_{ai,bj;\gamma wdk},$$

$$\gamma\in A_{\text{III}}, \quad w = 1,2,...,M_{\gamma_1\gamma_2;\gamma}, \quad d = 1,2, \text{ and } k = 1,2,...,n_{\gamma},$$
$$a = 1,2, \text{ and } i = 1,2,...,n_{\gamma_1},$$
$$b = 1,2, \text{ and } j = 1,2,...,n_{\gamma_2}. \tag{II.122}$$

Obviously, the vectors

$$\mathbf{W}^{\gamma w}_{dk}, \quad w = 1,2,...,M_{\gamma_1\gamma_2;\gamma},$$
$$d = 1,2, \text{ and } k = 1,2,...,n_{\gamma}, \tag{II.123}$$

form an orthonormal basis of

$$\mathscr{W}^{\gamma_1\gamma_2;\gamma} = \sum_i \{\mathbb{E}^{\gamma}_{ii} + \mathbb{E}^{\bar{\gamma}}_{ii}\}\mathscr{W}^{\gamma_1\gamma_2},$$

$$\dim\mathscr{W}^{\gamma_1\gamma_2;\gamma} = 2n_{\gamma}M_{\gamma_1\gamma_2;\gamma}, \tag{II.124}$$

where the units $\mathbb{E}^{\gamma}_{ij}$ and $\mathbb{E}^{\bar{\gamma}}_{ij}$ can be written as

$$\mathbb{E}^{\gamma}_{ij} = E^{\gamma_1\gamma_2;\gamma}_{ij}\oplus E^{\gamma_1\bar{\gamma}_2;\gamma}_{ij}\oplus E^{\bar{\gamma}_1\gamma_2;\gamma}_{ij}\oplus E^{\bar{\gamma}_1\bar{\gamma}_2;\gamma}_{ij}, \tag{II.125}$$

$$E^{\mu_1\mu_2;\gamma}_{ij} = \frac{n_{\gamma}}{|H|}\sum_h R^{\gamma*}_{ij}(h)R^{\mu_1\mu_2}(h)$$
$$\mu_1 = \gamma_1,\bar{\gamma}_1, \quad \mu_2 = \gamma_2,\bar{\gamma}_2, \tag{II.126}$$

and

$$\mathbb{E}^{\bar{\gamma}}_{ij} = E^{\gamma_1\gamma_2;\bar{\gamma}}_{ij}\oplus E^{\gamma_1\bar{\gamma}_2;\bar{\gamma}}_{ij}\oplus E^{\bar{\gamma}_1\gamma_2;\bar{\gamma}}_{ij}\oplus E^{\bar{\gamma}_1\bar{\gamma}_2;\bar{\gamma}}_{ij}, \tag{II.127}$$

$$E^{\mu_1\mu_2;\bar{\gamma}}_{ij} = \frac{n_{\gamma}}{|H|}\sum_h \{Z^{\gamma\dagger}R^{\bar{\gamma}}(h)Z^{\gamma}\}^*_{ij}R^{\mu_1\mu_2}(h),$$
$$\mu_1 = \gamma_1,\bar{\gamma}_1, \quad \mu_2 = \gamma_2,\bar{\gamma}_2, \tag{II.128}$$

respectively. The structure of $\mathbb{R}^{\gamma_1\gamma_2}\downarrow H$ suggests that one can introduce, by means of

$$\{\mathbf{Q}^{\gamma v11}_k\}_{ai,bj} = \delta_{a1}\delta_{b1}\{\mathbf{K}^{\gamma v}_k\}_{ij}, \tag{II.129}$$

$$\{\mathbf{Q}^{\gamma v12}_k\}_{ai,bj} = \delta_{a1}\delta_{b2}\{\mathbf{L}^{\gamma v}_k\}_{ij}, \tag{II.130}$$

$$\{\mathbf{Q}^{\gamma v21}_k\}_{ai,bj} = \delta_{a2}\delta_{b1}\{\mathbf{M}^{\gamma v}_k\}_{ij}, \tag{II.131}$$

$$\{\mathbf{Q}^{\gamma v22}_k\}_{ai,bj} = \delta_{a2}\delta_{b2}\{\mathbf{N}^{\gamma v}_k\}_{ij}, \tag{II.132}$$

and

$$\{\mathbf{Q}^{\bar{\gamma}v11}_k\}_{ai,bj} = \delta_{a1}\delta_{b1}\{\mathbf{K}^{\bar{\gamma}v}_k\}_{ij}, \tag{II.133}$$

$$\{\mathbf{Q}^{\bar{\gamma}v12}_k\}_{ai,bj} = \delta_{a1}\delta_{b2}\{\mathbf{L}^{\bar{\gamma}v}_k\}_{ij}, \tag{II.134}$$

$$\{\mathbf{Q}^{\bar{\gamma}v21}_k\}_{ai,bj} = \delta_{a2}\delta_{b1}\{\mathbf{M}^{\bar{\gamma}v}_k\}_{ij}, \tag{II.135}$$

$$\{\mathbf{Q}_k^{\bar{\gamma}v22}\}_{ai,bj} = \delta_{a2}\delta_{b2}\{\mathbf{N}_k^{\bar{\gamma}v}\}_{ij}, \tag{II.136}$$

the following orthonormal basis of $\mathscr{W}^{\gamma_1\gamma_2;\gamma}$:

$$\mathbf{Q}_k^{\gamma vab},\quad a,b=1,2,\quad v=1,2,...,m_1(a,b),$$
$$m_1(1,1)=m_{\gamma_1\gamma_2;\gamma},\quad m_1(1,2)=m_{\gamma_1\bar{\gamma}_2;\gamma},$$
$$m_1(2,1)=m_{\bar{\gamma}_1\gamma_2;\gamma},\quad m_1(2,2)=m_{\bar{\gamma}_1\bar{\gamma}_2;\gamma},$$
$$k=1,2,...,n_\gamma, \tag{II.137}$$

$$\mathbf{Q}_k^{\bar{\gamma}vab},\quad a,b=1,2,\quad v=1,2,...,m_2(a,b),$$
$$m_2(1,1)=m_{\bar{\gamma}_1\bar{\gamma}_2;\gamma},\quad m_2(1,2)=m_{\bar{\gamma}_1\gamma_2;\gamma},$$
$$m_2(2,1)=m_{\gamma_1\bar{\gamma}_2;\gamma},\quad m_2(2,2)=m_{\gamma_1\gamma_2;\gamma},$$
$$k=1,2,...,n_\gamma, \tag{II.138}$$

where the vectors $\mathbf{K}_k^{\gamma v}$, $\mathbf{L}_k^{\gamma v}$, $\mathbf{M}_k^{\gamma v}$, and $\mathbf{N}_k^{\gamma v}$, and $\mathbf{K}_k^{\bar{\gamma}v}$, $\mathbf{L}_k^{\bar{\gamma}v}$, $\mathbf{M}_k^{\bar{\gamma}v}$, and $\mathbf{N}_k^{\bar{\gamma}v}$, represent columns of the CG matrices K, L, M, and N. Because of their definitions, the vectors (II.137) and (II.138) transform with respect to H according to

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{Q}_k^{\gamma vab} = \sum_l R^{\,\gamma}_{lk}(h)\mathbf{Q}_l^{\gamma vab}, \tag{II.139}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(h)\mathbf{Q}_k^{\bar{\gamma}vab} = \sum_l \{Z^{\gamma\dagger}R^{\,\bar{\gamma}}(h)Z^{\gamma}\}_{lk}\mathbf{Q}_l^{\bar{\gamma}vab},$$
$$\text{for all } h\in H, \tag{II.140}$$

and are therefore especially suited to simplifying the following considerations, although they are not a solution of Eqs. (II.120) and (II.121).

In order to be able to satisfy also these conditions, we define the following $M_{\gamma_1\gamma_2;\gamma}$-dimensional unitary transformations

$$\mathbf{W}_{1k}^{\gamma w} = \sum_{abv} B_{abv;w}\mathbf{Q}_k^{\gamma vab}, \tag{II.141}$$

$$\mathbf{Q}_k^{\gamma vab} = \sum_w B^*_{abv;w}\mathbf{W}_{1k}^{\gamma w},\quad k=1,2,...,n_\gamma, \tag{II.142}$$

$$\mathbf{W}_{2k}^{\gamma w} = \sum_{abv} C_{abv;w}\mathbf{Q}_k^{\bar{\gamma}vab}, \tag{II.143}$$

$$\mathbf{Q}_k^{\bar{\gamma}vab} = \sum_w C^*_{abv;w}\mathbf{W}_{2k}^{\gamma w},\quad k=1,2,...,n_\gamma, \tag{II.144}$$

and determine them in such a way that the corresponding vectors (II.141) and (II.143) are a solution of Eqs. (II.120) and (II.121), respectively.

For this purpose we consider

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{Q}_k^{\gamma vab*} = \sum_{a'b'v'} F_{a'b'v';abv}\mathbf{Q}_k^{\bar{\gamma}v'a'b'}, \tag{II.145}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{Q}_k^{\bar{\gamma}vab*} = \sum_l R^{\,\gamma}_{lk}(s^2)\sum_{a'b'v'} F_{abv;a'b'v'}\mathbf{Q}_l^{\gamma v'a'b'}, \tag{II.146}$$

where

$$F_{a'b'v';abv} = \{CB^T\}_{a'b'v';abv},$$
$$a',b'=1,2,\quad v'=1,2,...,m_2(a',b'),$$
$$a,b=1,2,\quad v=1,2,...,m_1(a,b). \tag{II.147}$$

These relations allow one to transform Eqs. (II.120) and (II.121) as follows:

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}_{1k}^{\gamma w*} = \sum_{w'} \{C^\dagger FB^*\}_{w'w}\mathbf{W}_{2k}^{\gamma w'}, \tag{II.148}$$

$$\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{W}_{2k}^{\gamma w*} = \sum_l R^{\,\gamma}_{lk}(s^2)\{C^\dagger FB^*\}^T_{w'w}\mathbf{W}_{1l}^{\gamma w'}. \tag{II.149}$$

Hence, if we can find unitary matrices B and C satisfying

$$C = FB^*, \tag{II.150}$$

the corresponding CG coefficients of type III are readily obtained from Eqs. (II.153) and (II.155). Otherwise, since F is unitary, we can choose

$$B = \mathbb{1}_M \Longleftrightarrow C = F, \tag{II.151}$$

which implies that the multiplicity index w can be identified with the triplet (a,b,v), i.e.,

$$w=(a,b,v),\quad a,b=1,2,\text{ and } v=1,2,...,m_1(a,b). \tag{II.152}$$

Inserting the special values (II.151) into Eqs. (II.141) and (II.143), we obtain

$$\mathbf{W}_{1k}^{\gamma(abv)} = \mathbf{Q}_k^{\gamma vab},\quad a,b=1,2,\text{ and } v=1,2,...,m_1(a,b), \tag{II.153}$$

$$\mathbf{W}_{2k}^{\gamma(abv)} = \sum_{a'b'}\sum_{v'=1}^{m_2(a'b')} F_{a'b'v';abv}\mathbf{Q}_k^{\bar{\gamma}v'a'b'},$$
$$a,b=1,2,\quad v=1,2,...,m_1(a,b), \tag{II.154}$$

respectively.

Therefore, the last step is to compute the matrix elements of F. For this purpose one has to carry out the scalar products

$$\langle\mathbf{Q}_k^{\bar{\gamma}v'a'b'},\mathbb{R}^{\gamma_1\gamma_2}(s)\mathbf{Q}_k^{\gamma vab*}\rangle = F_{a'b'v';abv}, \tag{II.155}$$

whose values are independent of k. This can be shown by means of

$$\mathbb{R}^{\gamma_1\gamma_2}(s)^\dagger\mathbf{E}_{ij}^{\bar{\gamma}}\mathbb{R}^{\gamma_1\gamma_2}(s) = \mathbf{E}_{ij}^{\gamma*} \tag{II.156}$$

and Eq. (II.139) and (II.140). Because of Eqs. (II.129)–(II.136), we obtain for Eq. (II.155)

$$F_{11v',22v} = F_{v'v}^{\bar{\gamma}(\gamma_1\gamma_2)} = \langle\mathbf{K}_k^{\bar{\gamma}v'},R^{\,\gamma_1}(s^2)\otimes R^{\,\gamma_2}(s^2)\mathbf{N}_k^{\gamma v*}\rangle, \tag{II.157}$$

$$F_{12v';21v} = F_{v'v}^{\bar{\gamma}(\gamma_1\bar{\gamma}_2)} = \langle\mathbf{L}_k^{\bar{\gamma}v'},R^{\,\gamma_1}(s^2)\otimes\mathbb{1}_{\gamma_2}\mathbf{M}_k^{\gamma v*}\rangle, \tag{II.158}$$

$$F_{21v';12v} = F_{v'v}^{\bar{\gamma}(\bar{\gamma}_1\gamma_2)} = \langle\mathbf{M}_k^{\bar{\gamma}v'},\mathbb{1}_{\gamma_1}\otimes R^{\,\gamma_2}(s^2)\mathbf{L}_k^{\gamma v*}\rangle, \tag{II.159}$$

$$F_{22v';11v} = F_{v'v}^{\bar{\gamma}(\bar{\gamma}_1\bar{\gamma}_2)} = \langle\mathbf{N}_k^{\bar{\gamma}v'},\mathbf{K}_k^{\gamma v*}\rangle, \tag{II.160}$$

$$F_{a'b'v';abv} = 0,\quad\text{otherwise}, \tag{II.161}$$

where the scalar product is analogously defined. In order to verify that the matrix elements (II.157)–(II.160) are independent of k, one has to use

$$\{R^{\,\gamma_1}(s^2)\otimes R^{\,\gamma_2}(s^2)\}^\dagger E_{ij}^{\gamma_1\gamma_2;\bar{\gamma}}\{R^{\,\gamma_1}(s^2)\otimes R^{\,\gamma_2}(s^2)\}$$
$$=\{E_{ij}^{\bar{\gamma}_1\bar{\gamma}_2;\gamma}\}^*, \tag{II.162}$$

$$\{R^{\,\gamma_1}(s^2)\otimes\mathbb{1}_{\gamma_2}\}^\dagger E_{ij}^{\gamma_1\bar{\gamma}_2;\bar{\gamma}}\{R^{\,\gamma_1}(s^2)\otimes\mathbb{1}_{\gamma_2}\}$$
$$=\{E_{ij}^{\bar{\gamma}_1\gamma_2;\gamma}\}^*, \tag{II.163}$$

$$\{\mathbb{1}_{\gamma_1}\otimes R^{\,\gamma_2}(s^2)\}^\dagger E_{ij}^{\bar{\gamma}_1\gamma_2;\bar{\gamma}}\{\mathbb{1}_{\gamma_1}\otimes R^{\,\gamma_2}(s^2)\}$$
$$=\{E_{ij}^{\gamma_1\bar{\gamma}_2;\gamma}\}^*, \tag{II.164}$$

$$E_{ij}^{\bar{\gamma}_1\bar{\gamma}_2;\bar{\gamma}} = \{E_{ij}^{\gamma_1\gamma_2;\gamma}\}^*, \tag{II.165}$$

and the transformation properties of the vectors $\mathbf{K}_k^{\gamma v}$, $\mathbf{L}_k^{\gamma v}$, $\mathbf{M}_k^{\gamma v}$, and $\mathbf{N}_k^{\gamma v}$, and $\mathbf{K}_k^{\bar{\gamma}v}$, $\mathbf{L}_k^{\bar{\gamma}v}$, $\mathbf{M}_k^{\bar{\gamma}v}$, and $\mathbf{N}_k^{\bar{\gamma}v}$ with respect to H.

Finally, let us assume that it is possible to apply the method given in Ref. 2, in order to determine the corresponding matrix elements of the CG matrices K, L, M, and N:

$$\{\mathbf{K}_k^{\gamma v}\}_{ij} = \{\mathbf{K}_k^{\gamma_1\gamma_2;\gamma(i_v,j_v)}\}_{ij}$$
$$= \|\mathbf{B}_{a_0}^{\gamma_1\gamma_2;\gamma(i_v,j_v)}\|^{-1}\frac{n_\gamma}{|H|}$$
$$\times \sum_h R^{\gamma_1}_{ii_v}(h) R^{\gamma_2}_{jj_v}(h) R^{\gamma *}_{ka_0}(h),$$
$$v = 1,2,..,m_1(1,1),\quad k = 1,2,...,n_\gamma, \tag{II.166}$$

$$\{\mathbf{L}_k^{\gamma v}\}_{ij} = \{\mathbf{L}_k^{\gamma_1\bar{\gamma}_2;\gamma(i_v,j_v)}\}_{ij}$$
$$= \|\mathbf{B}_{a_0}^{\gamma_1\bar{\gamma}_2;\gamma(i_v,j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h R^{\gamma_1}_{ii_v}(h)$$
$$\times \{Z^{\gamma_2\dagger} R^{\bar{\gamma}_2}(h) Z^{\gamma_2}\}_{jj_v} R^{\gamma *}_{ka_0}(h),$$
$$v = 1,2,...,m_1(1,2),\quad k = 1,2,...,n_\gamma, \tag{II.167}$$

$$\{\mathbf{M}_k^{\gamma v}\}_{ij} = \{\mathbf{M}_k^{\bar{\gamma}_1\gamma_2;\gamma(i_v,j_v)}\}_{ij}$$
$$= \|\mathbf{B}_{a_0}^{\bar{\gamma}_1\gamma_2;\gamma(i_v,j_v)}\|^{-1}\frac{n_\gamma}{|H|}$$
$$\times \sum_h \{Z^{\gamma_1\dagger} R^{\bar{\gamma}_1}(h) Z^{\gamma_1}\}_{ii_v} R^{\gamma_2}_{jj_v}(h) R^{\gamma *}_{ka_0}(h),$$
$$v = 1,2,...,m_1(2,1),\quad k = 1,2,...,n_\gamma, \tag{II.168}$$

$$\{\mathbf{N}_k^{\gamma v}\}_{ij}$$
$$= \{\mathbf{N}_k^{\bar{\gamma}_1\bar{\gamma}_2;\gamma(i_v,j_v)}\}_{ij}$$
$$= \|\mathbf{B}_{a_0}^{\bar{\gamma}_1\bar{\gamma}_2;\gamma(i_v,j_v)}\|^{-1}\frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma_1\dagger} R^{\bar{\gamma}_1}(h) Z^{\gamma_1}\}_{ii_v}$$
$$\times \{Z^{\gamma_2\dagger} R^{\bar{\gamma}_2}(h) Z^{\gamma_2}\}_{jj_v} R^{\gamma *}_{ka_0}(h),$$
$$v = 1,2,...,m_1(2,2),\quad k = 1,2,...,n_\gamma, \tag{II.169}$$

respectively. The components of the remaining vectors $\mathbf{K}_k^{\bar{\gamma}v}$, $\mathbf{L}_k^{\bar{\gamma}v}$, $\mathbf{M}_k^{\bar{\gamma}v}$, and $\mathbf{N}_k^{\bar{\gamma}v}$ are obtainable in principle from Eqs. (II.166)–(II.169) by replacing $R^{\gamma *}_{ka_0}(h)$ through $\{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}^*_{ka_0}$, the multiplicities $m_1(a,b)$ through $m_2(a,b)$, and the index sets $\{(i_v,j_v): v = 1,2,...,m_1(a,b)\}$ through suitable ones. Inserting these special values into Eqs. (II.157)–(II.160), we obtain

$$F^{\bar{\gamma}(\gamma_1\gamma_2)}_{v'v} = \|\mathbf{B}_{a_0}^{\gamma_1\gamma_2;\bar{\gamma}(i_{v'},j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\bar{\gamma}_1\bar{\gamma}_2;\gamma(i_v,j_v)}\|^{-1}$$
$$\times \frac{n_\gamma}{|H|}\sum_h R^{\gamma_1}_{i_{v'}i_v}(hs^2) R^{\gamma_2}_{j_{v'}j_v}(hs^2)$$
$$\times \{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}^*_{a_0a_0}, \tag{II.170}$$

$$F^{\bar{\gamma}(\gamma_1\bar{\gamma}_2)}_{v'v} = \|\mathbf{B}_{a_0}^{\gamma_1\bar{\gamma}_2;\bar{\gamma}(i_{v'},j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\bar{\gamma}_1\gamma_2;\gamma(i_v,j_v)}\|^{-1}$$
$$\times \frac{n_\gamma}{|H|}\sum_h R^{\gamma_1}_{i_{v'}i_v}(hs^2)$$
$$\times \{Z^{\gamma_2\dagger} R^{\bar{\gamma}_2}(h) Z^{\gamma_2}\}_{j_{v'}j_v}\{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}^*_{a_0a_0}, \tag{II.171}$$

$$F^{\bar{\gamma}(\bar{\gamma}_1\gamma_2)}_{v'v} = \|\mathbf{B}_{a_0}^{\bar{\gamma}_1\gamma_2;\bar{\gamma}(i_{v'},j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\gamma_1\bar{\gamma}_2;\gamma(i_v,j_v)}\|^{-1}$$
$$\times \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma_1\dagger} R^{\bar{\gamma}_1}(h) Z^{\gamma_1}\}_{i_{v'}i_v}$$
$$\times R^{\gamma_2}_{j_{v'}j_v}(hs^2)\{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}^*_{a_0a_0}, \tag{II.172}$$

$$F^{\bar{\gamma}(\bar{\gamma}_1\bar{\gamma}_2)}_{v'v} = \|\mathbf{B}_{a_0}^{\bar{\gamma}_1\bar{\gamma}_2;\bar{\gamma}(i_{v'},j_{v'})}\|^{-1}\|\mathbf{B}_{a_0}^{\gamma_1\gamma_2;\gamma(i_v,j_v)}\|^{-1}$$
$$\times \frac{n_\gamma}{|H|}\sum_h \{Z^{\gamma_1\dagger} R^{\bar{\gamma}_1}(h) Z^{\gamma_1}\}_{i_{v'}i_v}$$
$$\times \{Z^{\gamma_2\dagger} R^{\bar{\gamma}_2}(h) Z^{\gamma_2}\}_{j_{v'}j_v}$$
$$\times \{Z^{\gamma\dagger} R^{\bar{\gamma}}(h) Z^{\gamma}\}^*_{a_0a_0} = \delta_{vv'}, \tag{II.173}$$

respectively, whose values are of course independent of k. The special result (II.173) is a consequence of the method given in Ref. 2.

Apart from these special values for the matrix elements (II.157)–(II.160), the final formulas read as

$$\mathbf{W}_{1k}^{\gamma(abv)} = \mathbf{Q}_k^{\gamma vab},\quad a,b = 1,2,\ \text{and}\ v = 1,2,...,m_1(a,b), \tag{II.174}$$

$$\mathbf{W}_{2k}^{\gamma(11v)} = \sum_{v'} F^{\bar{\gamma}(\bar{\gamma}_1\bar{\gamma}_2)}_{v'v}\mathbf{Q}_k^{\bar{\gamma}v'22},\quad v = 1,2,...,m_1(1,1), \tag{II.175}$$

$$\mathbf{W}_{2k}^{\gamma(12v)} = \sum_{v'} F^{\bar{\gamma}(\bar{\gamma}_1\gamma_2)}_{v'v}\mathbf{Q}_k^{\bar{\gamma}v'21},\quad v = 1,2,...,m_1(1,2), \tag{II.176}$$

$$\mathbf{W}_{2k}^{\gamma(21v)} = \sum_{v'} F^{\bar{\gamma}(\gamma_1\bar{\gamma}_2)}_{v'v}\mathbf{Q}_k^{\bar{\gamma}v'12},\quad v = 1,2,...,m_1(2,1), \tag{II.177}$$

$$\mathbf{W}_{2k}^{\gamma(22v)} = \sum_{v'} F^{\bar{\gamma}(\gamma_1\gamma_2)}_{v'v}\mathbf{Q}_k^{\bar{\gamma}v'11},\quad v = 1,2,...,m_1(2,2), \tag{II.178}$$

which show that CG coefficients of type III also for the most complicated case are connected by simple unitary transformations with convenient CG coefficients for H. Similar to the other cases, the only problem is to compute the submatrices $F^{\bar{\gamma}(\cdot\cdot)}$, but whose dimensions are not in general equal.

SUMMARY

Within the last of this series of papers we investigated the most complicated situation by considering Kronecker products which are composed of co-unirreps of type III only. Due to our approach the first task is to determine convenient CG matrices which decompose the corresponding terms of $\mathbb{R}^{\gamma_1\gamma_2}\downarrow H = R^{\gamma_1\gamma_2}\oplus R^{\gamma_1\bar{\gamma}_2}\oplus R^{\bar{\gamma}_1\gamma_2}\oplus R^{\bar{\gamma}_1\bar{\gamma}_2}$ into direct sums of unirreps of H. These CG matrices are denoted by K, L, M, and N. Provided such matrices have been determined, CG coefficients for corepresentations are readily obtained as follows.

In the case of CG coefficients of type I it is only necessary to compute the submatrices $F^{\alpha(\gamma_1\gamma_2)}$ and $F^{\alpha(\gamma_1\bar{\gamma}_2)}$ of B:

$$B = \frac{1}{\sqrt{2}}\begin{bmatrix} i\mathbf{1}_m & 0 & 0 & \mathbf{1}_m \\ 0 & i\mathbf{1}_{m'} & \mathbf{1}_{m'} & 0 \\ 0 & -i\mathbf{G}^T & \mathbf{G}^T & 0 \\ -i\mathbf{F}^T & 0 & 0 & \mathbf{F}^T \end{bmatrix},$$

where $m = m_{\gamma_1\gamma_2;\alpha}$ and $m' = m_{\gamma_1\bar{\gamma}_2;\alpha}$, and $\mathbf{F} = F^{\alpha(\gamma_1\gamma_2)}$ and $\mathbf{G} = F^{\alpha(\gamma_1\bar{\gamma}_2)}$, since the corresponding CG coefficients follow from Eqs. (II.60)–(II.63). Thereby, one has to use the definitions (II.17)–(II.20). The matrix elements of these matrices should be reasonably calculated with the aid of Eq. (II.34) and (II.35) together with the relations (II.42) and (II.43). Besides this the symmetry relations $m_{\gamma_1\gamma_2;\alpha} = m_{\bar{\gamma}_1\bar{\gamma}_2;\alpha}$ and

$m_{\gamma_1\bar{\gamma}_2;\alpha} = m_{\bar{\gamma}_1\gamma_2;\alpha}$ give rise to simple solutions for the multiplicity problem, although $m_{\gamma_1\gamma_2;\alpha}$ and $m_{\gamma_1\bar{\gamma}_2;\alpha}$ are not necessarily equal.

CG coefficients of type II are given by Eqs. (II.115)–(II.117), where the definitions (II.71)–(II.74) have to be taken into account. Consequently, it suffices to compute the submatrices $F^{\beta(\gamma_1\gamma_2)}$ and $F^{\beta(\gamma_1\bar{\gamma}_2)}$ of B:

$$B = \begin{bmatrix} \mathbf{1}_m & & 0 & 0 \\ 0 & \mathbf{1}_{m'} & 0 & 0 \\ 0 & 0 & 0 & F^{\beta(\gamma_1\bar{\gamma}_2)^T} \\ 0 & 0 & F^{\beta(\gamma_1\gamma_2)^T} & 0 \end{bmatrix},$$

where however, the multiplicities $m = m_{\gamma_1\gamma_2;\beta}$ and $m' = m_{\gamma_1\bar{\gamma}_2;\beta}$ will not in general be equal. The matrix elements of these matrices are defined by Eqs. (II.90) and (II.88), where the symmetry relation (II.97) should be utilized in any case. Nevertheless, we arrive also for this case at a simple solution for the multiplicity problem.

For the last case we have found the special solution

$$B = \mathbf{1}_M$$

and

$$C = \begin{bmatrix} 0 & 0 & 0 & F^{\bar{\gamma}(\gamma_1\gamma_2)} \\ 0 & 0 & F^{\bar{\gamma}(\gamma_1\bar{\gamma}_2)} & 0 \\ 0 & F^{\bar{\gamma}(\bar{\gamma}_1\gamma_2)} & 0 & 0 \\ F^{\bar{\gamma}(\bar{\gamma}_1\bar{\gamma}_2)} & 0 & 0 & 0 \end{bmatrix},$$

which yields the final formulas (II.174)–(II.178), where the definitions (II.129)–(II.136) have to be used. Thus, the only problem is to compute the unitary submatrices $F^{\bar{\gamma}(\cdots)}$ of F by means of Eqs. (II.157)–(II.160).

Besides this let us summarize the main points of the present approach. Considering the results it is reasonable to divide the determination of CG coefficients for corepresentations into two steps. The first step is to compute convenient CG coefficients for the normal subgroup. The second step is to use these CG coefficients in order to define special bases, whose elements transform according to unirreps of the normal subgroup and are therefore especially suited to determine CG coefficients for corepresentations. Their transformation properties with respect to a special representative of the antiunitary group elements are utilized to derive simple defining equations for those unitary transformations which link these special bases with corresponding columns of CG matrices for corepresentations. Thereby, we obtain for each type of co-unirrep in principle the same type of defining equations, namely,

$$\text{type I:} \quad FB^* = B, \quad FF^* = +\mathbf{1}_M,$$

$$\text{typeII:} \quad FB^* = BG^T, \quad FF^* = -\mathbf{1}_{2M},$$

$$\text{typeIII:} \quad FB^* = C, \quad FF^\dagger = \mathbf{1}_M,$$

but whose structure depends essentially on the considered Kronecker product and lead therefore to quite different solutions. Apart from the first two cases, which are contained in the first paper, we were able to find solutions for these equations. These give rise to special solutions for the multiplicity problem, where the corresponding multiplicity problems referring to subductions with respect to the normal subgroup play an important role. The only problem for the present method is to compute unitary submatrices $F^{\mu(\cdots)}$ whose dimensions are much smaller than that of the considered Kronecker product and which are uniquely fixed through the given CG coefficients for the normal subgroup.

Apart from the general solution for the underlying problem, we discussed in each paper the possibility that the CG coefficients for the normal subgroup are representable in a very special way. This led to special values for the matrix elements of $F^{\mu(\cdots)}$. The reason for considering this possibility arises therefrom that CG coefficients for space groups can be traced back for nearly all cases to this special form, so that the present results can be directly transferred to magnetic space groups. This transfer to projective corepresentations and its application to magnetic space groups should be discussed in a forthcoming paper.

[1]C.J. Bradley and A.P. Cracknell, *The Mathematical Theory of Symmetry in Solids* (Clarendon, Oxford, 1972).
[2]R. Dirl, J. Math. Phys. 20, 659 (1979).

A connection between nonlinear evolution equations and ordinary differential equations of P-type. II

M. J. Ablowitz
Mathematics and Computer Science, Clarkson College, Potsdam, New York 13676
A. Ramani[a]
Program in Applied Mathematics, Princeton University, Princeton, New Jersey 08540
H. Segur
Aeronautical Research Associates of Princeton, Inc. Princeton, New Jersey 08540

(Received 15 May 1979; accepted for publication 17 December 1979)

It is known through the inverse scattering transform that certain nonlinear differential equations can be solved via linear integral equations. Here it is demonstrated "directly," i.e., without the Jost-function formalism that the solution of the linear integral equation actually solves the nonlinear differential equation. In particular, this extends the scope of inverse scattering methods to ordinary differential equations which are found to be of Painlevé type. Some global properties of these nonlinear ODE's are obtained rather easily by this approach.

1. INTRODUCTION

The purpose of this paper is to show, following the work of Zakharov and Shabat,[1] that certain nonlinear differential equations can be solved via linear integral equations. This method does not require any analytic properties of an associated scattering problem, which are required to solve nonlinear partial differential equations by inverse scattering transforms (IST; cf. Ref. 2). Because of this extra freedom, the range of solutions obtained by the present approach is larger than that obtained by IST. In particular, self-similar solutions of these equations, which satisfy nonlinear *ordinary* differential equations, can be obtained. Thus, the method provides exact linearization of both nonlinear PDE's and ODE's.

Our interests in these nonlinear ODE's are twofold. First, they often have "physical" significance. Indeed, for many of the well-known, physically interesting nonlinear evolution equations, the solution for large times separates into a finite number of solitons plus a remainder which asymptotically (as $t\rightarrow\infty$) approaches a modulated similarity solution (see, for example, Ref. 3). Second, the nonlinear ODE's in question turn out to have no movable branch points or essential singularities.[4–6] (Hereafter we shall refer to ODE's with this property as being of P-type; P is for Painlevé). In Refs. 4 and 5, a conjecture is given regarding the relationship between nonlinear evolution equations solvable by IST and nonlinear ODE's of P-type. In fact, for many of the well-known nonlinear evolution equations (KdV, MKdV, sine-Gordon, etc.), the similarity solution is one of the Painlevé transcendents (see also Refs. 4–7).

In Sec. 2, we discuss a procedure for deriving the differential equation directly from an associated linear integral equation, which is in Gel'fand–Levitan–Marchenko form. The derivation applies to partial as well as ordinary differential equations. We require only that the solutions decay rapidly enough as $x\rightarrow +\infty$ (say) that the integral operators are defined. We stress, however, that in general a solution which decays as $x\rightarrow+\infty$ may diverge at some finite value of x, diverge as $x\rightarrow-\infty$, or have weak decay as $x\rightarrow-\infty$. In any of these cases, the classical analysis of inverse scattering using the analytic properties of the Jost functions is not applicable, since it requires "nice" properties of the potential on the whole line (see Refs. 8 and 9).

Examples of the method are studied in detail in Sec. 3. Various ways of proving that the solution of the associated integral equation actually exists and is unique are given in Sec. 4.

2. DERIVATION OF THE DIFFERENTIAL EQUATION

Consider the linear integral equation

$$K(x,y)=F(x,y)+\int_x^\infty K(x,z)N(x;z,y)dz,\quad y\geqslant x. \tag{2.1}$$

Besides the arguments (x,y,z) which explicitly appear in Eq. (2.1), F, N, and K may depend on other parameters $(t,\lambda,\cdots)$. Derivatives with respect to these extra variables may appear in the differential equations that F and K satisfy, but Eq. (2.1) is understood to be solved at fixed given values of these parameters.

In each specific case N is explicitly given in terms of F. For instance,

(A) $N(x;z,y)=F(z,y)$ (KdV, higher order KdV's),

(B) $N(x;z,y)=\pm\int_x^\infty F(z,s)F(s,y)ds$ (modified KdV, sine-Gordon, etc.),

(B') $N(x;z,y)=\pm\int_x^\infty F^*(z,s)F(s,y)ds$ (nonlinear Schrödinger, etc.),

where * refers to complex conjugation

[a] Permanent address: L.P.T.H.E. Université Paris-Sud, 91405 Orsay, France.

(C) $N(x;z,y) = \pm i \int_x^\infty \left[\frac{\partial_z + \partial_s}{2} F^*(z,s) \right] F(s,y) ds,$

(massive Thirring, derivative nonlinear Schrödinger, etc.),

(D) $N(x;z,y) = \int_x^\infty \int_x^\infty F(z,s) F(s,v) F(v,y) ds\, dv.$

Throughout this paper, we consider only those cases in which

$$F(x,y) = F(x+y). \tag{2.2}$$

However, other possibilities certainly exist.

In the usual approach, F is constructed from the scattering data of the "direct problem" and the scattering potential $u(x)$ is reconstructed from K [e.g., $u(x) = K(x,x)$ or $u(x) = (d/dx)K(x,x)$]. Here we do not give to F any interpretation, but only demand that it satisfies some linear (partial or ordinary) differential equation.

Define the operator A_x by

$$A_x f(y) = \int_x^\infty f(z) N(x;z,y) dz, \quad y \geqslant x,$$
$$= 0, \quad y < x. \tag{2.3}$$

We assume that for each specific choice of N, one can prove that $(I - A_x)$ is invertible. More precisely, there is an x large enough and a function space on which $(I - A_x)$ is invertible, and $(I - A_x)^{-1}$ is continuous. Moreover, we assume that the operators obtained by differentiating (2.3) with respect to x or y also are defined on this function space. It will be shown in Sec. 4 that these assumptions are valid in a large variety of problems.

Subject to these assumptions and the fact that F satisfies some linear differential equation, we prove in this section that $u(x)$ (defined above) satisfies a nonlinear differential equation. We shall say that this equation is of "inverse scattering type" even though no reference is made to the direct scattering problem.

The outline of this approach can be stated rather simply:

(i) F satisfies two linear (partial or ordinary) differential equations

$$\mathscr{L}_i F = 0, \quad i = 1,2. \tag{2.4}$$

These two operators can be related to the x dependence and t dependence of one of the eigenfunctions in the usual IST approach (e.g., see Ref. 2). Through this paper,

$$\mathscr{L}_1 = (\partial_x - \partial_y),$$

so that $\mathscr{L}_1 F = 0$ implies Eq. (2.2), but this choice is not essential. (ii) K is related to F through (2.1), which we may write in the form

$$(I - A_x) K = F. \tag{2.1a}$$

(iii) Applying $\mathscr{L}_i$ to this equation yields

$$\mathscr{L}_i (I - A_x) K = 0, \quad i = 1,2. \tag{2.5}$$

This can be rewritten as

$$(I - A_x)(\mathscr{L}_i K) = R_i,$$

where R_i contains all the remaining terms after operating with $\mathscr{L}_i$. However, (2.1) and (2.4) may be such that

$$R_i = (I - A_x) M_i(K), \quad i = 1,2, \tag{2.6}$$

where $M_i(K)$ are nonlinear functionals of K. (iv) Therefore,

$$(I - A_x)[\mathscr{L}_i K - M_i(K)] = 0.$$

However, $(I - A_x)$ is invertible, so K must satisfy the nonlinear differential equation

$$\mathscr{L}_i K - M_i(K) = 0, \quad i = 1,2. \tag{2.7}$$

Therefore, every solution of the linear integral equation (2.1) is also a solution of the nonlinear differential equations (2.7). Making use of the results for $i = 1$ allows us to find a differential equation on line $y = x$ for $i = 2$.

Thus, the basic ingredients to this approach are a linear integral equation (2.1) and two linear differential operators $\mathscr{L}_1$, $\mathscr{L}_2$. However, some more information is required in order to make the method effective. First, we must identify a class of suitable differential operators $\mathscr{L}$. Second, we must establish that K, defined by Eq. (2.1), is differentiable enough that $(\mathscr{L}_i K)$ even exists. Third, in order to implement Eq. (2.6), it is convenient to establish a "dictionary" of the terms that may appear in R. These last two steps depend only on the order of $\mathscr{L}$, not on $\mathscr{L}$ itself. Then given this information, we may show how applying $\mathscr{L}$ to Eq. (2.1) generates Eq. (2.7), the nonlinear equation satisfied by K.

Here are some of the possible types of differential equations for F (i.e., $\mathscr{L}_2 F = 0$):

(i) There are evolution equations of the type

$$i\partial_t F = Q(i\partial_x) F,$$

where t is an auxiliary argument of F, and Q is a polynomial with real constant coefficients which is otherwise arbitrary in cases (B') and (C), but must be odd in cases (A), (B), and (D).

(ii) Linear ODE's for F are obtained by reducing one of the equations by a suitable ansatz (similarity form).[4] For instance, the ansatz for traveling waves

$$F(x+y;t) = \mathscr{F}(x+y-2vt),$$
$$K(x,y;t) = \mathscr{K}(x-vt, y-vt)$$

is always compatible with the integral equation.

(iii) If $Q(i\partial_x) = (i\partial_x)^p$, one can also look for a similarity solution of scaling type. Compatibility with the integral equation fixes the behavior of F:

$$F(x,t) = \frac{1}{x} \tilde{f}(xt^{-1/p}) = t^{-1/p} f(xt^{-1/p}),$$
in case (A), (B), (B'), (D),

$$F(x,t) = x^{-1/2} \tilde{f}(xt^{-1/p}) = t^{-1/2p} f(xt^{-1/p}),$$
in case (C).

Indeed, N scales like x^{-1}, and like F in case (A). Similarly, N scales like $F^2 x$ in case (B) and (B'), like F^2 in case (C) and like $F^3 x^2$ in case (D). More complicated ansatzes can also be found in some cases. If F satisfies an ODE which is obtained by such an ansatz, the correct operator $\mathscr{L}$ is found by using the same reduction. For instance, if

$$F(x+y;t) = t^{-1/p} f((x+y)t^{-1/p}),$$
$$K(x,y;t) = t^{-1/p} \mathscr{K}(xt^{-1/p}, yt^{-1/p}),$$

the operator $i\partial_t$ leads to the operator

$-i/p\{I+\xi(\partial/\partial\xi)+\eta(\partial/\partial\eta)\}$ where

$\xi = xt^{-1/p}, \quad \eta = yt^{-1/p}.$

(iv) The equation $\partial_x\partial_t F = \lambda F$, λ constant, does not lead directly to a differential equation for u but to an integrodifferential equation. However, by changes of variables it is then possible to recover the sine-Gordon and sinh-Gordon equations from case (B) and the equations of the massive Thirring model from case (C). Here also, ODE's can be found from the PDE using either the traveling wave or scaling similarity ansatz.

Consider next the differentiability of K. It is convenient to define $\eta = y - x$, so that (2.1) becomes

$$K(x,x+\eta) = F(2x+\eta) + \int_0^\infty K(x,x+\zeta) \times N(x;x+\zeta;x+\eta)d\zeta, \quad \eta \geqslant 0. \tag{2.8}$$

Now F must be differentiable, by (2.4). We may assume as well that for fixed (large enough) x, $(I - A_x)$ is invertible and its inverse is bounded and therefore continuous. Then (2.8) defines $K(x,y)$ as an element of a certain function space. We also assume that the operators defined by taking x or y derivatives of A_x are defined on this same function space. Then derivatives of K are defined by forming the appropriate difference quotients in Eq. (2.8) and taking limits. In this way, the following results can be established:

(i) $\partial_\eta K(x,x+\eta)$ exists and may be computed directly from Eq. (2.8) in terms of $\partial_\eta F$, $\partial_\eta N$, and K.

(ii) For $\eta > 0$, $\partial_x K$ exists; it is the solution of an equation of the form

$(I - A_x)(\partial_x K) = R(x, x+\eta).$

(iii) The procedure in (ii) fails at $\eta = 0$, and Eq. (2.3) shows that K is discontinuous there. However, $(\partial_x + \partial_\eta)K$ exists along $\eta = 0$, and $(\partial_x K)$ exists as a one-sided derivative there.

(iv) Higher derivatives can be established in the same way, provided the required derivatives of F are sufficiently well behaved.

After these preliminary steps, it remains to establish a "dictionary" for each linear equation, and to apply $\mathscr{L}_2$. This is the heart of the method, which we illustrate with examples in the next section.

3. EXAMPLES

A. Modified KdV and Painlevé II

Consider the linear integral equation (2.1) for case (B). If $\mathscr{L}_1 = (\partial_x - \partial_y)$, then we may write the solution of $\mathscr{L}_1 F = 0$ as $F = F((x+y)/2)$, and put Eq. (2.1) in the form ($\sigma = \pm 1$)

$$K(x,y) = F\left(\frac{x+y}{2}\right) + \frac{\sigma}{4}\int\int_x^\infty K(x,z)F\left(\frac{z+u}{2}\right)F\left(\frac{u+y}{2}\right)dz\,du.$$

For this example, the second operator for $F(x,t)$ is

$$\mathscr{L}_2 F = (\partial_t + \partial_x^3)F = 0. \tag{3.1}$$

We begin by constructing a dictionary, which we will again use in example B.

We shift the origin to avoid boundary terms

$$K(x,y) = F\left(\frac{x+y}{2}\right) + \frac{\sigma}{4}\int\int_0^\infty K(x,x+\zeta) \times F\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right)d\zeta\,d\eta. \tag{3.2}$$

or

$$[(I-\sigma A_x)K](x,y) = F\left(\frac{x+y}{2}\right),$$

where A_x is defined as in (2.3). Defining

$$K_2(x,z) = \int_0^\infty K(x,x+\zeta)F\left(\frac{x+\zeta+z}{2}\right)d\zeta, \tag{3.3}$$

one easily shows that

$$(I-\sigma A_x)K_2(x,z) = \int_0^\infty F\left(\frac{2x+\zeta}{2}\right)F\left(\frac{x+\zeta+z}{2}\right)d\zeta. \tag{3.4}$$

It is convenient to write the integral equation (3.2) as

$$K(x,y) = F\left(\frac{x+y}{2}\right) + \frac{\sigma}{4}\int_0^\infty K_2(x,x+\eta)F\left(\frac{x+\eta+y}{2}\right)d\eta. \tag{3.5}$$

Applying the operator $(\partial_x - \partial_y)$ to (3.5) yields

$$(\partial_x - \partial_y)K(x,y) = \frac{\sigma}{4}\int_0^\infty [(\partial_1+\partial_2)K_2(x,x+\eta)]F\left(\frac{x+\eta+y}{2}\right)d\eta, \tag{3.6}$$

where ∂_1 and ∂_2 are derivatives with respect to the first and second arguments of K_2, respectively. Similarly, applying $(\partial_x + \partial_z)$ to Eq. (3.3) yields

$$\begin{aligned}(\partial_x+\partial_z)K_2(x,z) &= \int_0^\infty \left\{(\partial_1+\partial_2)K(x,x+\zeta)F\left(\frac{x+\zeta+z}{2}\right) + K(x,x+\zeta)F'\left(\frac{x+\zeta+z}{2}\right)\right\}d\zeta \\ &= \int_0^\infty [(\partial_1-\partial_2)K(x,x+\zeta)]F\left(\frac{x+\zeta+z}{2}\right)d\zeta \\ &\quad -2K(x,x)\,F\left(\frac{x+z}{2}\right).\end{aligned} \tag{3.7}$$

Substituting (3.6) into (3.7), we see that

$$\begin{aligned}(I-\sigma A_x)(\partial_x+\partial_z)K_2(x,z) &= -2K(x,x)F\left(\frac{x+z}{2}\right) \\ &= -2K(x,x)(I-\sigma A_x)K(x,z).\end{aligned}$$

Substituting (3.7) into (3.6) leads to

$$\begin{aligned}(I-\sigma A_x)(\partial_x-\partial_y)K(x,y) &= -\frac{\sigma}{2}K(x,x)\int_0^\infty F\left(\frac{2x+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right)d\eta \\ &= -\frac{\sigma}{2}K(x,x)(I-\sigma A_x)K_2(x,y).\end{aligned}$$

Note from (2.3) that A_x commutes with multiplication by a function of x only. Thus, if $(I-\sigma A_x)$ is invertible, we have proven that

$$(\partial_x + \partial_z)K_2(x,z) = -2K(x,x)K(x,z),$$
$$(\partial_x - \partial_y)K(x,y) = -\frac{\sigma}{2}K(x,x)K_2(x,y). \tag{3.8}$$

There are the results expected from the inverse scattering analysis [Ref. 2, Eq. (4.46)]. However, here we obtained them using only invertibility of $(I - \sigma A_x)$, which is a much weaker condition than what is required to apply the usual analytic approach.

Now apply $(\partial_x + \partial_y)$ to (3.2):

$$(\partial_x + \partial_y)K(x+y) = F' + \frac{\sigma}{4}\int\int_0^\infty [(\partial_1 + \partial_2)K(x,x+z)]F\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right)d\zeta\, d\eta$$
$$+ \frac{\sigma}{4}\int\int_0^\infty K(x,x+\zeta)(\partial_x + \partial_y)\left[F\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right)\right]d\zeta\, d\eta. \tag{3.9}$$

However,

$$(\partial_x + \partial_y)F\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right) = F'\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right) + F\left(\frac{2x+\zeta+\eta}{2}\right)F'\left(\frac{x+\eta+y}{2}\right)$$
$$= 2\partial_\eta\left\{F\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right)\right\}. \tag{3.10}$$

Performing the η integration in (3.9) leads to

$$(I - \sigma A_x)(\partial_x + \partial_y)K(x,y) = F'\left(\frac{x+y}{2}\right) - \frac{\sigma}{2}\int_0^\infty K(x,x+\zeta)F(2x+\zeta)d\zeta F(x+y)$$
$$= F'\left(\frac{x+y}{2}\right) - \frac{\sigma}{2}K_2(x,x)(I - \sigma A_x)K(x,y),$$

i.e.,

$$F'\left(\frac{x+y}{2}\right) = (I - \sigma A_x)\left\{(\partial_x + \partial_y)K(x,y) + \frac{\sigma}{2}K_2(x,x)K(x,y)\right\}. \tag{3.11}$$

This completes the dictionary required for this problem.

The final step makes use of the fact that F satisfies (3.1), i.e., $L_2F = 0$. Apply the operator $L_2 = \{\partial_t + (\partial_x + \partial_y)^3\}$ to (3.2):

$$\{\partial_t + (\partial_x + \partial_y)^3\}K(x,y) = 0 + \frac{\sigma}{4}\{\partial_t + (\partial_x + \partial_y)^3\}\int\int_0^\infty K(x,x+\zeta)F\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\eta+y}{2}\right)d\eta\, d\zeta. \tag{3.12}$$

The terms on the right side of (3.12) proliferate when the differentiation is performed under the integral, but several simplifying cancellations occur. For example, using (3.1) leads to

$$\{\partial_t + (\partial_x + \partial_y)^3\}F\left(\frac{2x+\zeta+\eta}{2}\right)F\left(\frac{x+\zeta+y}{2}\right) = 6\partial_\eta\left[F'\left(\frac{2x+\zeta+\eta}{2}\right)F'\left(\frac{x+\eta+y}{2}\right)\right].$$

It follows that (3.12) is equivalent to

$$(I - \sigma A_x)\{\partial_t + (\partial_x + \partial_y)^3\}K(x,y) = -\frac{3\sigma}{2}\left[\partial_x\int_0^\infty d\zeta\left\{\frac{\partial_1 + \partial_2}{2}K(x,x+\zeta)\right\}F\left(\frac{2x+\zeta}{2}\right)\right]F\left(\frac{x+y}{2}\right)$$
$$-\frac{3\sigma}{2}[\partial_x K_2(x,x)]F'\left(\frac{x+y}{2}\right). \tag{3.13}$$

However from (3.8),

$$\partial_x K_2(x,x) = -2K^2(x,x),$$

and

$$\partial_x\int_0^\infty d\zeta F\left(\frac{2x+\zeta}{2}\right)\frac{\partial_1 + \partial_2}{2}K(x,x+\zeta) = \partial_x[(\partial_x - \partial_y)K_2(x,y)]_{y=x}$$
$$= [(\partial_x + \partial_y)(\partial_x - \partial_y)K_2(x,y)]_{y=x}$$
$$= (\partial_x - \partial_y)\{-2K(x,x)K(x,y)\}_{y=x}$$
$$= -2[\partial_x K(x,x)]K(x,x) + \sigma K^2(x,x)K_2(x,x).$$

Then using (3.2) and (3.11) and the invertibility of $(I - \sigma A_x)$, (3.13) becomes (for $y \geqslant x$)

$$\{\partial_t + (\partial_x + \partial_y)^3\}K(x,y) = 3\sigma K(x,x)K(x,y)\partial_x K(x,x) + 3\sigma K^2(x,x)(\partial_x + \partial_y)K(x,y). \tag{3.14}$$

If we define

$$q(x,t) = K(x,x;t) \tag{3.15}$$

and evaluate (3.14) along $y = x$, then

$$\partial_t q + \partial_x^3 q = 6\sigma q^2 q_x, \tag{3.16}$$

i.e., q satisfies the modified Korteweg–de Vries equation.

Thus, every solution of $\mathscr{L}_i F = 0, i = 1, 2$ [i.e., Eq. (3.1) with $F = F((x+y)/2)$] that decays fast enough as $x \to \infty$ defines a solution of Eq. (3.16), via the linear integral equation (3.2). No global properties (on $-\infty < x < \infty$) are required. A special case of interest is obtained if F and K are self-similar:

$$K(x,y;t) = (3t)^{-1/3}\mathscr{K}(\xi,\eta), \quad F\left(\frac{x+y}{2};t\right) = (3t)^{-1/3}\mathscr{F}\left(\frac{\xi+\eta}{2}\right), \tag{3.17}$$

where

$$\xi = x/(3t)^{1/3}, \quad \eta = y/(3t)^{1/3}.$$

substituting these into Eq. (3.2) shows that K satisfies an equation of the same form:

$$\mathscr{K}(\xi,\eta) = \mathscr{F}\left(\frac{\xi+\eta}{2}\right) + \frac{\sigma}{4}\iint_\xi^\infty \mathscr{K}(\xi,\zeta)\mathscr{F}\left(\frac{\zeta+\psi}{2}\right)\mathscr{F}\left(\frac{\psi+\eta}{2}\right)d\zeta\, d\psi, \quad \eta \geqslant \xi. \tag{3.18}$$

Substituting (3.17) into (3.1) yields

$$\mathscr{F}'''(\xi) - [\mathscr{F}(\xi) + \xi\mathscr{F}'(\xi)] = 0,$$

which can be integrated once:

$$\mathscr{F}''(\xi) - \xi\mathscr{F}(\xi) = C_1. \tag{3.19}$$

If $C_1 = 0$, the solutions of (3.19) that vanish as $\xi \to \infty$ are multiples of the Airy function:

$$\mathscr{F}\left(\frac{\xi+\eta}{2}\right) = r\mathrm{Ai}\left(\frac{\xi+\eta}{2}\right). \tag{3.20}$$

Meanwhile, $Q(\xi) = \mathscr{K}(\xi,\xi)$ must solve the similarity form of Eq. (3.16):

$$Q''' - [Q + \xi Q'] = 6\sigma Q^2 Q',$$

which can also be integrated once:

$$Q'' = \xi Q + 2\sigma Q^3 + C_2. \tag{3.21}$$

This nonlinear ODE is the second equation of Painlevé ($\mathrm{P_{II}}$). What we have shown here is that every solution of the linear integral equation (3.18), in which $\mathscr{F}$ is defined by (3.19), is also a solution of (3.21).

In particular, if $C_1 = 0$ then it follows from (3.18) that $Q(\xi)$ is exponentially small as $\xi \to \infty$, so that C_2 vanishes in (3.21), which becomes

$$Q'' = \xi Q + 2\sigma Q^3. \tag{3.22}$$

Then if (3.20) is used in Eq. (3.18), a one parameter family of solutions of (3.22) is obtained from the solutions of

$$[I - \sigma r^2 \widetilde{A}_\xi]\mathscr{K}(\xi,\eta;r) = r\,\mathrm{Ai}\left(\frac{\xi+\eta}{2}\right), \tag{3.23}$$

where

$$\widetilde{A}_\xi f(\eta) = \frac{1}{4}\iint_\xi^\infty f(\xi)\mathrm{Ai}\left(\frac{\zeta+\psi}{2}\right)\mathrm{Ai}\left(\frac{\psi+\eta}{2}\right)d\zeta\, d\psi,$$

using $Q(\xi;r) = \mathscr{K}(\xi,\xi;r)$. This result was first obtained in Ref. 3. Let us now show how simple it is to obtain some of the global properties of these solutions of (3.21) that decay like $r\,\mathrm{Ai}(\xi)$ as $\xi \to \infty$ (see also Ref. 10).

(i) The only singularities of these solutions in the complex plane are poles. This result was first obtained by Painlevé (cf. Ref. 11), using his α-method. It follows from (3.23) by using the Fredholm theory of integral equations (see Ref. 6 for details). A direct consequence of this theory and the fact that the Airy function is entire is the fact that the solution of Eq. (3.23) can have only poles.

(ii) For $\sigma = -1$, the solutions of (3.22) that vanish as $\xi \to +\infty$ are bounded for all real ξ and for all real r. For real ξ, $\widetilde{A}_\xi$ is a positive operator (as shown in Sec. 4). Hence, for real r, $(I + r^2\widetilde{A}_\xi)$ is always invertible, and the solution of (3.23) exists and is bounded (cf. Ref. 12).

(iii) For $\sigma = +1$, the solutions of Eq. (3.22) that decay as $\xi \to \infty$ are bounded for all real ξ if $-1 < r < 1$. $\mathrm{Ai}(\eta)$ is an L_2 function on $[\xi, \infty)$, for every real ξ. As shown in Appendix B, it then follows that $\widetilde{A}_\xi$ is bounded on $L_2([\xi,\infty))$:

$$a(\xi) \equiv |||\widetilde{A}_\xi|||_{L_2([\xi,\infty))} \equiv \sup_{f \in L_2([\xi,\infty))} \frac{\|A_\xi f\|_2}{\|f\|_2} \leqslant 1, \tag{3.24}$$

with $a(\xi) \to 0$ as $\xi \to +\infty$, $a(\xi) \to 1$ as $\xi \to -\infty$, and $a(\xi)$ monotonically decreasing for ξ real. Thus, for $-1 < r < 1$,

$$|||r^2\widetilde{A}_\xi|||_{L_2} \leqslant r^2 < 1,$$

$(I - r^2\widetilde{A}_\xi)$ is invertible for all real ξ, and

$$|||(I - r^2\widetilde{A}_\xi)^{-1}|||_{L_2} \leqslant (1 - r^2)^{-1}. \tag{3.25}$$

(iv) For $\sigma = +1$, $r > 1$, there is a real ξ_0 such that $|||\widetilde{A}_{\xi_0}|||_{L_2} = 1/r^2$. The solution of Eq. (3.22) that vanishes as $\xi \to \infty$ is finite for real $\xi > \xi_0$. The result also follows from (3.25). We suspect, but have not proven, that $Q(\xi)$ has a singularity (and therefore a pole) at ξ_0.

(v) In the general case, $C_1 \neq 0$ in (3.19) and

$$F(\xi) \sim C_1/\xi, \quad \text{as } \xi \to \infty. \tag{3.26}$$

As shown in Appendix A, this decay rate is enough to establish that $(I - \sigma A_x)$ is invertible for x large enough. Then (3.18) provides a family of solutions of (3.21).

It may be worth mentioning that it was not necessary to obtain (3.21) via (3.16). The operator $[I + \xi\partial_\xi + \eta\partial_\eta - (\partial_\xi + \partial_\eta)^3]$ when applied to (3.18) leads directly to (3.21). However, the algebra seems slightly simpler by the present route.

B. Sine-Gordon equation and Painlevé III

The appropriate integral equation is still (3.2), but now F satisfies

$$\mathscr{L}_2 F = -\partial_x\partial_t F(x,t) + F(x,t) = 0. \tag{3.27}$$

The dictionary constructed above still applies, but two additions are needed. First, from (3.4),

$$\begin{aligned}(I - \sigma A_x)\partial_t K_2(x,z) = \partial_t \int_0^\infty F\left(\frac{2x+\zeta}{2}\right) F\left(\frac{x+\zeta+z}{2}\right) d\zeta \\ + \frac{\sigma}{4}\int^\infty \int_0 K_2(x,x+\zeta)\partial_t \\ \times \left\{F\left(\frac{2x+\zeta+\eta}{2}\right) F\left(\frac{x+\eta+z}{2}\right)\right\} d\zeta\, d\eta.\end{aligned} \tag{3.28}$$

Second, it follows from (3.27) that

$$\begin{aligned}F\left(\frac{2x+\zeta+\eta}{2}\right)\left[\partial_t(\partial_x + \partial_y) F\left(\frac{x+\eta+y}{2}\right)\right] &= F\left(\frac{2x+\zeta+\eta}{2}\right) F\left(\frac{x+\eta+y}{2}\right) \\ &= \left[\partial_t(\partial_x + \partial_\eta) F\left(\frac{2x+\zeta+\eta}{2}\right)\right] F\left(\frac{x+\eta+y}{2}\right).\end{aligned} \tag{3.29}$$

Now apply $\mathscr{L}_2 = [I - \partial_t(\partial_x + \partial_y)]$ to (3.2). Suppressing all arguments, the result may be written schematically as

$$\begin{aligned}\mathscr{L}_2 K = 0 + \frac{\sigma}{4}\int\int (\mathscr{L}_2 K) FF - \frac{\sigma}{4}\int\int (\partial_t K)(\partial_x + \partial_y)(FF) - \frac{\sigma}{4}\int\int [(\partial_1 - \partial_2)K]\partial_t(FF) \\ - \frac{\sigma}{2}\int\int (\partial_2 K)\partial_t(FF) - \frac{\sigma}{4}\int\int K\partial_t[\{(\partial_x + \partial_y)F\}F] - \frac{\sigma}{4}\int\int K\partial_t\{F(\partial_x + \partial_y)F\}.\end{aligned}$$

Use (3.10) to integrate (once) the second nonzero term on the right. Use (3.8) to simplify the next term. The next two terms may be combined and then integrated once. Use (3.29) to simplify the last term, and it may be integrated as well. The results may be combined and written as

$$(I - \sigma A_x)\mathscr{L} K(x,y) = \frac{\sigma}{2} F\left(\frac{x+y}{2}\right)\partial_t K_2(x,x) + \frac{\sigma}{2} K(x,x)\left[\frac{\sigma}{4}\int\int K_2 \partial_t(FF) + \partial_t \int FF\right].$$

Then using (3.2) and (3.28) and the invertibility of $(I - \sigma A_x)$, we obtain

$$[I - \partial_t(\partial_x + \partial_y)]K(x,y) = \frac{\sigma}{2}[\partial_t K_2(x,x)]K(x,y) + \frac{\sigma}{2}K(x,x)[\partial_t K_2(x,y)]. \tag{3.30}$$

On $y = x$, set

$$q(x,t) = K(x,x;t), \qquad R(x,t) = K_2(x,x;t), \tag{3.31}$$

and (3.30) becomes

$$q - q_{xt} = \sigma q R_t. \tag{3.32}$$

Also on $y = x$, (3.8a) becomes

$$R_x = -2q^2. \tag{3.33}$$

These are the same equations generated by the IST approach.[2] If desired, they can be combined into a single integrodifferential equation for q. They may be related to the sine-(sinh-) Gordon equation through a somewhat obscure transformation. Define

$$q = \frac{\sigma}{2} u_x(x,t). \tag{3.34}$$

If $\sigma = +1$, set

$$R = \sigma \int^t (1 - \cosh u) d\tau \tag{3.35}$$

and note that (3.32) and (3.33) are both satisfied if

$$u_{xt} = \sinh u. \tag{3.36}$$

Similarly, if $\sigma = -1$, set

$$R = \sigma \int^t (1 - \cos u) d\tau \tag{3.37}$$

and find that

$$u_{xt} = \sin u. \tag{3.38}$$

We emphasize again that the solutions of (3.36) and (3.38) obtained via (3.2) need not satisfy any of the global $(-\infty < x < \infty)$ requirements inherent in IST.

Now let us consider the self-similar form of these results. Let

$$F(x,t) = t\mathscr{F}(\xi), \tag{3.39}$$

where $\xi = xt$. It follows from (3.27) that

$$\xi\mathscr{F}'' + 2\mathscr{F}' - \mathscr{F} = 0. \tag{3.40}$$

For $\xi > 0$, if we set

$$\xi = (\rho/2)^2, \quad \mathscr{F}(\xi) = \rho^{-1} g(\rho),$$

then (3.40) becomes Bessel's equation:

$$\rho^2 g'' + \rho g' - (\rho^2 + 1) g = 0. \tag{3.41}$$

For the linear integral equation, we will need the solution of (3.41) that decays as $\rho \to \infty$, so that

$$\mathscr{F}(\xi) \sim C\xi^{-3/4} \exp(-2\xi^{1/2}), \quad \xi \to \infty. \tag{3.42}$$

Corresponding to (3.39), set

$$K(x,y,t) = t\mathscr{K}(\xi,\eta), \tag{3.43}$$

where $\eta = yt$, and (3.2) becomes

$$\begin{aligned}\mathscr{K}(\xi,\eta) = \mathscr{F}\left(\frac{\xi+\eta}{2}\right) + \frac{\sigma}{4}\iint_0^\infty &\mathscr{K}(\xi,\xi+\zeta)\\ \times &\mathscr{F}\left(\frac{2\xi+\zeta+\psi}{2}\right)\\ \times &\mathscr{F}\left(\frac{\xi+\psi+\eta}{2}\right) d\zeta\, d\psi, \quad \eta \geqslant 0.\end{aligned} \tag{3.44}$$

With $\mathscr{F}$ defined by (3.40) and (3.42), $\mathscr{K}$ is then defined by Eq. (3.44). Let

$$\begin{aligned}&K_2(x,y;t) = t\mathscr{K}_2(\xi,\eta), \mathscr{R}(\xi) = \mathscr{K}_2(\xi,\xi),\\ &Q(\xi) = \mathscr{K}(\xi,\xi), U(\xi) = u(x,t).\end{aligned} \tag{3.45}$$

Then (3.32) and (3.33) become, respectively,

$$\begin{aligned}&\xi Q'' + 2Q' - Q + \sigma Q(\mathscr{R}\xi)' = 0,\\ &R' + 2Q^2 = 0.\end{aligned} \tag{3.46}$$

If desired, these can be combined into a single third-order equation for Q. In any case, every solution of the linear integral equation (3.44) also provides a solution of this third-order system of nonlinear ODE's.

To obtain the third equation of Painlevé (P_{III}), set

$$Q(\xi) = \frac{\sigma}{2} U'(\xi) = \frac{\sqrt{\sigma}}{2} (\ln W(\xi))'. \tag{3.47}$$

Then either (3.36) or (3.38) yields P_{III}:

$$W'' = \frac{1}{W}(W')^2 - \frac{1}{\xi} W' + \frac{1}{2\xi}(W^2 - 1), \tag{3.48}$$

and solutions of this equation also can be obtained via (3.44). Two immediate conclusions about these solutions of (3.48) are the following:

(i) The solutions of (3.48) that decay as $\xi \to \infty$ may have a fixed branch point at $\xi = 0$, but have no movable singularities other than poles. The possible branch point at $\xi = 0$ is evident from (3.48). As shown in Ref. 6 (cf. Sec. 4), the decay rate in (3.42) insures that Fredholm theory applies to (3.44), and its solution can have only poles, apart from whatever fixed singularities $\mathscr{F}$ may have (at $\xi = 0$ in this case). Thus, Q has only (simple) poles, and from (3.47) this is true of W as well. (This result, of course, is consistent with that of Painlevé.[11]) However, it is *not* true of $U(\xi)$, which has movable logarithmic branch points. This shows an important aspect of the conjecture formulated in Ref. 5. The nonlinear PDE's solved by IST are (3.32) and (3.33), and their similarity solutions (i.e., Q) have the Painlevé property. The sine-Gordon equation is a consequence of (3.32) and (3.33); its similarity solution (i.e., U, which is the integral of Q) has movable logarithmic branch points which were introduced by (3.47).

(ii) There is a ξ_0 large enough that Eq. (3.44) has a convergent Neumann series for $\eta \geqslant \xi > \xi_0$:

$$\mathscr{K} = \sum_{n=0}^{\infty} (\sigma A_\xi)^n \mathscr{F}. \tag{3.49}$$

Naturally, this also dictates the behavior of $W(\xi)$ for $\xi > \xi_0$. ξ_0 must be large enough that $|||A_\xi|||_{L_2} < 1$ for $\xi > \xi_0$. The rapid decay given by (3.42) makes it fairly easy to find an adequate ξ_0. This result was first obtained in Ref. 13; the proof given there required 133 numbered equations and 22 figures! [To put their work in perspective, we should also note that they obtained global connection formulas for (3.48) by their approach.]

C. Derivative nonlinear Schrödinger equation and Painlevé IV

Consider a different linear integral equation[14]:

$$\begin{aligned}K(x,y) = &F^*(x+y)\\ &+ i\iint_x^\infty K(x,s) F'(s+z) F^*(z+y) dz\, ds.\end{aligned} \tag{3.50}$$

Here we have already assumed that $\mathscr{L}_1 F = (\partial_x - \partial_y) F = 0$. From this we must develop a new dictionary. Define

$$K_2^*(x,z) = i\int_x^\infty K(x,s) F'(s+z) ds. \tag{3.51}$$

Then one may show that

$$\begin{aligned}&(\partial_x - \partial_y) K(x,y)\\ &\quad = -2K(x,x) K_2(x,y) + 2i|K(x,x)|^2 K(x,y),\end{aligned} \tag{3.52}$$

$$(\partial_x + \partial_y)K_2^*(x,y) = 2i|K(x,x)|^2K_2^*(x,y) - 2iK(x,x)\,\partial_y K^*(x,y).$$

Let $F(x+y;t)$ also satisfy

$$\mathscr{L}_2 F = i\partial_t F + 4\,\partial_x^2 F = 0. \tag{3.53}$$

Then one may show, using the present approach, that the solution of (3.50) satisfies

$$\begin{aligned} i\partial_t K(x,y) &= (\partial_x + \partial_y)^2 K(x,y) \\ &\quad -4iK(x,x)\,[\partial_x K^*(x,x)] \\ &\quad \times K(x,y) + 8|K(x,x)|^4 K(x,y). \end{aligned} \tag{3.54}$$

If $q(x) = K(x,x)$, then

$$i\partial_t q = \partial_x^2 q - 4iq^2 q_x^* + 8|q|^4 q. \tag{3.55}$$

The similarity form is in this case

$$F(x,t) = (2t)^{-1/4}\mathscr{F}(\xi), \quad \xi = x/(2t)^{1/2}.$$

Then if we define

$$K(x,y;t) = (2t)^{-1/4}\mathscr{K}(\xi,\eta), \quad \eta = y/(2t)^{1/2},$$

$$q(x,t) = (2t)^{-1/4}Q(\xi),$$

then (3.55) becomes

$$-i\xi Q' - i\frac{Q}{2} = Q'' - 4iQ^2(Q^*)' + 8|Q|^4 Q. \tag{3.56}$$

This equation can be transformed into the fourth Painlevé transcendent (P_{IV}) by writing

$$Q = \rho e^{i\theta},$$

$$\rho\theta'' + 2\rho'\theta' - 4\rho^2\rho' + \xi\rho' + \frac{\rho}{2} = 0, \tag{3.57}$$

$$\rho'' + \rho\theta'^2 - 4\rho^3\theta' + 8\rho^5 - \xi\rho\theta' = 0.$$

The first equation can be integrated once after multiplication by ρ:

$$\rho^2\theta' - \rho^4 + \xi\rho^2/2 + C = 0,$$

which gives θ' in terms of ρ^2. Then defining $u = \rho^2$ and substituting for its expression in terms of u,

$$u'' - \frac{(u')^2}{2u} + 6u^3 + 4u^2\xi + \left(12C + \frac{\xi^2}{2}\right)u - \frac{2C^2}{u} = 0. \tag{3.58}$$

This equation defines the fourth Painlevé transcendent, after scaling u and ξ. Thus, one more inverse scattering problem has been shown to be connected with an equation of P-type.

4. INVERTIBILITY

The method presented here relies fundamentally on our being able to invert $(I - A_x)$. In this section we discuss some methods to prove invertibility.

There are cases where the power of Fredholm theory can be used to explicitly invert the operator $(I - A_x)$ anywhere in the complex plane except for fixed cuts and movable isolated points x_i.[6] Moreover, one can prove that the movable singularities at the x_i's are poles.

The conditions for applicability of Fredholm theory are the existence of a region along the positive real axis, say a strip of finite width, such that for any path C with x as one end point and going to infinity in this region, the following hold:

$$|N(x;y,z)| < M(x;y) \text{ for all } z$$

for some M such that $\oint M(x;y)\,dy < +\infty$,

$$\oint F(x+y)M(x;y)dy < +\infty.$$

This is satisfied in many cases where the decay of F (and thus N) is exponential provided its argument goes to $+\infty$ in a strip. For example, this was the case in Sec. 3 in example (A) if $C_1 = 0$ in (3.19), and in example (B).

There are cases, however, where these conditions are not satisfied, but where it is still possible to prove that $(I - A_x)$ is invertible in some L_p space to which F belongs. For instance, $(I - A_x)$ is certainly invertible, $(I - A_x)^{-1}$ bounded, if $(-A_x)$ is a positive operator. In case (B) of Sec. 2, if only real functions are considered, and in case (B′) for complex functions, provided the minus sign is taken in the definition of N, it is enough to assume that $xF(x)$ remains bounded as $x \to +\infty$ in order to prove that $(-A_x)$ exists and is positive on L_2 (Appendix A). Then $(I - A_x)$ is invertible in L_2, and since the boundedness of $xF(x)$ implies that F is in L_2 as well, it follows that the solution of Eq. (2.1) in these cases exists for all x.

If one can find an upper bound to A_x in some L_p space to which F belongs, one can again prove something about the invertibility of $(I - A_x)$. We must remember here that F satisfies some *linear* differential equation and that it is always possible to multiply F by some number λ. Of course, the solution of (2.1) is not just multiplied by λ since it is not linear in F.

Let us choose one solution F of the relevant linear equation. If there exists some L_p space that contains F and on which A_x is bounded, then the function $a(x)$, defined by (3.24), is certainly a monotonic nonincreasing function of x and $\lim_{x\to\infty} a(x) = 0$. This will be true in cases (A), (B), (B′), and (D) for $p = 2$ provided $xF(x)$ remains bounded as $x \to +\infty$, and in case (C) for $p = 4$ provided $x^{1/2}F(x)$ and $x^{3/2}F'(x)$ remain bounded as $x \to +\infty$ (Appendix A). There are cases where

$$\lim_{x\to-\infty} a(x) = C < +\infty \quad \text{(Appendix B)}.$$

Then, for all λ such that $|\lambda| < C^{-1}$ the operator $(I - \lambda A_x)$ is invertible in the appropriate L_p space, for all x, and $(I - \lambda A_x)^{-1}$ is bounded by $(1 - |\lambda|C)^{-1}$. The solution of (2.1) where F is replaced by $\lambda^{1/r}F$ [$r = 1$ in case (A); 2 in cases (B), (B′), and (C); 3 in case (D)] exists for all values of the argument. Thus, we are again in a case where we find global existence of the solution with a restriction on F. If $(-A_x)$ is positive, the solution of (2.1) exists for all decaying F's.

If $\lim_{x\to-\infty} a(x) = +\infty$, however small λ is, we cannot prove global existence by this approach. However, define

$$x_0 = \sup\{x \,|\, a(x) \geqslant 1/|\lambda|\}.$$

Then for all $x > x_0$, $(I - \lambda A_x)$ is invertible. Thus, for all λ one can find some interval (x_0, ∞) on which $(I - \lambda A_x)^{-1}$ exists and is bounded, and therefore continuous.

ACKNOWLEDGMENTS

We gratefully acknowledge conversations with Professor Elliott Lieb, which were especially useful in obtaining the results in Appendix A. We are also indebted to Dr. J. Satsuma for many useful discussion at very early stages of this work. This work was partially supported by the U.S. Army Research Office, the Air Force Office of Scientific Research, and the National Science Foundation.

APPENDIX A

We will prove that provided $xF(x)$ remains bounded as $x \to +\infty$, the operator B_x defined on L_2 by

$$(B_x g)(y) = \int_x^\infty g(u)F(u+y)du, \quad y \geqslant x,$$
$$= 0, \quad y < x,$$

exists and is bounded.

This result on B_x will give us information on cases (A), (B), (B'), and (D) since the relevant operators A_x as defined in Sec. 2 are, respectively, B_x, σB_x^2, $\sigma B_x^\dagger B_x$, and B_x^3, where $B_x^\dagger$ is the adjoint of B_x. This is enough to prove that $(-A_x)$ is positive in cases (B) and (B') if only real functions are considered. It also gives an upper bound to $a(x) = |||A_x|||_{L_2}$.

The idea is to prove that $B_x g$ is in L_2 by bounding it and using Fubini's theorem to show that the integral

$$J = \int_{-\infty}^{+\infty} dy (B_x g)^*(y)(B_x g)(y)$$

exists and is bounded.

We assume that $xF(x)$ remains bounded as $x \to +\infty$. More precisely, this means that there exists some x_0 and M such that

$$|F(z)| \leqslant \frac{M}{z - 2x_0}, \quad \text{for} \quad z > 2x_0. \tag{A1}$$

Note that in (2.1) the argument of F is never smaller than $2x$. From now on, we take $x \geqslant x_0$.

Consider the integral

$$I = \int_x^\infty \int_x^\infty \int_x^\infty |g(y)||g(z)||F(y+u)|\,|F(z+u)|\,du\,dy\,dz. \tag{A2}$$

The convergence of this integral implies that the integral J is absolutely convergent and $J \leqslant I$. However, the integrand of I is positive and by Fubini's theorem it is enough to prove integrability for some choice of the orders of integration. In Eq. (A2) integrate in u first, using relation (A1):

$$I \leqslant M^2 \int_x^\infty \int_x^\infty dz\,dy |g(y)||g(z)| \times \int_x^\infty \frac{du}{(z+u-2x_0)(y+u-2x_0)}. \tag{A3}$$

The last integral in (A3) is explicitly calculated:

$$\frac{\ln(z+x-2x_0) - \ln(y+x-2x_0)}{z-y}.$$

Changing variables by

$$e^r = z + x - 2x_0, \quad e^s = y + x - 2x_0,$$

gives

$$I \leqslant M^2 \int_{\ln 2(x-x_0)}^\infty e^r dr \int_{\ln 2(x-x_0)}^\infty e^s ds \times |g(e^r + 2x_0 - x)||g(e^s + 2x_0 - x)| \frac{r-s}{e^r - e^s}.$$

Define

$$\gamma(r) = e^{r/2} g(e^r + 2x_0 - x), \quad r \geqslant \ln 2(x - x_0),$$
$$= 0, \quad r < \ln 2(x - x_0). \tag{A4}$$

Then

$$\|\gamma\|_2^2 = \int_{-\infty}^{+\infty} |\gamma(r)|^2 dr = \int_{\ln 2(x-x_0)}^\infty dr\, e^r |g(e^r + 2x_0 - x)|^2$$
$$= \int_x^\infty |g(z)|^2 dz \leqslant \|g\|_2^2. \tag{A5}$$

Thus, $\gamma \in L_2$ and $\|\gamma\|_2 \leqslant \|g\|_2$. Then

$$I \leqslant M^2 \int_{\ln 2(x-x_0)}^\infty dr \int_{\ln 2(x-x_0)}^\infty ds \times |\gamma(r)||\gamma(s)| \frac{r-s}{e^{(r-s)/2} - e^{(s-r)/2}}.$$

Observe that the function

$$\zeta(t) = \frac{t}{2\sinh(t/2)}$$

is in L_1 and

$$\|\zeta\|_1 = \int_{-\infty}^{+\infty} \zeta(t) dt = \pi^2.$$

Since $\zeta \in L_1$ and $\gamma \in L_2$,

$$I \leqslant M^2 \int \int_{-\infty}^{+\infty} |\gamma(r)||\gamma(s)| \zeta(r-s) dr\,ds \leqslant M^2 \|\gamma\|_2^2 \|\zeta\|_1$$

by Young's inequality.

Thus, I exists. Since $J \leqslant I$, J is absolutely convergent. Hence,

$$(B_x g) \in L_2, \quad \|B_x g\|_2 \leqslant M\pi \|g\|_2,$$
$$\text{and} \quad |||B_x|||_{L_2} \leqslant M\pi \quad \text{for} \quad x \geqslant x_0.$$

however, M depends on x_0 and it is not generally possible to find a finite upper bound to $|||B_x|||$ for all x.

Nevertheless, this proves that in cases (A), (B), (B'), and (D), whenever $xF(x)$ remains bounded as $x \to +\infty$, the operator $(I - \lambda A_x)$ is invertible on (x_0, ∞) for some λ. The same can be proven in case (C) but the condition is now that $x^{1/2}F(x)$ and $x^{3/2}F'(x)$ remain bounded. To show this, suppose that

$$|F(z)| \leqslant \frac{M}{\sqrt{z - 2x_0}}, \quad |F'(z)| \leqslant \frac{N}{(z-2x_0)^{3/2}}, \quad \text{for } z > 2x_0. \tag{A6}$$

We will prove that the operator A_x exists and is bounded on L_4. For this we consider $g \in L_4$ and $h \in L_{4/3}$. We will again show by using Fubini's theorem that

$$(h, A_x g) = \int_{-\infty}^{+\infty} h^*(y) A_x g(y) dy$$

is absolutely convergent by bounding the integral

$$I=\int_x^{+\infty}\int_x^{+\infty}\int_x^{+\infty}|h(y)||g(z)||F'(z+u)|$$
$$|F(y+\omega)|du\,dy\,dz,$$
$$I\leqslant H=MN\int_x^{\infty}\int_x^{\infty}|h(y)||g(z)|$$
$$\times\int_x^{\infty}\frac{du}{(z+u-2x_0)^{3/2}(y+u-2x_0)^{1/2}}.$$

The idea is again to perform the u integration first, and to define new independent and dependent variables

$$x+y-2x_0=e^r,\quad x+z-2x_0=e^s,$$
$$\eta(t)=e^{3t/4}h(e^t+2x_0-x),\quad \gamma(t)=e^{t/4}g(e^t+2x_0-x),$$
$$t\geqslant \ln 2(x-x_0),$$
$$\eta(t)=0,\quad \gamma(t)=0,$$
$$t<\ln 2(x-x_0).$$

We now have

$\eta\in L_{4/3}$, $\|\eta\|_{4/3}\leqslant\|h\|_{4/3}$, $\gamma\in L_4$ and $\|\gamma\|_4\leqslant\|g\|_4$.

H can be rewritten as

$$H=MN\int_{\ln 2(x-x_0)}^{\infty}dr\int_{\ln 2(x-x_0)}^{\infty}ds|\eta(r)||\gamma(s)|$$
$$\times\frac{1}{2}\operatorname{sech}\left(\frac{r-s}{4}\right).$$

The function

$\beta(z)=\frac{1}{2}\operatorname{sech}(z/4)$ is in L_1 and $\|\beta\|_1=4\pi$.

Then by Young's inequality

$$H\leqslant MN\|\eta\|_{4/3}\|\gamma\|_4\|\beta\|_1,$$
$$J\leqslant I\leqslant H\leqslant 4\pi MN\|h\|_{4/3}\|g\|_4,$$
$$(h,A_xg)\leqslant 4\pi MN\|g\|_4,$$
$$\|A_xg\|_4=\sup_{g\in\mathscr{L}_4}\|h\|_{4/3}=1,$$

or

$$|||A_x|||_{L_4}<4\pi MN.$$

APPENDIX B

Let

$$A_xf(y)=\frac{1}{4}\int_x^{\infty}f(\xi)\mathrm{Ai}\left(\frac{\xi+\eta}{2}\right)\mathrm{Ai}\left(\frac{\eta+y}{2}\right)d\eta\,d\xi. \tag{B1}$$

We want to show that $a(x)\leqslant 1$ for all real x, where $a(x)$ is defined by (3.24). If $x<z$, then $L_2[x,\infty)\supset L_2[z,\infty)$. Therefore, $a(x)$ is a nonincreasing function of x, and $a(x)\to 0$ as $x\to\infty$. Therefore, it is suficient to show that $a(x)\leqslant 1$ as $x\to-\infty$.

Consider

$$B_xf(y)=\frac{1}{2}\int_x^{\infty}f(z)\mathrm{Ai}\left(\frac{z+y}{2}\right)dz,$$
$$B_xf(2q)=\int_{x/2}^{\infty}f(2p)\mathrm{Ai}(p+q)dp. \tag{B2}$$

We will show that $B_{-\infty}$ exists on $L_2(-\infty,\infty)$ and that $|||B_x|||_{L_2}\leqslant|||B_{-\infty}|||_{L_2}=1$. For $f\in L_2$, define

$$I(f)=(f,B_{-\infty}f)$$
$$=\frac{1}{2}\int_{-\infty}^{\infty}f^*(z)f(y)\mathrm{Ai}\left(\frac{z+y}{2}\right)dz\,dy. \tag{B3}$$

Because $B_{-\infty}$ is symmetric and real,

$$|||B|||_{L_2}=\sup_{f\in L_2}\frac{I(f)}{\|f\|_2}. \tag{B4}$$

In the sense of distributions, the Fourier transform of the Airy Function is $(1/\sqrt{2\pi})\exp(ik^3/3)$. This means that we can formally replace $\mathrm{Ai}(z+y/2)$ in I by

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}dk\frac{1}{\sqrt{2\pi}}e^{ik^3/3}e^{ik(z+y)/2}$$

and interchange the order of integration. Thus,

$$I=\frac{1}{2}\int_{-\infty}^{+\infty}dk\,e^{ik^3/3}\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}dz$$
$$\times\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}dyf^*(y)f(y)e^{ik(z+y)/2},$$
$$I=\frac{1}{2}\int_{-\infty}^{+\infty}dk\,e^{ik^3/3}\tilde{f}(-k/2)\tilde{f}^*(k/2)$$
$$\leqslant\frac{1}{2}\int_{-\infty}^{+\infty}dk\,|\tilde{f}(-k/2)||\tilde{f}(k/2)|,$$

where $\tilde{f}$ is the Fourier transform of f. By Young's inequality

$$I\leqslant\int_{-\infty}^{+\infty}\frac{dk}{2}|\tilde{f}(k/2)|^2=\|\tilde{f}\|^2=\|f\|_2^2,$$

the last equality being Parseval's identity. Equality is attained, for instance, for

$$\tilde{f}(k)=e^{+4ik^3/3}\phi(k),$$

with $\phi(k)$ real even, and $\in L_2$.

It follows from this last remark that

$$|||B_{-\infty}|||_{L_2}=1.$$

Now $A_x=B_x^2$; therefore

$$|||A_x|||_{L_2}=|||B_x^2|||_{L_2}\leqslant|||B_{-\infty}|||_{L_2}^2=1.$$

[1] V.E. Zakharov and A.B. Shabat, Funct. Anal. Appl. **8**, 226 (1974).
[2] M.J. Ablowitz, D.J. Kaup, A.C. Newell, and H. Segur, Studies Appl. Math. **53**, 249 (1974).
[3] M.J. Ablowitz, and H. Segur, Studies Appl. Math. **57**, 13 (1977).
[4] M.J. Ablowitz, and H. Segur, Phys. Rev. Lett. **38**, 1103 (1977).
[5] M.J. Ablowitz, A. Ramani, and H. Segur, Lett. Nuovo Cimento **23**, 333 (1978).
[6] M.J. Ablowitz, A. Ramani, and H. Segur, J. Math Phys. **21**, 715 (1980).
[7] R. Nakach, Technical Report No. 72, School of Electrical Engineering, Chalmers University of Technology, Göteborg, Sweden, 1977.
[8] L.D. Fadeev, J. Math. Phys. **4**, 72 (1963).
[9] P.D. Deift, Duke Math. J. **45**, 267 (1978).
[10] S.P. Hastings, and J.B. McLeod, MRC Report #1861 (1978).
[11] E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1944).
[12] J.W. Miles, Proc. R. Soc. (London) Ser. A **361**, 277 (1978).
[13] B.M. McCoy, C.A. Tracy, and T.T. Wu, J. Math. Phys. **18**, 1058 (1977).
[14] D.J. Kaup, and A.C. Newell, J. Math. Phys. **19**, 798 (1978).

A commutator representation of Painlevé equations

H. Flaschka

Department of Mathematics, The University of Arizona, Tucson, Arizona 85721

(Received 15 May 1979; accepted for publication 17 December 1979)

This paper further develops the connection between partial differential equations solvable by inverse scattering methods and ordinary differential equations of Painlevé type. The main result given here is that Painlevé equations have a concise algebraic formulation $[L,B] = L$ when written in the Lax representation.

1. INTRODUCTION

This note is a supplement to the preceding work[1] of Ablowitz, Ramani, and Segur (ARS). That paper deals with classes of Painlevé-type equations solvable by linear integral equations; the nonlinear o.d.e.'s are there derived by a "direct method" related to the Zakharov–Shabat technique for identifying integrable evolution equations.[2] Some time ago, while trying to understand the original paper of Ablowitz and Segur[3] on Painlevé equations and inverse scattering, I had observed that by scaling the time t out of the Marchenko equations of the general inverse problems treated by Zakharov and Shabat, one could generate and "solve" (in the sense of reduction to a Volterra equations) a whole series of nonlinear o.d.e.'s. The resulting method is a straightforward adaptation of the Zakharov–Shabat technique, and the Painlevé equations are simultaneously derived and linearized. This variant of the ARS ideas is perhaps of methodological interest in that it provides a different perspective on the material in earlier papers.[1,3-5] The real point, however, lies in the final algebraic formulation of the Painlevé-type equations; *they appear as a condition on the commutator of two differential operators, namely,*

$$[L,B] = L. \tag{1}$$

As example, take $L = D^2 + f\,(D = d/dx)$ and $B = D^3 + (\tfrac{3}{2}f + \tfrac{1}{2}x)\,D + \tfrac{3}{4}f'$. Then Eq. (1) becomes

$$f''' + 6ff' - 4f + 2xf' = 0,$$

which is the equation governing the self-similar solution $q(\xi,t) = (3t)^{-2/3}f(\xi(3t)^{-1/3})$ of KdV, $q_t + 3qq_\xi + \tfrac{1}{2}q_{\xi\xi\xi} = 0$.

It is well known that the general Zakharov–Shabat equation in one space variable can be written in the Lax representation

$$[L,A] + L_t = 0. \tag{2}$$

The stationary solutions of Eq. (2), which include the n-soliton and n-phase quasiperiodic solutions of Eq. (2), satisfy the commutator equation of the form

$$[L,B] = 0. \tag{3}$$

Equation (3) is essentially an algebraic equation (which makes sense even in the setting of differential operators over commutative rings). With some constraints on the orders of L and B, Eq. (3) implies the existence of a polynomial $Q(l,b)$ such that

$$Q(L,B) = 0,$$

and the explicit solution of Eq. (3) reduces to function theory on the algebraic curve $Q(l,b) = 0$. The case of arbitrary L,B leads to vector bundles of rank $\geqslant 1$ over compact Riemann surfaces.[6] It is intriguing that another algebraic condition $[L,B] = L$ appears to be intimately related to the Painlevé equations which, together with equations for elliptic functions [included among Eq. (3)] and certain elementary o.d.e.'s, essentially exhaust all known "well-behaved" nonlinear o.d.e.'s.

A hint of the function-theoretic significance of the equation $[L,B] = L$ can be gotten from the following considerations: On the one hand, differentiation of the eigenvalue equation $Lv = \zeta v$ with respect to ζ leads to

$$(L - \zeta)v_\zeta = v. \tag{4}$$

On the other hand, application of $[L,B] - L = 0$ to an eigenvector v gives

$$LBv - \zeta Bv - \zeta v = 0,$$

or

$$(L - \zeta)\left(\frac{1}{\zeta}Bv\right) = v. \tag{5}$$

It follows from Eqs. (4) and (5) that $(1/\zeta)Bv = v_\zeta - \tilde{v}$, where $\tilde{v}$ is an appropriate solution of $L\tilde{v} = \zeta\tilde{v}$. So $\zeta v_\zeta = Bv + \zeta\tilde{v}$. Typically, Bv can be re-expressed in terms of ζ, if one replace x derivatives of v by multiples of ζ according to $Lv = \zeta v$. The result is an equation governing the dependence of v on the eigenvalue parameter ζ, e.g.,

$$\zeta v_\zeta = B(\zeta)v, \tag{6}$$

where $B(\zeta)$ is some matrix *depending polynomially on* ζ. A forthcoming paper by Alan Newell and me[7] will discuss the relevance of certain systems of the form (6) to Painlevé equations. The structures which emerge are again very algebraic, and in fact reminiscent of the vector-bundle theory associated with $[L,B] = 0$. [The analysis in Ref. 7 centers on the behavior of solutions of Eq. (6) near irregular singular points.]

The only object of the present note, however, is to get as far as $[L,B] = L$. To shorten the discussion, which in any case is formal and straightforward, I will assume familiarity with ARS[1,4,5] and with the paper of Zakharov–Shabat.[2] It should be noted that some of the equations below in abstract form might reduce to "0 = 0". The Zakharov–Shabat method can impose too many interrelated conditions on the transformation kernels $K_\pm(x,y)$, the kernels $F(x,y)$ of the Vol-

 0022-2488/80/051016-03$01.00

terra equations, and on the unknowns themselves. Even if everything else is consistent, the analysis of the resulting Volterra equations may be highly nontrivial.[1,3,4,8] It is best to think of my variant of the ARS method as providing probable Painlevé-type equations, with more detailed investigations being required in specific examples.

Finally, it should be kept in mind that (as with inverse scattering in general) there will be various ways of deriving and linearizing Painlevé equations—some more efficient than others, depending on the example considered. For example, the technique of Newell[9,10] (modelled explicitly on inverse scattering) leads very quickly to families of nonlinear o.d.e.'s equivalent to the ones discussed here.

2. RESULTS

Following Ref. 2, consider an integral operator $\widehat{F}$:

$$(\widehat{F}\psi)(x) = \int_{-\infty}^{\infty} F(x,y)\psi(y)\,dy,$$

which admits a triangular factorization

$$I + \widehat{F} = (I + \widehat{K}_{+})^{-1}(I + \widehat{K}_{-}). \tag{7}$$

$\widehat{K}_{\pm}$ are upper and lower triangular Volterra operators. (In what follows, I will not distinguish notationally between operators $\widehat{F}$ and their kernels F). Zakharov–Shabat introduce (matrix) constant-coefficient linear operators L_0, B_0 of the form $L_0 = lD^n$, $B_0 = bD^m$ and require that

$$[L_0,B_0] = 0, \tag{8}$$

$$[F,L_0] = 0, \tag{9}$$

$$\left[F,B_0 + \frac{\partial}{\partial t}\right] = 0.$$

Time dependence is not relevant for the self-similar solutions of interest here, and the last condition is replaced by

$$[F,B_0 + T_0] = 0, \tag{10}$$

with T_0 denoting an operator whose coefficients will generally depend on x in a simple fashion.

Example 1: $L_0 = D^2$, $B_0 = D^3$, $T_0 = \frac{1}{2}xD$. Equation (8) is trivial, and Eqs. (9) and (10) impose conditions on the kernel $F(x,y)$:

$$(9) \Rightarrow F_{xx} - F_{yy} = 0, \tag{11}$$

$$(10) \Rightarrow F_{xxx} + F_{yyy} + \tfrac{1}{2}xF_x + \tfrac{1}{2}yF_y + \tfrac{1}{2}F = 0. \tag{12}$$

If, as is customary, we take the solution of Eq. (11) in the form $F = F(x + y)$, then with $x + y = \xi$, $' = d/d\xi$, Eq. (12) becomes $4F''' + \xi F' + F = 0$.

It follows from Eqs. (8)–(10) by a formal application of the Jacobi identity for commutators, that

$$[F,[L_0,T_0]] = 0. \tag{13}$$

This is a necessary condition on T_0, inasmuch as F is pretty much fixed by Eqs. (9) and (10). One can ensure Eq. (13) without imposing further requirements on F by choosing T_0 so that

$$[L_0,T_0] = L_0. \tag{14}$$

Then Eq. (13) is an automatic consequence of Eq. (9). We assume Eq. (14) to hold for the rest of the paper.

Example 1 (*continued*): $[D^2,\frac{1}{2}xD] = D^2$.

There are now two additional steps in the Zakharov–Shabat method.

Step 1: *Obtain perturbations L,B,T of the free operators L_0,B_0,T_0*. Remembering that F, and hence $K_{\pm}$, are fixed once and for all [subject to Eqs. (8)–(10)], we introduce L,B,T by the conditions

$$X(I + K_+) - (I + K_+)X_0 = \text{integral operator},$$

where $X = L$, B, or T, and $X_0 = L_0$, B_0, or T_0. This condition simply defines X in terms of K_+.

Example 1 (*continued*): From "$L(I + K_+) - (I + K_+)L_0 =$ integral operator" follows $L = D^2 + q$, where $q(x) = 2(d/dx)K_+(x,x)$. For B one finds

$$B = D^3 + \tfrac{3}{2}qD + \tfrac{3}{4}q' + r,$$

where

$$q(x) = 2\frac{d}{dx}\xi_0(x), \quad r(x) = \frac{3}{2}\frac{d}{dx}(\xi_1(x) + \xi_0(x)^2), \tag{15}$$

with

$$\xi_j(x) = \left(\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right)^j K_+(x,y)\bigg|_{x=y} \quad \text{(see Ref. 2).}$$

Finally, $T = T_0 = \frac{1}{2}xD$.

The map $X_0 \rightarrow X$ is linear; in particular, it will be important that

$$(B + T)(I + K_+) - (I + K_+)(B_0 + T_0) = \text{integral operator},$$

with precisely the B and T defined above. Just as in Ref. 2, one now has the following:

Theorem 1: *If Eqs.* (8)–(10) *hold, then*

$$L(I + K_{\pm}) - (I + K_{\pm})L_0 = 0, \tag{16}$$

$$(B + T)(I + K_{\pm}) - (I + K_{\pm})(B_0 + T_0) = 0. \tag{17}$$

This imposes conditions on $K_{\pm}(x,y)$:

Example 1 (*continued*): Equation (16) implies the familiar

$$\frac{\partial^2}{\partial x^2}K_{\pm} = \frac{\partial^2}{\partial y^2}K_{\pm} + q(x)K_{\pm} = 0, \quad x \lessgtr y.$$

This forces $r = 0$ in Eq. (15).[2]

Step 2: *integrability condition.*

Theorem 2: *Suppose that Eqs.* (8)–(10) *and* (14) *hold. Define $L,B + T$ as in Eq.* (15). *Then*

$$[L,B + T] = L. \tag{18}$$

As explained in the Introduction, this is the main formula of this paper. The proof of Theorem 2 differs a little from the proof of the corresponding Theorem 2 in Ref. 2, but it is still tedious and so is relegated to an Appendix. We now turn to further examples.

3. EXAMPLES

Example 2: $L_0 = lD$, with $l = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $B_0 = -4D^3$, $T_0 = xD$ (times the 2×2 identity, of course). Then $L = lD + \left(\begin{smallmatrix} 0 & q \\ r & 0 \end{smallmatrix}\right)$. The requirement $r = q$ turns out to be consistent. Then $B = -4D^3 + 6\left[\begin{smallmatrix} q^2 & q' \\ q' & q^2 \end{smallmatrix}\right]D + \left[\begin{smallmatrix} 6qq' & 3q'' \\ 3q'' & 6qq' \end{smallmatrix}\right]$ and $T = xD$. Finally, $[L,B + T] = L$ reduces to

$$q''' = 6q^2q' = (xq)', \text{ or } q'' = 2q^3 + xq + c. \tag{19}$$

Equation (10) is the second Painlevé equation,[1,4] which governs the self-similar solution of modified KdV. The Marchenko kernel $F(x,y)$, by Eqs. (9) and (10) satisfies

$$lF_x + F_y l = 0, \tag{20}$$

$$-4(F_{xxx} + F_{yyy}) + xF_x + yF_y + F = 0. \tag{21}$$

For the entries F_{ij} of the 2×2 matrix F one finds

$$F_{11_x} + F_{11_y} = 0 \text{ from Eq. (20)},$$

so that $F_{11} = F_{11}(x-y)$, which turns Eq. (21) into

$$\xi F'_{11} + F_{11} = 0$$

($\xi = x - y$, $' = d/d\xi$). $F_{11} \equiv 0$ is the nonsingular solution. Next,

$$-F_{12_x} + F_{12_y} = 0,$$

so $F_{12} = F_{12}(x+y)$. With $x + y = \xi$ and $' = d/d\xi$,

$$-8F'''_{12} + \xi F'_{12} + F_{12} = 0.$$

The solution decaying at $x = +\infty$ is the Airy function $\mathrm{Ai}(\xi/2)$. At this point, the study of Eq. (19) can be continued along the lines of Refs. 3–5 and 8.

Example 3: $L_0 = D^3$, $B_0 = D^2$, $T_0 = \frac{1}{3}xD$. Then $L = D^3 + \frac{3}{2}uD + \frac{3}{4}u' + w$, $B = D^2 + u$, $T = \frac{1}{3}xD$ (compare Example 1). The resulting equation $[L, B + T] = L$ is somewhat peculiar:

$$u'''' + 3(u^2)'' + \tfrac{8}{3}u + \tfrac{7}{3}xu' + \tfrac{1}{3}x^2u'' = 0.$$

As far as I know, this is *not* related to any self-similar behavior of the Boussinesq equation. (The Boussinesq equation is obtained from this L_0, B_0 in Ref. 2.

APPENDIX: HOW TO PROVE THEOREM 2

By definition and by Theorem 1,

$$(B + T)(I + K) = (I + K)(B_0 + T_0), \tag{A1}$$

$$L(I + K) = (I + K)L_0, \tag{A2}$$

where $K = K_+$ or K_-. Multiply Eq. (A2) first on the left by T_0, then again on the right by T_0, and subtract:

$$[T_0, L](I + K) + L[T_0, K] = (I + K)[T_0, L_0] + [T_0, K]L_0. \tag{A3}$$

Rewrite Eq. (A1) as

$$[T_0, K] + (T - T_0)(I + K) + B(I + K) = (I + K)B_0. \tag{A4}$$

Multiply Eq. (A4) on the left by L, and again on the right by L_0. Solve these expressions for $L[T_0, K]$ and $[T_0, K]L_0$ and substitute into Eq. (A3). To the resulting identity add

$$BL(I + K) = B(I + K)L_0,$$

subtract

$$L(I + K)B_0 = (I + K)L_0B_0,$$

and use $[L_0, B_0] = 0$, to get

$$\{[T_0, L] + L(T_0 - T) + [B, L]\}(I + K) = (I + K)[T_0, L_0] + (T_0 - T)(I + K)L_0.$$

Now from Eq. (A2),

$$(T_0 - T)(I + K)L_0 = (T_0 - T)L(I + K),$$

so

$$\begin{aligned}([T_0, L] + [L, T_0 - T] + [B, L])(I + K) &= (I + K)[T_0, L_0]\\ &= (I + K)L_0\\ &= -L(I + K).\end{aligned}$$

Multiply on the right by the operator $(I + K)^{-1}$, to get

$$[T_0, L] + [L, T_0 - T] + [B, L] + L = 0.$$

This simplifies to $[L, B + T] = L$, as required.

ACKNOWLEDGMENTS

This work was supported in part by NSF Grant MPS75-07530 and U.S. Army contract DAAG29-78-G-0059 at the University of Arizona, and completed while I was on sabbatical leave at Clarkson College. I wish to thank Alan Newell for many useful discussions on this topic, and Mark Ablowitz for helpful explanations (and for convincing me that the above observations might be worth writing up).

[1] M.J. Ablowitz, A. Ramani, and H. Segur, J. Math. Phys. **21**, xxx (1980).

[2] V.E. Zakharov and A.B. Shabat, Funkt. Anal. Prilozhen. **8**, 43 (1974).

[3] M.J. Ablowitz and H. Segur, Phys. Rev. Lett. **38**, 1103 (1977).

[4] M.J. Ablowitz, A. Ramani, and H. Segur, J. Math. Phys. **21**, 715 (1980).

[5] M.J. Ablowitz, A. Ramani, and H. Segur, Lett. Nuovo Cimento **23**, 333 (1978).

[6] I.M. Krichever, Funkt. Anal. Prilozhen. **12**, 20 (1978).

[7] H. Flaschka and A.C. Newell, "Monodromy-preserving Deformations, I" (to appear in Commun. Math. Phys.).

[8] S.P. Hastings and J.B. McLeod, MRC Report 1861 (1978).

[9] A.C. Newell, "The General Structure of Integrable Evolution Equations." Proc. R. Soc. (London) Ser. A **365**, 283 (1979).

[10] A.C. Newell, "Further Connections Between Integrable Partial Differential Equations and Nonlinear Ordinary Differential Equations" (unpublished manuscript).

On the remarkable nonlinear diffusion equation $(\partial/\partial x)[a(u+b)^{-2}(\partial u/\partial x)] - (\partial u/\partial t) = 0$

George Bluman and Sukeyuki Kumei
Department of Mathematics and Institute of Applied Mathematics and Statistics, University of British Columbia, Vancouver, British Columbia, Canada V6T1W5

(Received 18 July 1979; accepted for publication 14 December 1979)

We study the invariance properties (in the sense of Lie–Bäcklund groups) of the nonlinear diffusion equation $(\partial/\partial x)[C(u)(\partial u/\partial x)] - (\partial u/\partial t) = 0$. We show that an infinite number of one-parameter Lie–Bäcklund groups are admitted if and only if the conductivity $C(u) = a(u+b)^{-2}$. In this special case a one-to-one transformation maps such an equation into the linear diffusion equation with constant conductivity, $(\partial^2\bar{u}/\partial\bar{x}^2) - (\partial\bar{u}/\partial\bar{t}) = 0$. We show some interesting properties of this mapping for the solution of boundary value problems.

1. INTRODUCTION

In recent years nonlinear diffusion processes described by the partial differential equation (p.d.e)

$$\frac{\partial}{\partial x}\left[C(u)\frac{\partial u}{\partial x}\right] - \frac{\partial u}{\partial t} = 0, \tag{1}$$

with a variable conductivity $C(u)$, have appeared in problems related to plasma and solid state physics.[1,2] Interest in such processes has long occurred in other fields such as metallurgy and polymer science.[3–5]

Some exact solutions are well known for such equations.[6] These can be shown to be included in the class of all similarity solutions to such equations obtained from invariance under a Lie group of point transformations.[7,8]

Recently, it has been shown that differential equations can be invariant under continuous group transformations beyond point or contact transformation Lie groups which act on a finite dimensional space.[9] These new continuous group transformations act on an infinite dimensional space. Such infinite dimensional contact transformations have been called Noether transformations[10] or Lie–Bäcklund (LB) transformations[11] (Noether mentioned the possibility of such transformations in her celebrated paper on conservation laws[12]). Well known nonlinear partial differential equations admitting LB transformations include the Korteweg–deVries,[13,14] sine-Gordon,[10,15] cubic Schrödinger,[14] and Burgers' equations.[16] All of these known examples admit an infinite number of one-parameter LB transformations. Moreover, many of their important properties (existence of an infinite number of conservation laws,[13,14] existence of solitons,[14] and existence[17] of Bäcklund transformations[18]) are related to their invariance under LB transformations.

Any linear differential equation which admits a nontrivial one-parameter point Lie group is invariant under an infinite number of one-parameter LB transformations through superposition. Moreover, every known nonlinear p.d.e., invariant under LB transformations, can be associated with some corresponding linear p.d.e.

With the above views in mind we study the invariance properties of Eq. (1). Previously,[7,8,19] it had been shown that Eq. (1) is invariant under

a) a three-parameter point Lie group for arbitrary $C(u)$,
b) a four-parameter point Lie group if $C(u) = a\cdot(u+b)^{\nu}$,
c) a five-parameter point Lie group if $\nu = -\frac{4}{3}$.

[It is well known that a six-parameter point Lie group leaves invariant Eq. (1) in the case $C(u) = \text{const.}$[20]]

In the present work, we show that Eq. (1) *is invariant under LB transformations if and only if the conductivity is of the form*

$$C(u) = a\cdot(u+b)^{-2}, \tag{2}$$

i.e., *if* Eq. (1) *is of the form*

$$\frac{\partial}{\partial x}\left[a\cdot(u+b)^{-2}\frac{\partial u}{\partial x}\right] - \frac{\partial u}{\partial t} = 0. \tag{3}$$

Furthermore, this equation admits an infinite number of LB transformations.

In this special case, there exists a one-to-one transformation which maps Eq. (3) *into the linear diffusion equation with constant conductivity, namely, the heat equation*

$$\frac{\partial^2\bar{u}}{\partial\bar{x}^2} - \frac{\partial\bar{u}}{\partial\bar{t}} = 0. \tag{4}$$

In the course of this paper, we find an operator connecting two infinitesimal LB transformations leaving Eq. (3) invariant. We prove that this operator is a recursion operator which generates an infinite sequence of one-parameter infinitesimal LB transformations leaving Eq. (3) invariant. Moreover, we show that no other LB transformation leaves Eq. (3) invariant.

By examining the linearization of Eq. (3), we are led to construct the transformation mapping Eq. (3) into Eq. (4). It is shown that this transformation maps the recursion operator of Eq. (3) into the spatial translation operator of Eq. (4), giving a simple interpretation of the transformation relating Eq. (3) to Eq. (4). We use this transformation to connect boundary value problems of Eq. (3) to those of Eq. (4).

We construct a new similarity solution of Eq. (3) corresponding to invariance under LB transformations.

 0022-2488/80/051019-05$01.00

2. DERIVATION OF THE CLASS OF NONLINEAR DIFFUSION EQUATIONS INVARIANT UNDER LB TRANSFORMATIONS

LB transformations include Lie groups of point transformations and finite dimensional contact transformations.[11] The algorithm for calculating infinitesimal LB transformations leaving differential equations invariant is essentially the same as Lie's method[8] for calculating infinitesimal point groups.

Consider the most general one-parameter infinitesimal LB transformation that can leave invariant a time-evolution equation,[21] namely;

$$u^* = u + \epsilon U(x,t,u,u_1,\ldots,u_n) + O(\epsilon^2),$$
$$x^* = x, \tag{5}$$
$$t^* = t,$$

where $u_i = \partial^i u/\partial x^i$, $i = 1,2,\cdots$. Let $\partial u/\partial t = u_t$, $\partial u_i/\partial t = u_{it}$, $\partial U/\partial u = U_0$, $\partial U/\partial u_i = U_i$, $\partial^2 U/\partial u_i \partial u_j = U_{i,j}$, $C' = dC/du$, and $C'' = d^2C/du^2$.

In the above notation Eq. (1) becomes

$$u_t = C'(u_1)^2 + Cu_2. \tag{6}$$

Under Eqs. (5) the derivatives of u appearing in Eq. (6) transform as follows:

$$(u_t)^* = u_t + \epsilon U^t + O(\epsilon^2),$$
$$(u_1)^* = u_1 + \epsilon U^x + O(\epsilon^2),$$
$$(u_2)^* = u_2 + \epsilon U^{xx} + O(\epsilon^2),$$

where

$$U^t = D_t U = \frac{\partial U}{\partial t} + U_0 u_t + \sum_{i=1}^{n} U_i u_{it}, \tag{7}$$
$$U^x = D_x U = \frac{\partial U}{\partial x} + \sum_{i=0}^{n} U_i u_{i+1},$$
$$U^{xx} = (D_x)^2 U = \frac{\partial^2 U}{\partial x^2} + 2\sum_{i=0}^{n} \frac{\partial U_i}{\partial x} u_{i+1} + \sum_{i,j=0}^{n} U_{i,j} u_{i+1} u_{j+1} + \sum_{i=0}^{n} U_i u_{i+2}.$$

D_t and D_x are total derivative operators with respect to t and x, respectively.

The transformation (5) is said to leave Eq. (6) invariant if and only if for every solution $u = \theta(x,t)$ of Eq. (6)

$$U^t = C''U(u_1)^2 + 2C'U^x u_1 + C'Uu_2 + CU^{xx}. \tag{8}$$

The fact that U must satisfy Eq. (8) for any solution of Eq. (6) imposes severe restrictions on U. Using Eq. (6) the derivatives of u_i with respect to t, i.e., u_{it}, can be eliminated in Eq. (8). Since the invariance condition (8) must hold for every solution of Eq. (6), Eq. (8) becomes a polynomial form in u_{n+1} and u_{n+2}. As a result the coefficients of each term in this form must vanish. This leads us to the determining equations for the infinitesimal LB transformations (5).

If in Eq. (5), $n \leqslant 2$, we obtain the Lie group of point transformations leaving Eq. (6) invariant. Without loss of generality we assume $n \geqslant 3$ in Eq. (5). It turns out that for $n \geqslant 3$, U is independent of x and t.

In our polynomial form, the coefficient of u_{n+2} vanishes and the coefficients of $(u_{n+1})^2$ and u_{n+1}, respectively, lead to determining equations

$$CU_{n,n} = 0, \tag{9}$$
$$nC'U_n u_1 = 2C\sum_{i=0}^{n-1} U_{n,i} u_{i+1}. \tag{10}$$

Solving Eqs. (9) and (10) we find that

$$U = \alpha (C)^{(1/2)n} u_n + E(u,u_1,\ldots,u_{n-1}), \tag{11}$$

where E is undetermined, and $\alpha =$ arbitrary constant.

The substitution of Eq. (11) into the remaining terms of Eq. (8) leads to a polynomial form in u_n whose coefficients of $(u_n)^2$ and u_n, respectively, lead to determining equations

$$CE_{n-1,n-1} = 0, \tag{12}$$
$$2C\left[\sum_{i=0}^{n-2} E_{n-1,i} u_{i+1}\right] + (1-n)C'E_{n-1} u_1 - \frac{\alpha}{4} n(n+3)C'(C)^{(1/2)n} u_2 + \alpha\,[\tfrac{1}{4}n^2(C')^2(C)^{(1/2)n-1} - \tfrac{1}{2}n(n+2)C''(C)^{(1/2)n}](u_1)^2 = 0. \tag{13}$$

Solving Eqs. (12) and (13) we find that

$$U = \alpha\,[(C)^{(1/2)n} u_n + \tfrac{1}{4}n(n+3)C'(C)^{(1/2)n-1} u_1 u_{n-1}] + F(u)u_{n-1} + G(u,u_1,\ldots,u_{n-2}), \tag{14}$$

where F and G are undetermined and, more importantly, for $\alpha \neq 0$ it is necessary that the conductivity $C(u)$ satisfy the differential equation

$$2CC'' = 3(C')^2. \tag{15}$$

Hence, it is necessary that

$$C(u) = a\cdot(u+b)^{-2}, \tag{16}$$

where a and b are arbitrary constants for the invariance of Eq. (1) under LB transformations. Without loss of generality we can set $a = 1$, $b = 0$, i.e., from now on we consider the equivalent p.d.e.

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(u^{-2}\frac{\partial u}{\partial x}\right) \equiv B. \tag{17}$$

This particular equation has been considered as a model equation of diffusion in high-polymeric systems.[4,5]

3. CONSTRUCTION OF A RECURSION OPERATOR; AN INFINITE SEQUENCE OF INVARIANT LB TRANSFORMATIONS OF EQ. (17)

For $n = 3$ it is easy to solve the rest of the determining equations and show that the only LB transformation leaving Eq. (17) invariant is

$$U = U^{(1)} = u^{-3}u_3 - 9u^{-4}u_1u_2 + 12u^{-5}(u_1)^3. \tag{18}$$

For $n = 4$ we obtain two linearly independent LB transformations $U^{(1)}$ and

$$U^{(2)} = u^{-4}u_4 - 14u^{-5}u_1u_3 - 10u^{-5}(u_2)^2 + 95u^{-6}(u_1)^2u_2 - 90u^{-7}(u_1)^4. \tag{19}$$

The existence of $U^{(1)}$ and $U^{(2)}$, combined with the work of Olver,[16] motivates us to seek a linear recursion operator $\mathscr{D}$ leading to infinitesimal LB transformations $U^{(k)}$ defined as follows:

$$(\mathscr{D})^k B = U^{(k)}, \quad k = 1,2,\cdots. \tag{20}$$

The character of $\{B, U^{(1)}, U^{(2)}\}$ leads one to consider for $\mathscr{D}$ the form

$$\mathscr{D} = pD_x + q + r(D_x)^{-1}, \tag{21}$$

where D_x is a total derivative operator, $(D_x)\cdot(D_x)^{-1}$ is the identity operator, and $\{p,q,r\}$ are functions of $\{u,u_1,u_2\}$. Then one can show that $\mathscr{D}B = U^{(1)}$ if and only if

$$p = u^{-1}, \tag{22}$$

and

$$q[u^{-2}u_2 - 2u^{-3}(u_1)^2] + ru^{-2}u_1 = -3u^{-4}u_1u_2 + 6u^{-5}(u_1)^3. \tag{23}$$

Furthermore, $(\mathscr{D})^2 B = U^{(2)}$ if and only if

$$q = -2u^{-2}u_1, \tag{24}$$

and

$$r = -u^{-2}u_2 + 2u^{-3}(u_1)^2. \tag{25}$$

A more concise expression for the operator is

$$\mathscr{D} = (D_x)^2\cdot(u^{-1})\cdot(D_x)^{-1}. \tag{26}$$

We now show that the constructed operator $\mathscr{D}$ is indeed a recursion operator. Let the operator

$$\begin{aligned} A &= \sum_{i=0}^{2} B_i (D_x)^i \\ &= u^{-2}(D_x)^2 - 4u^{-3}u_1 D_x + 6u^{-4}(u_1)^2 - 2u^{-3}u_2 \\ &= (D_x)^2\cdot u^{-2}, \end{aligned} \tag{27}$$

where $B_i = (\partial/\partial u_i)B$. Olver's work[16] shows that $\mathscr{D}$ is a recursion operator for Eq. (17) if and only if the commutator

$$[A - D_t, \mathscr{D}] = 0, \tag{28}$$

for any solution $u = \theta(x,t)$ of Eq. (17). Moreover, if $\mathscr{D}$ is a recursion operator, then the sequence $\{U^{(1)}, U^{(2)}, \cdots\}$ given by Eq. (20) is an infinite sequence of LB transformations leaving Eq. (17) invariant. It is straightforward to show that A and $\mathscr{D}$ satisfy Eq. (28).

The nature of $U^{(1)}$ and the form of a general U given by Eq. (11) show that for $n = l+2$, there are at most $k \leqslant l$ linearly independent LB transformations leaving Eq. (17) invariant since U must depend uniquely on u_{l+2}.

The proof that $\mathscr{D}$ is a recursion operator demonstrates that $k = l$ and hence we have found all possible LB transformations leaving Eq. (17) invariant, namely, $\{U^{(k)}\}$, $k = 1,2,\cdots$.

4. A MAPPING TO THE LINEAR DIFFUSION EQUATION

As far as we know all p.d.e.'s invariant under LB transformations have a recursion operator and, moreover, can be related to linear p.d.e.'s. This suggests the possibility of seeking a transformation relating Eq. (17) to a linear equation. This leads us to consider the linearization of Eq. (17), namely,

$$(A - \partial/\partial t) f = 0, \tag{29}$$

where A is given by Eq. (27) for any solution $u = \theta(x,t)$ of Eq. (17). Introducing a new variable $\bar{u}$ by

$$f = \frac{\partial}{\partial x}(u\bar{u}), \tag{30}$$

we obtain from Eq. (29) the equation

$$\left[\left(u^{-1}\frac{\partial}{\partial x}\right)^2 + u^{-3}u_1\frac{\partial}{\partial x} - \frac{\partial}{\partial t}\right]\bar{u} = 0 \tag{31}$$

and if we set

$$\begin{aligned} \frac{\partial}{\partial \bar{x}} &= u^{-1}\frac{\partial}{\partial x}, \\ \frac{\partial}{\partial \bar{t}} &= \frac{\partial}{\partial t} - u^{-3}u_1\frac{\partial}{\partial x}, \end{aligned} \tag{32}$$

Eq. (31) becomes

$$\frac{\partial^2 \bar{u}}{\partial \bar{x}^2} - \frac{\partial \bar{u}}{\partial \bar{t}} = 0. \tag{33}$$

Since $f = 0$ is always a solution of Eq. (29), the relation (30) suggests that we set $u\bar{u} =$ constant. This and Eqs. (32) lead us to the transformation

$$\begin{aligned} d\bar{x} &= u\,dx + u^{-2}u_1\,dt, \\ d\bar{t} &= dt, \\ \bar{u} &= u^{-1}, \end{aligned} \tag{34}$$

relating solutions $u = \theta(x,t)$ of Eq. (17) to solutions $\bar{u} = \bar{\theta}(\bar{x},\bar{t})$ of Eq. (33). Choosing a fixed point (x_0,t_0), we have the following integrated form of Eqs. (34):

$$\begin{aligned} \bar{x} &= \int_{x_0}^{x} u\,dx' - \int_{t_0}^{t}\left(\frac{\partial}{\partial x}u^{-1}\right)_{x = x_0} dt', \\ \bar{t} &= t - t_0, \\ \bar{u} &= u^{-1}. \end{aligned} \tag{35}$$

It is easy to check that Eqs. (35) indeed transform Eq. (17) to Eq. (33), and define a map relating the solutions of Eqs. (17) and (33). Moreover, if $u > 0$ $(\bar{u} > 0)$, Eqs. (35) define a one-to-one map since $\partial\bar{x}/\partial x > 0$ for each fixed t.[22]

We now show that under the transformation (34) the recursion operator $\mathscr{D}$ of Eq. (17) is transformed into the recursion operator

$$\bar{\mathscr{D}} = D_{\bar{x}}, \tag{36}$$

leading to an infinite sequence of LB transformations of the heat equation (33). The proof is as follows:

An LB transformation of the form (5) induces an LB transformation on the variables $\{\bar{x},\bar{t},\bar{u}\}$ through Eqs. (34), namely,

$$\begin{aligned} \bar{x}^* &= \bar{x} + \epsilon\bar{\xi} + O(\epsilon^2), \\ \bar{t}^* &= \bar{t}, \\ \bar{u}^* &= \bar{u} + \epsilon\bar{\eta} + O(\epsilon^2), \end{aligned} \tag{37}$$

where $\bar{\xi}$ and $\bar{\eta}$ are defined by

$$\begin{aligned} d\bar{\xi} &= \mathscr{A}\,d\bar{x} + \mathscr{B}\,d\bar{t}, \\ \mathscr{A} &= \bar{u}U, \quad \mathscr{B} = \bar{u}_{\bar{x}}U + (\bar{u})^2(U^x - 2U), \\ \bar{\eta} &= -(\bar{u})^2 U. \end{aligned} \tag{38}$$

It turns out that for any solution $\bar{u} = \bar{\theta}(\bar{x},\bar{t})$ of Eq. (33), $\mathscr{A}$ and $\mathscr{B}$ satisfy the integrability condition $D_{\bar{t}}\mathscr{A} = D_{\bar{x}}\mathscr{B}$, so that $d\bar{\xi}$ is an exact differential. The integrated form of $\bar{\xi}$ is

$$\bar{\xi} = -(D_{\bar{x}})^{-1}[\bar{\eta}\cdot(\bar{u})^{-1}] + c, \tag{39}$$

where c is an arbitrary constant. Since $U^{(i+1)} = \mathscr{D}\, U^{(i)}$, where $\mathscr{D}$ is given by Eq. (26), for $c = 0$ we get a corresponding infinite sequence of invariant infinitesimal LB transformations $\{\bar{U}^{(i)}\}$ for Eq. (33), namely,

$$\bar{U}^{(i)} = \bar{\eta}^i - \bar{u}_{\bar{1}}\bar{\xi}^i,$$

where

$$\bar{\eta}^i = -(\bar{u})^2 U^{(i)}, \tag{40}$$
$$\bar{\xi}^i = -(D_{\bar{x}})^{-1}[\bar{\eta}^i\cdot(\bar{u})^{-1}],$$

and $\bar{u}_{\bar{i}} = (\partial/\partial \bar{x})^i \bar{u}$. From Eqs. (40) it is simple to show that

$$\bar{U}^{(i+1)} = D_{\bar{x}}\bar{U}^{(i)}, \tag{41}$$

leading to Eq. (36). Moreover,

$$\bar{U}^{(i)} = D_{\bar{x}}(\bar{u}\bar{\xi}^i) = (D_{\bar{x}})^i \bar{u}_{\bar{2}}, \quad i = 1,2,\cdots. \tag{42}$$

$D_{\bar{x}}$ corrsponds to the obvious invariance of Eq. (33) under translations in $\bar{x}$.

It is interesting to note that the recursion operator for the invariant LB transformations of Burgers' equation is also mapped into the space translation operator under the Hopf–Cole transformation relating Burgers' equation to the heat equation. Moreover, we can obtain the Hopf–Cole transformation by examining the linearization equation (29) corresponding to Burgers' equation.

5. PROPERTIES OF SOLUTIONS OF EQ. (17) FROM THE MAPPING

We now consider the use of Eqs. (34) in constructing solutions to Eq. (17). It is easy to show that Eqs. (34) are equivalent to

$$dx = \bar{u}\, d\bar{x} + \bar{u}_{\bar{1}}\, d\bar{t},$$
$$dt = d\bar{t}, \tag{43}$$
$$u = (\bar{u})^{-1},$$

with an integrated form

$$x = \int_{\bar{x}_0}^{\bar{x}} \bar{u}\, d\bar{x}' + \int_{\bar{t}_0}^{\bar{t}} (\bar{u}_{\bar{1}})_{\bar{x}=\bar{x}_0}\, d\bar{t}',$$
$$t = \bar{t} - \bar{t}_0, \tag{44}$$
$$u = (\bar{u})^{-1},$$

for some fixed point $(\bar{x}_0,\bar{t}_0)$. In the following, we assume $u > 0$ ($\bar{u} > 0$). Without loss of generality, we set $\bar{x}_0 = \bar{t}_0 = 0$.

A. Explicit formula connecting solutions; examples

First we consider the problem of giving a more explicit formula for relating solutions of Eq. (33) to those of Eq. (17). Let $\bar{u} = \bar{\theta}(\bar{x},\bar{t})$ be a solution of Eq. (33) on the domain $\bar{t} > 0$, $\bar{x}\in(\bar{x}_1,\bar{x}_2)$. By Eqs. (43),

$$x = X(\bar{x},\bar{t}) = \int_0^{\bar{x}} \bar{\theta}(\bar{x}',\bar{t})\, d\bar{x}' + \int_0^{\bar{t}} \left(\frac{\partial\bar{\theta}(\bar{x},\bar{t}')}{\partial\bar{x}}\right)_{\bar{x}=0} d\bar{t}'. \tag{45}$$

This uniquely determines the function X^{-1}, $\bar{x} = X^{-1}(x,t)$, where $\bar{t} = t$. Now Eqs. (44) lead to the following solution of Eq. (17):

$$u = \theta(x,t) = \frac{1}{\bar{\theta}(X^{-1}(x,t),t)},$$

on the domain $x\in(x_1(t),\ x_2(t))$, $t > 0$, where (46)

$$x_1(t) = X(\bar{x}_1,t), x_2(t) = X(\bar{x}_2,t).$$

In a similar manner, Eqs. (35) map a solution $u = \theta(x,t)$ of Eq. (17) to

$$\bar{u} = \bar{\theta}(\bar{x},\bar{t}) = \frac{1}{\theta(\bar{X}^{-1}(\bar{x},\bar{t}),\bar{t})}$$

on the domain $\bar{x}\in(\bar{x}_1(\bar{t}),\bar{x}_2(\bar{t}))$, $\bar{t} > 0$ where (47)

$$\bar{x}_1(\bar{t}) = \bar{X}(x_1,\bar{t}),\quad \bar{x}_2(\bar{t}) = \bar{X}(x_2,\bar{t}),$$

$$\bar{x} = \bar{X}(x,t) = \int_0^x \theta(x',t)\, dx' - \int_0^t \left[\frac{\partial}{\partial x}(\theta(x,t'))^{-1}\right]_{x=0} dt', \tag{48}$$

with the corresponding definition of the function $\bar{X}^{-1}(\bar{x},\bar{t}) = x$.

Example 1: The source solution of Eq. (33), i.e., $\bar{u} = \bar{\theta}(\bar{x},\bar{t}) = a(4\pi\bar{t})^{-1/2}e^{-(\bar{x}^2/4\bar{t})}$ on the domain $-\infty < \bar{x} < \infty, \bar{t} > 0$, is mapped by Eqs. (45) and (46) into the following separable solution of Eq. (17):

$$u = \theta(x,t) = a^{-1}(4\pi t)^{1/2}e^{v^2},$$

on the domain $-\frac{1}{2}a < x < \frac{1}{2}a$, $t > 0$, where $v(x)$ is defined by (49)

$$x = \frac{a}{\sqrt{\pi}}\int_0^v e^{-y^2}\, dy.$$

Note that $\lim_{x\to\pm\frac{1}{2}a}\theta(x,t) = +\infty$.

Example 2. The dipole solution of Eq. (33), i.e.,

$$\bar{u} = \bar{\theta}(\bar{x},\bar{t}) = -\frac{\partial}{\partial\bar{x}}[a(4\pi\bar{t})^{-1/2}e^{-(\bar{x}^2/4\bar{t})}],$$

on the domain $0 < \bar{x} < \infty, \bar{t} > 0$, is mapped by Eqs. (45) and (46) into the following self-similar solution of Eq. (17):

$$u = \theta(x,t) = x^{-1}(2t)^{1/2}\left[\ln\left(\frac{a^2}{4\pi t x^2}\right)\right]^{-1/2} \tag{50}$$

on the shrinking domain $0 < x < a(4\pi t)^{-1/2}$, $t > 0$.

B. Connection between initial conditions; connection between boundary conditions

The mapping formulas (34) and (43) demonstrate a one-to-one correspondence (within translation of x,t) between initial conditions for Eq. (17) and those for Eq. (33). As for the connection between boundary conditions, from the same formula it is easy to see that $x = s(t)$ is an insulating boundary of Eq. (17), i.e., $[\partial\theta(x,t)/\partial x]_{x=s(t)} = 0$, if and only if the corresponding boundary $\bar{x} = \bar{s}(\bar{t})$ is an insulating boundary of Eq. (33), i.e., the corresponding solution $\bar{u} = \bar{\theta}(\bar{x},\bar{t})$ satisfies $[\partial\bar{\theta}(\bar{x},\bar{t})/\partial\bar{x}]_{\bar{x}=\bar{s}(\bar{t})} = 0$. Moreover, $s(t) = \text{const}$ if and only if $\bar{s}(\bar{t}) = \text{const}$, i.e., there is a one-to-one correspondence between fixed insulating boundaries of Eqs. (17) and (33).

In general, a noninsulating boundary condition for Eq. (17), on a fixed boundary $x = \text{const} = c$, is mapped into a noninsulating boundary condition of Eq. (33) with a corresponding moving boundary $\bar{x} = \bar{s}(\bar{t}) \neq$const with speed

$$\frac{d\bar{s}}{d\bar{t}} = \left[[\theta(x,t)]^{-2} \frac{\partial \theta(x,t)}{\partial x} \right]_{\substack{x=c \\ t=\bar{t}}}, \tag{51}$$

where, as previously mentioned, $u = \theta(x,t) > 0$.

6. CONCLUDING REMARKS

(a) From invariance under the LB transformations $\{U^{(i)}\}$, $i = 1,2,\cdots$, there exist similarity solutions of Eq. (17), i.e., $u = \theta(x,t;n)$, whose similarity forms satisfy

$$U^{(n)} + \sum_{k=1}^{n-1} c_k U^{(k)} = 0, \tag{52}$$

where $\{c_1, c_2, ..., c_{n-1}\}$ are arbitrary constants, $n = 1,2,\cdots$. For example, for $n = 1$, Eq. (52) leads to the similarity form

$$u = \theta(x,t;1) = [a(t)\cdot(x + b(t))^2 + c(t)]^{-1/2}, \tag{53}$$

where $\{a(t), b(t), c(t)\}$ are arbitrary. Substituting Eq. (53) into Eq. (17) we find that Eq. (53) solves Eq. (17) if and only if $a = \alpha$, $b = \beta$, and $c = \gamma e^{2\alpha t}$, where $\{\alpha,\beta,\gamma\}$ are arbitrary constants. This solution is not contained in the class of similarity solutions of Eq. (17) obtained from invariance under a four-parameter point Lie group.[7,8]

(b) The infinitesimal transformations (5) of the four-parameter point group of Eq. (17) are given by

$$U^a = u + xu_1, \quad U^b = xu_1 + 2tu_t,$$
$$U^c = u_1, \quad U^d = B. \tag{54}$$

Under the mapping (34), these are transformed, respectively, to corresponding infinitesimals of invariant point group transformations of Eq. (33):

$$\bar{U}^a = \bar{u}, \quad \bar{U}^b = \bar{x}\bar{u}_{\bar{1}} + 2\bar{t}\bar{u}_{\bar{t}},$$
$$\bar{U}^c = 0, \quad \bar{U}^d = \bar{B} = \bar{u}_{\bar{2}}. \tag{55}$$

Conversely, the mapping (34) transforms the six-parameter point Lie group of Eq. (33) as follows: The three-parameter subgroup of infinitesimals given by Eq. (55) transforms to $\{U^a, U^b, U^d\}$ given by Eqs. (54) and $\bar{U} = \bar{u}_{\bar{1}}$ transforms to $U = 0$; the remaining infinitesimal point group transformations $\bar{U}^e = \bar{x}\bar{u} + 2\,\bar{t}\bar{u}_{\bar{t}}$ and $\bar{U}^f = (\frac{1}{4}\bar{x}^2 + \frac{1}{2}\bar{t})\bar{u} + \bar{x}\bar{t}\bar{u}_{\bar{1}} + \bar{t}^2\bar{u}_{\bar{t}}$ are mapped, respectively, into infinitesimals which depend on $\{x,t,u,u_1\}$ and integrals of u.

(c) Generally speaking, the action of a recursion operator $\mathscr{D}$ on any infinitesimal invariance transformation U of the form (5) (whether of point group or LB type) yields a new infinitesimal transformation $U' = \mathscr{D}U$ if $\mathscr{D}U \neq 0$. For Eq. (17), we can show that $\mathscr{D}U^a = \mathscr{D}U^b = \mathscr{D}U^c = 0$.

(d) The heat equation is a special limiting case of Eq. (3) obtained by setting $a = b^2$ and then observing $\lim_{b\to\infty} b^2(u+b)^{-2} = 1$. As one might expect if $a = b^2$, for the corresponding recursion operator $\mathscr{D}$, $\lim_{b\to\infty}\mathscr{D} = \partial/\partial x$, and the mapping formulas reduce to identity mappings.

(e) Since Eq. (1) admits an infinite sequence of LB transformations if and only if $C(u)$ satisfies Eq. (15) with associated mapping (34) whereas Eq. (4) admits an infinite sequence of LB transformations, there is no point transformation of the form

$$\bar{x} = K(x,t,u),$$
$$\bar{t} = L(x,t,u),$$
$$\bar{u} = M(x,t,u),$$

relating solutions of Eq. (1) and those of Eq. (4).

ACKNOWLEDGMENTS

It is a pleasure to thank Professor C.E. Wulfman for stimulating discussions. We thank the National Research Council of Canada for financial support.

[1] J.G. Berryman and C.J. Holland, Phys. Rev. Lett. **40**, 1720 (1978).
[2] P. Baeri, S.U. Campisano, G. Foti, and E. Rimini, J. Appl. Phys. **50**, 788 (1979).
[3] Y.S. Touloukian, P.W. Powell, C.Y. Ho, and P.G. Klemens, *Thermophysical Properties of Matter* (Plenum, New York, 1970), Vol. 1.
[4] H. Fujita, Text. Res. J. **22**, 757, 823 (1952).
[5] R.H. Peters, in *Diffusion in Polymers*, edited by J. Crank and G.S. Park (Academic, New York, 1968); R. McGregor, R.H. Peters, and J.H. Petropolous, Trans. Faraday Soc. **58**, 1054 (1962).
[6] J. Crank, *The Mathematics of Diffusion* (Clarendon, Oxford, 1956).
[7] G.W. Bluman, Ph.D. Thesis, California Institute of Technology, 1967.
[8] G.W. Bluman and J.D. Cole, *Similarity Methods for Differential Equations* (Springer, New York, 1974).
[9] R.L. Anderson, S. Kumei, and C.E. Wulfman, Rev. Mex. Fis. **21**, 1, 35 (1972); Phys. Rev. Lett. **28**, 988 (1972).
[10] H. Steudel, Ann. Physik **32**, 205 (1975).
[11] N. Ibragimov and R.L. Anderson, J. Math. Anal. Appl. **59**, 145 (1977).
[12] E. Noether, Nachr. Akad. Wiss. Goettingen Math. Phys. K1.2 234 (1918).
[13] H. Steudel, Ann. Physik **32**, 445 (1975).
[14] S. Kumei, J. Math. Phys. **18**, 256 (1977).
[15] S. Kumei, J. Math. Phys. **16**, 2461 (1975).
[16] P. Olver, J. Math. Phys. **18**, 1212 (1977).
[17] A.A. Fokas, Ph.D. Thesis, California Institute of Technology, 1979, Chap. V.
[18] *Bäcklund Transformations*, edited by R.M. Miura (Springer, New York, 1976).
[19] L.V. Ovsjannikov, Dokl. Akad. Nauk SSSR **125**, 492 (1959).
[20] S. Lie, Arch. Math. **6**, 328 (1881).
[21] An infinitesimal LB transformation

$$x^* = x + \epsilon\xi + O(\epsilon^2),$$
$$t^* = t + \epsilon\tau + O(\epsilon^2),$$
$$u^* = u + \epsilon\mu + O(\epsilon^2),$$

acts on a surface $F(x,t,u) = 0$ in the same manner as

$$x^* = x,$$
$$t^* = t,$$
$$u^* = u + \epsilon U + O(\epsilon^2),$$

where $U = \mu - \xi u_x - \tau u_t$.
[22] G. Rosen, Phys. Rev. B **19**, 2398 (1979). After submitting this paper we discovered the above reference through Nonlinear Science Abstracts. Here Rosen discovered transformation (35) and worked out some examples. We are also grateful to the referee for bringing the above paper to our attention.

Towards a factorization of M_4 [a)]

Patrick L. Nash
Department of Physics and Astronomy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514

(Received 23 August 1979; accepted for publication 21 January 1980)

It may be desirable to eliminate M_4 as the underlying manifold in which some physical theories are cast, and recast these theories in the space "$\sqrt{M_4}$." We investigate some of the properties of one such space, which we denote by S_8. S_8 can be coordinatized by real eight-component spinors. The spinor algebra on this space is developed in this paper. It is shown that a (nondegenerate) spinor in S_8 determines an orthogonal tetrad on M_4 (the set of these spinors determines the space of orthogonal frames over M_4), and that this spinor corresponds to a "particle." A simple geometrical interpretation of the Dirac equation arises in arriving at this correspondence.

1. INTRODUCTION

When the gravitational interaction is neglected the physical world is usually modeled by M_4. For example, relativistic quantum mechanics as presently formulated depends crucially on the fact that the background space in which it is cast is M_4, regarded as either the flat spacetime manifold or the momentum dual vector space. Implicit in such models is a presumably physically meaningful isomorphism between the "points" of the universe and the points of the mathematical continuum of M_4. However, when one speaks of the universe itself as being "continuous" or a "continuum," then one is usually referring to an (apparent) ability to *continuously* vary the relative distances between an arbitrary number of macroscopic bodies situated in our universe. This idea does not generalize to a satisfactory operational *microscopic* definition of spacetime continuity. Einstein warned of "removing certain fundamental concepts from the domain of empiricism, where they are under our control, to the intangible heights of the *a priori*."[1] One such concept may be the notion of spacetime continuity. Accordingly, it is of great importance to construct a theory of interacting particles which is not cast in a mathematical space that is "isomorphic" to our universe. Instead, the theory is to be formulated in a space from which the concepts of spacetime and spacetime continuity can be derived, being operationally defined by properties of interacting particles evolving in this space. This idea is not new; to mention the most important example, significant progress has been made by Penrose and his co-workers in eliminating the spacetime continuum concept from physics using the twistor formalism, and in deriving space from interacting spin-networks.[2]

Taking a somewhat different approach, one might attempt to genealize some physical theories by eliminating M_4 as their underlying manifold and replacing it by, crudely speaking, $\sqrt{M_4}$. More precisely, by $\sqrt{M_4}$ we mean a triple (X, ω, G) such that (i) X is a real even-dimensional C^∞ manifold, (ii) ω is a real nondegenerate closed 2-form on X (ω is a symplectic form on X and (X, ω) is a symplectic manifold), (iii) $\omega/2\pi$ represents an integral de Rham cohomology class [(X, ω) is a *quantizable* symplectic manifold], and (iv) G is a Lie group which acts on X as a group of diffeomorphisms, and G contains $\overline{\mathscr{L}^{\uparrow}_{+}} \times \mathbb{R}^*$, where $\overline{\mathscr{L}^{\uparrow}_{+}}$ is the universal covering group of the proper orthochronous Lorentz group, and $\mathbb{R}^* = \mathbb{R} - \{0\}$ is the group of nonzero real numbers under multiplication ($\mathbb{R}^*$ corresponds to dilations). Qualitatively, a classical theory cast on $\sqrt{M_4}$ gives rise to a quantum theory as follows: On account of (iv) there exists a set $B(X)$ of bilinear mappings from X into (the components of elements in) the tensor algebra of M_4. $B(X)$ is contained in the ring of smooth real-valued functions on X. Classically, the observables of a theory are to be found among those functions in this ring that take values as the components of elements in the tensor algebra of M_4; the space of functions $B(X)$ is the generator of this set of (classical) observables. We denote the ring of observables by $O(X)$. The symplectic form ω defines a Poisson bracket on X which gives $O(X)$ the structure of a Lie algebra over $\mathbb{R}$. Since (X, ω) is a quantizable symplectic manifold, there exists a representation of the Lie algebra of classical observables by Hermitian operators on an appropriate Hilbert space, such that the mapping is a Lie algebra isomorphism (i.e., the Poisson bracket of two classical observables is mapped into the commutator of the corresponding two Hermitian operators.).[3]

This paper is devoted to a study of a particular choice for $\sqrt{M_4}$, and to the construction of the space of functions $B(X)$. For the space X we consider an eight-dimensional differentiable manifold that admits a global canonical atlas, and whose coordinate functions in this atlas are real eight-component spins. Henceforth we denote X by S_8. For G we take $\overline{\mathrm{SO}(3,3)} \times \mathbb{R}^*$)($\overline{\mathrm{SO}(3,3)}$ is the universal covering group of the special Lorentz group on a flat spacetime with three space and three time dimensions). The emphasis in the first portion of this paper is on developing the algebra of these spinors and in constructing geometrical objects on M_4 from spinors in S_8. In order to forge a simple and direct link between M_4 and S_8 we begin our work with the construction of an orthonormal tetrad of vectors from spinors. We show that except in certain degenerate cases, *every* spinor in S_8 determines an orthogonal tetrad of vectors $e^{\alpha}_{(\mu)}$; moreover, the (future-pointing) timelike member of the tetrad $e^{\alpha}_{(4)}$ is the sum of two linearly independent (future-pointing) null vectors, while $e^{\alpha}_{(3)}$ is the difference of these null vectors. In the degenerate case a spinor determines a unique null vector, and we unimaginatively call this spinor a null spinor.

[a)] This paper is dedicated to my Mother and Father.

 0022-2488/80/051024-08$01.00

In Sec. 8 (equivalence classes of) nondegenerate spinors are identified with linear combinations of two types of massive spin-$\frac{1}{2}$ fermions (the two types can *possibly* be characterized according to whether they carry an intrinsic electric charge or an intrinsic magnetic charge). The correspondence between these spinors and particles is established by a comparison with the usual Dirac theory, which also provides some insight into the Dirac equation itself. However the existence of this correspondence relies on an imbedding in M_4, and is therefore only of heuristic value.

2. CONSTRUCTION OF THE TETRAD

The real-valued spinor ψ with components ψ^a $a, b, \cdots = 1, \ldots, 8$ coordinatizes an eight-dimensional space S_8. The tetrad is to be constructed from quadratic products of the components of this spinor. It happens that this construction proceeds most clearly by temporarily leaving M_4 and going to a larger space M_6 with "signature" $(1, 1, 1, -1, -1, -1)$. This is because there exists a natural homomorphism between Lorentz transformations on S_8 and M_6. The matrix generators of the special Lorentz transformations SO(3, 3) (for brevity we shall henceforth omit the characterization "universal covering group") on S_8 provide us with a representation of the SO(3, 3) Lie algebra. By saturating the S_8 indices of these matrices with $\psi^a \psi^b$ we generate a realization of the SO(3, 3) Lie algebra. Two of the spacelike members of the tetrad, $e^\alpha_{(1)}$ and $e^\alpha_{(2)}$, can be identified with eight of the members of this realization. The remaining members of the algebra also have physical interpretations. Therefore we shall proceed rather formally beginning with a discussion of the above mentioned homomorphism.

3. CONNECTION BETWEEN LORENTZ TRANSFORMATIONS ON M_6 and S_8 [4]

The homomorphism between Lorentz transformations on M_6 and S_8 may be established as follows.[5] Let Γ^A be six real matrices which generate an irreducible representation of the Clifford algebra C_6:

$$\Gamma^A \Gamma^B + \Gamma^B \Gamma^A = 2g^{AB}, \tag{1}$$

where $g_{AB} = \mathrm{diag}\,(1, 1, 1, -1, -1, -1)$ is the metric tensor on M_6. Suppose that $\Lambda^A{}_B \in O\,(3, 3)$ is an arbitrary Lorentz transformation on M_6, $x^A \to x'^A = \Lambda^A{}_B x^B$. The $\Lambda^A{}_B$ satisfy the conditions $g_{RS} = g_{AB} \Lambda^A{}_R \Lambda^B{}_S$. Since $\Lambda^A{}_R \Gamma^R \Lambda^B{}_S \Gamma^S + \Lambda^B{}_S \Gamma^S \Lambda^A{}_R \Gamma^R = 2g^{AB}$ and since the Γ^A are irreducible, there exists a nonsingular matrix $M(\Lambda)$ such that $\Lambda^A{}_B \Gamma^B = M^{-1} \Gamma^A M$, or

$$\Gamma^A = \Lambda^A{}_B M \Gamma^B M^{-1}. \tag{2}$$

M may be assumed to be real and unimodular and so is determined up to a factor of ± 1. The set of all such matrices gives a faithful 2-1 representation of O (3, 3). For every such matrix M there exists a unique Λ given by

$$\Lambda^A{}_B = \tfrac{1}{8}\mathrm{trace}(M^{-1} \Gamma^A M \Gamma_B). \tag{3}$$

A Lorentz transformation $x \to x' = \Lambda x$ on M_6 coincides with the transformation $\psi \to \psi' = M\psi$ in S_8, up to the above-mentioned factor of ± 1.

For the special Lorentz tranformations on S_8, M is generated by the antisymmetrized products of the Γ^A, denoted by M^{AB}. Let

$$M^{AB} = -\tfrac{1}{4}(\Gamma^A \Gamma^B - \Gamma^B \Gamma^A) = -\tfrac{1}{4}[\Gamma^A, \Gamma^B]. \tag{4}$$

using Eq. (1) we find that

$$[M^{AB}, \Gamma_R] = \delta^A{}_R \Gamma^B - \delta^B{}_R \Gamma^A, \tag{5}$$

and

$$[M^{AB}, M^{RS}] = g^{AR} M^{BS} - g^{AS} M^{BR} - g^{BR} M^{AS} + g^{BS} M^{AR}. \tag{6}$$

Therefore by Eq. (6) the M^{AB} are the infinitesimal generators of an eight-dimensional representation of SO(3, 3). If $\Lambda^A{}_B = \delta^A{}_B - \omega^A{}_B + \cdots = (e^{-\omega})^A{}_B$, where $\omega_{AB} = -\omega_{BA}$ are parameters, then

$$M = 1 + \tfrac{1}{2}\omega_{AB} M^{AB} + \cdots = \exp(\tfrac{1}{2}\omega_{AB} M^{AB}), \tag{7}$$

which follows from Eqs. (1 and 5).

This representation of the SO (3, 3) Lie algebra is completely reducible[5] so that the M^{AB} are of the form

$$M^{AB} = \begin{pmatrix} S^{AB} & \\ & T^{AB} \end{pmatrix} \tag{8}$$

Therefore a matrix representing a rotation of SO (3, 3) in S_8 is of the form $\left(\begin{smallmatrix} S & \\ & T \end{smallmatrix}\right)$, where S is generated by S^{AB} and T by T^{AB}. The spinor ψ lives in the eight-dimensional space that carries this representation of SO (3, 3). Because the representation is reducible this eight-component spinor may be decomposed into a direct sum

$$\psi = \begin{pmatrix} \lambda \\ \xi \end{pmatrix} \tag{9}$$

of two distinct four-component spinors. Special Lorentz transformations such that $\psi \to \psi'$ carry $\lambda \to \lambda'$ and $\xi \to \xi'$ with no mixing of the components λ and ξ; the decomposition into distinct four-component objects is preserved under SO (3, 3). Moreover it is possible to pick an irreducible representation of the Γ^A to that $\tilde{\xi}\lambda$ ($\tilde{\xi}$ denotes the transpose of ξ) is an invariant under SO (3, 3). However, improper Lorentz transformations do not preserve this decomposition; nevertheless it is often convenient to write ψ as $\psi = \left(\begin{smallmatrix} \lambda \\ \xi \end{smallmatrix}\right)$, keeping in mind when this has an invariant meaning.

4 METRIC BISPINOR AND A REALIZATION OF THE SO (3, 3) LIE ALGEBRA

The M^{AB} defined in Eq. (4) give a (reducible) representation of the SO (3, 3) Lie algebra, with the Lie Bracket being the usual commutator bracket. We may also construct a realization of this Lie algebra from homogeneous quadratic polynomials in ψ^a. The Lie bracket in this case will be a sort of Poisson bracket. We shall see that the elements of the Lie algebra in this realization have simple geometrical interpretations once the O (3, 3) symmetry is broken down to O (3, 1). The first step in the construction of this realization is to define an appropriate metric bispinor on S_8 [which will be used in forming the bilinear products in ψ^a to ensure that they transform covariantly under SO (3, 3)]. There exist essentially two distinct choices for the metric bispinor. The first, denoted by Ω, is defined by

$$\Omega\Gamma^A = \widetilde{\Gamma}^A\Omega \tag{10}$$

and is consequently skew-symmetric ($\widetilde{\Gamma}^A$ denotes the transpose of Γ^A). The second, D, is defined by

$$D\Gamma^A = -\widetilde{\Gamma}^A D \tag{11}$$

and hence is symmetric. D may be obtained from Ω by right translation with Γ^7, i.e., $D = \Omega\Gamma^7$, where Γ^7 is defined by

$$\begin{aligned}\Gamma^7 &= \Gamma^1\Gamma^2\Gamma^3\Gamma^4\Gamma^5\Gamma^6\\ &= (1/6!)\epsilon_{ABCDEF}\Gamma^A\Gamma^B\Gamma^C\Gamma^D\Gamma^E\Gamma^F.\end{aligned} \tag{12}$$

Γ^7 has the properties that

$$\Gamma^7\Gamma^A + \Gamma^A\Gamma^7 = 0, \tag{13}$$

$$(\Gamma^7)^2 = 1, \tag{14}$$

and

$$\Omega\Gamma^7 = -\widetilde{\Gamma}^7\Omega. \tag{15}$$

We shall employ Ω as the metric bispinor. In addition, due to its skew symmetry, Ω is a symplectic structure on S_8.[6] We shall use this fact when defining a Lie bracket for this algebra.

At this point it is convenient to introduce some simple index notation to compliment the matrix notation which has been used. Let ψ denote the column matrix of the ψ^a and $\tilde{\psi}$ the transpose of ψ. Associate indices as follows: $\psi \longleftrightarrow \psi^a$; $\Omega \longleftrightarrow \Omega_{ab} = -\Omega_{ba}$; $\tilde{\psi}\Omega \longleftrightarrow \psi_b = \psi^a\Omega_{ab}$; $-\Omega^{-1} \longleftrightarrow \Omega^{ab} = -\Omega^{ba}$, and so the convention is $\Omega^{ab}\Omega_{bc} = -\delta^a{}_c$; $\Gamma^A \longleftrightarrow \Gamma^{Aa}{}_b$. In this notation Eq. (10) reads $\Omega_{ac}\Gamma^{Ac}{}_b = \Gamma^{Ac}{}_a\Omega_{cb}$. Introducing the convention that spinor indices are to be raised according as

$$X^a = \Omega^{ab}X_b, \tag{16}$$

and lowered as

$$X_b = X^a\Omega_{ab}, \tag{17}$$

(note position of indices), this equation becomes

$$-\Gamma^A{}_{ab} = \Gamma^A{}_{ba}. \tag{18}$$

Similarly Eq. (15) becomes

$$\Gamma^7{}_{ab} = \Gamma^7{}_{ba}. \tag{19}$$

From the definition of M^{AB} given in Eq. (4) and using Eq. (10) we have that

$$\Omega M^{AB} = -\widetilde{M}^{AB}\Omega, \tag{20}$$

i.e.,

$$M^{AB}{}_{ab} = M^{AB}{}_{ba}. \tag{21}$$

Consider now a matrix M representing a rotation of SO (3, 3). By Eqs. (7) and (20) Ω is invariant under this transformation:

$$\begin{aligned}\Omega \rightarrow \Omega' &= \widetilde{M}\Omega M\\ &= \Omega.\end{aligned} \tag{22}$$

In index notation this reads

$$\Omega_{ab} = M^c{}_a\Omega_{cd}M^d{}_b. \tag{23}$$

The final step before giving a realization of the SO (3, 3) Lie algebra is to define a Lie bracket. For any functions F and G only of ψ, their Lie bracket is defined to be

$$\{F,G\} = \frac{\partial F}{\partial\psi^a}\Omega^{ab}\frac{\partial G}{\partial\psi^b}. \tag{24}$$

We now define the realization of the SO (3, 3) Lie algebra, denoted by m^{AB}, as

$$\begin{aligned}m^{AB} &= -\tfrac{1}{2}\tilde{\psi}\,\Omega\, M^{AB}\psi\\ &= \tfrac{1}{2}M^{AB}{}_{ab}\psi^a\psi^b.\end{aligned} \tag{25}$$

To reveal the transformation properties of m^{AB} under SO (3, 3) we first consider M^{AB}. By Eqs. (2) and (4) M^{AB} is invariant under the combined SO (3, 3) transformations on M_6 and S_8. We write this as

$$M^{-1}M^{AB}M = \Lambda^A{}_R\Lambda^B{}_S M^{RS}. \tag{26}$$

Let ψ and ψ' be related by a SO (3, 3) rotation, $\psi' = M\psi$; we find by multiplying Eq. (26) from the left with $\tilde{\psi}\Omega$ and from the right with ψ that

$$m'^{AB} = \Lambda^A{}_R\Lambda^B{}_S m^{RS}, \tag{27}$$

where $m'^{AB} = -\frac{1}{2}\tilde{\psi}'\Omega M^{AB}\psi'$. Thus m^{AB} transforms as a type (2, 0) skew-symmetric tensor under SO (3, 3). Under an improper Lorentz transformation contained in O (3, 3), say a reflection about the x^A axis in M_6, a corresponding transformation matrix on S_8 is given by $\Gamma^A\Gamma^7$, which follows from Eq. (2). Equations (10) and (15) then imply that $-g^{AA}\Omega = \widetilde{\Gamma^A\Gamma^7}\Omega\Gamma^A\Gamma^7$ (no summation on A) under this reflection. In particular under a spatial reflection in M_6 this means that

$$m^{AB} \rightarrow m'^{AB} = -\Lambda^A{}_R\Lambda^B{}_S m^{RS}. \tag{28}$$

We pick up a minus sign under such spatial reflections. However, under an inversion of one of the time axes the factor $-g^{AA}$ is one and no such minus sign appears.

The m^{AB} may be grouped into sets transforming as tensors under SO (3, 1) once the O (3, 3) symmetry is broken. To achieve this we shall henceforth restrict our attention to the affine subspace of M_6 given by

$$\begin{aligned}x^5 &= \text{constant},\\ x^6 &= \text{constant},\end{aligned} \tag{29}$$

thus breaking the O (3, 3) symmetry. The only physically admissible O (3, 3) transformations are those which leave these constraint equations invariant. For SO (3, 3) transformations this is equivalent to $\omega^A{}_5 = 0 = \omega^A{}_6$. Under these restrictions the m^{AB} may be decomposed into a type (2, 0) skew-symmetric tensor, two vectors, and a scalar [under SO (3, 1)] as follows:

$$m^{\alpha\beta} = \Sigma^{\alpha\beta}, \tag{30}$$

$$m^{\alpha 5} = e^\alpha_{(1)}, \tag{31}$$

$$m^{\alpha 6} = e^\alpha_{(2)}, \tag{32}$$

and

$$m^{56} = N. \tag{33}$$

We shall see that $e^\alpha_{(1)}$ and $e^\alpha_{(2)}$ are spacelike vectors and mutually orthogonal, $e^{\alpha}_{(1)}e_{(2)\alpha} = 0$, except in the degenerate cases.

The Lie bracket relations among the m^{AB} are invariant under O (3, 3). They may be evaluated using Eqs. (24), (25), and (6); we find that

$$\{m^{AB},m^{RS}\} = g^{AR}m^{BS} - g^{AS}m^{BR} - g^{BR}m^{AS} + g^{BS}m^{AR}. \quad (34)$$

For $(A, B, R, S) = (\alpha, \beta, \mu, \nu)$ these relationships show that the $\Sigma^{\alpha\beta}$ by themselves possess a closed algebra and generate a realization of the special Lorentz group SO (3, 1). In fact $\hbar\Sigma^{\alpha\beta}$ turns out to be the spin tensor of a *classical* spinning point electron if one writes down a Lagrangian for the electron constructed from $\{\psi, \dot{\psi}, \dot{X}^{\alpha}, A_{\alpha} = \text{vector potential}\}$. This assertion should be plausible in light of these "commutation" relations.

5. The γ MATRICES[4]

The matrices defined below will be used in the next section in the construction of an irreducible representation of the Clifford algebra C_6. Dirac[7] devised the labeling scheme and multiplication rules upon which this process relies.

Consider a set of 15 real 4×4 matrices $\gamma^{AB} = -\gamma^{BA}$, where $A, B, \cdots = 1,\ldots,6$. The row and column indices are suppressed; $A, B, \cdots$ label the individual matrices. Let γ_0 be the 4×4 unit matrix. The γ^{AB} are defined to be skew-symmetric when A and B are both from the set $\{1, 2, 3\}$ or from the set $\{4, 5, 6\}$, and symmetric otherwise. For example γ^{34} is symmetric and γ^{23} is skew-symmetric. There are six skew-symmetric matrices and nine symmetric.

We quote from Dirac's work[7] the following multiplication rule for the γ^{AB}:

"We use the notation $\gamma^{AB}\gamma^{CD} = \gamma^{ABCD}$, and so on for products of more than two factors. Thus any product appears as a γ with an even number of suffixes. There are two general rules: (i) any two different suffixes may be interchanged, if one brings in the factor -1. (ii) A suffix A occurring in two consecutive positions may be suppressed but one must then bring in the factor -1 for $A = 4, 5, 6$."

For example, $\gamma^{12}\gamma^{13} = \gamma^{1213} = -\gamma^{1123} = -\gamma^{23}$; $\gamma^{24}\gamma^{34} = \gamma^{2434} = -\gamma^{2344} = \gamma^{23}$. As a consequence of these rules, $(\gamma^{AB})^2 = -\gamma_0$ if γ^{AB} is skew, $(\gamma^{AB})^2 = \gamma_0$ if γ^{AB} is symmetric. Also since $(\gamma^{123456})^2 = \gamma_0$ and γ^{123456} commutes with all γ^{AB}, we may choose $\gamma^{123456} = \pm\gamma_0$. Therefore in this six-dimensional Minkowski spacetime M_6, using a right-handed Cartesian coordinate system in which the metric tensor g_{AB} has components $g_{AB} = \text{diag}(1, 1, 1, -1, -1, -1)$, let us define

$$\gamma^{123456} = +\gamma_0. \quad (35)$$

Because of the multiplication rules we find that

$$\gamma^{AB}\gamma^{CD} + \gamma^{CD}\gamma^{AB} = 2\gamma_0(g^{AD}g^{BC} - g^{AC}g^{BD}) - \epsilon^{ABCDEF}g_{EG}g_{FH}\gamma^{GH}, \quad (36)$$

and

$$\gamma^{AB}\gamma^{CD} - \gamma^{CD}\gamma^{AB} = -2(g^{AC}\gamma^{BD} - g^{AD}\gamma^{BC} - g^{BC}\gamma^{AD} + g^{BD}\gamma^{AC}). \quad (37)$$

It follows from Eq. (36) that the γ^{AB} are trace-free, and Eq. (37) shows that $-\frac{1}{2}\gamma^{AB}$ generate a representation of the proper orthochronous Lorentz group on M_6. Upon contracting Eq. (36) with ϵ_{ABCDRS} we find that

$$\gamma_{RS} = -(1/4!)\epsilon_{ABCDRS}\gamma^{AB}\gamma^{CD}, \quad (38)$$

where $\gamma_{RS} = g_{AR}g_{BS}\gamma^{AB}$. This identity is used to reduce a γ matrix with four distinct indices to a γ matrix with two distinct indices. For example

$$\gamma_{36} = -(1/4!)\epsilon_{ABCD36}\gamma^{AB}\gamma^{CE} = -\epsilon_{124536}\gamma^{1245} = -\gamma^{1245} = -\gamma^{36}.$$

In [7] Dirac proved a simple but extremely useful pair of identities:

Lemma. For any symmetrical 4×4 matrix S,

$$\gamma^{12}S\gamma^{12} + \gamma^{23}S\gamma^{23} + \gamma^{31}S\gamma^{31} = S - \gamma_0\text{tr}S, \quad (39)$$

and similarly,

$$\gamma^{45}S\gamma^{45} + \gamma^{56}S\gamma^{56} + \gamma^{64}S\gamma^{64} = S - \gamma_0\text{tr}S. \quad (40)$$

A very useful identity can be obtained from Eq. (39). Since $\gamma^{34}(\gamma^{14}S\gamma^{14} + \gamma^{24}S\gamma^{24})\gamma^{34} = -\gamma^{31}S\gamma^{31} - \gamma^{23}S\gamma^{23} = \gamma^{12}S\gamma^{12} - S + \gamma_0\text{tr}S$, we have $\gamma^{14}S\gamma^{14} + \gamma^{24}S\gamma^{24} = \gamma^{34}\{\gamma^{12}S\gamma^{12} - S + \gamma_0\text{tr}S\}\gamma^{34} = \gamma^{56}S\gamma^{56} - \gamma^{34}S\gamma^{34} + \gamma_0\text{tr}S$, so that $\gamma^{j4}S\gamma^{j4} = \gamma^{56}S\gamma^{56} + \gamma_0\text{tr}S$. Because S is an arbitrary symmetric matrix this equation implies that

$$\gamma^{j4}_{pq}\gamma^{j4}_{rs} + \gamma^{j4}_{pr}\gamma^{j4}_{qs} = \gamma^{56}_{pq}\gamma^{56}_{rs} + \gamma^{56}_{pr}\gamma^{56}_{qs} + 2\delta_{ps}\delta_{qr}. \quad (41)$$

Holding s fixed one may obtain two similar equations by cyclically permuting (p, q, r). Upon adding two of these equations and subtracting the third one finds that

$$\gamma^{h4}_{pq}\gamma^{h4}_{rs} = -\delta_{pq}\delta_{rs} + \delta_{pr}\delta_{qs} + \delta_{ps}\delta_{qr} + \gamma^{56}_{pr}\gamma^{56}_{qs} + \gamma^{56}_{ps}\gamma^{56}_{qr}. \quad (42)$$

This identity is also valid for any cyclic permutation of (4, 5, 6), and under the replacement $(4, 5, 6)\rightarrow(1, 2, 3)$, $h\rightarrow\dot{h}$. Other similar identities may be obtained in this fashion, one of which is

$$\tfrac{1}{2}\epsilon^{jhk}\gamma^{hk}_{pq}\gamma^{j4}_{rs} = \delta_{rs}\gamma^{56}_{pq} + \delta_{pr}\gamma^{56}_{qs} - \delta_{qs}\gamma^{56}_{pr} + \delta_{ps}\gamma^{56}_{qr} - \delta_{qr}\gamma^{56}_{ps}. \quad (43)$$

6. A PARTICULAR REPRESENTATION OF C_6: SOME IDENTITIES

Let $\epsilon = \gamma^{45}$, $\gamma^{\alpha} = \gamma^{\alpha 6}$, and $\gamma^5 = \gamma^{56}$; the γ^{AB} are as in the previous section. All quantities are real.

We choose an irreducible Weyl representation for the Γ^A:

$$\Gamma^{\alpha} = \begin{pmatrix} 0 & \gamma^{\alpha}\epsilon \\ -\epsilon\gamma^{\alpha} & 0 \end{pmatrix}, \quad (44)$$

$$\Gamma^{5} = \begin{pmatrix} 0 & \gamma^{5}\epsilon \\ -\epsilon\gamma^{5} & 0 \end{pmatrix}, \quad (45)$$

$$\Gamma^{6} = \begin{pmatrix} 0 & -\epsilon \\ -\epsilon & 0 \end{pmatrix}; \quad (46)$$

then

$$\Gamma^{7} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (47)$$

and we take

$$\Omega = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (48)$$

The M^{AB} are in this representation

$$M^{AB} = -\tfrac{1}{2}\begin{pmatrix} \gamma^{AB} & \\ & -\tilde{\gamma}^{AB} \end{pmatrix}. \tag{49}$$

The $M^{\alpha\beta}$ may also be expressed as

$$M^{\alpha\beta} = \begin{pmatrix} S^{\alpha\beta} & \\ & -\tilde{S}^{\alpha\beta} \end{pmatrix}, \tag{50}$$

where $S^{\alpha\beta} = -\tfrac{1}{4}[\gamma^\alpha, \gamma^\beta]$.

In this representation of the Γ^A the m^{AB} are given by

$$m^{AB} = \tfrac{1}{2}\tilde{\xi}\gamma^{AB}\lambda. \tag{51}$$

Let us briefly consider the form of the matrices M defined by Eq. (2). Under SO (3, 1), $x \to x' = e^{-\omega}x$ (passive viewpoint) and $\psi \to \psi' = M\psi$; we see from Eqs. (7) and (50) that

$$M = \begin{pmatrix} S & \\ & \tilde{S}^{-1} \end{pmatrix}, \tag{52}$$

where $\omega \longleftrightarrow \omega^\alpha{}_\beta$ and $S = \exp(\tfrac{1}{2}\omega_{\alpha\beta}S^{\alpha\beta})$. $\psi \to \psi' = M\psi$ reads

$$\lambda \to \lambda' = S\lambda,$$

or

$$\lambda'_p = S_{pq}\lambda_q; \tag{53}$$

and

$$\xi \to \xi' = \tilde{S}^{-1}\xi,$$
$$\tilde{\xi}' = \tilde{\xi}S^{-1},$$

or

$$\xi'_p = \xi_q S^{-1}_{qp} = \tilde{S}^{-1}_{pq}\xi_q. \tag{54}$$

Under the parity operation $x^j \to -x^j$ and $x^4 \to x^4$, and

$$M_P = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}; \tag{55}$$

for a time reversal transformation $x^j \to x^j$, $x^4 \to -x^4$, and

$$M_T = \begin{pmatrix} & \gamma^5 \\ -\gamma^5 & \end{pmatrix}. \tag{56}$$

We now state a basic identity,

$$m^{AE}m^B{}_E = -g^{AB}(\tfrac{1}{2}\tilde{\xi}\lambda)^2. \tag{57}$$

We omit the proof since it is not particularly instructive and somewhat tedious.

In covariant notation

$$\tilde{\xi}\lambda = \tfrac{1}{2}\tilde{\psi}\Omega\Gamma^7\psi. \tag{58}$$

Since $\{m^{AB}, \tilde{\psi}\Omega\Gamma^7\psi\} = 0$, one might expect to obtain some additional information upon taking the Lie bracket of m^{RS} with Eq. (57). However the identity arising from the calculation of this bracket vanishes trivially due to Eq. (57).

7. THE TETRAD

Define

$$e^A_{(1)} = m^{A5} = m^{AB}\delta^5{}_B \tag{59}$$

and

$$e^A_{(2)} = m^{A6} = m^{AB}\delta^6{}_B, \tag{60}$$

which are, respectively, the inner products of m^{AB} with the timelike unit vectors $\delta^A{}_5$ and $\delta^A{}_6$. In this frame $e^5_{(1)} = 0 = e^6_{(2)}$. By Eq. (57), $e^A_{(1)}e_{(1)A} = e^\alpha_{(1)}e_{(1)\alpha} - N^2 = (\tilde{\xi}\lambda/2)^2 = e^A_{(2)}e_{(2)A} = e^\alpha_{(2)}e_{(2)\alpha} - N^2$, so that

$$e^\alpha_{(1)}e_{(1)\alpha} = e^\alpha_{(2)}e_{(2)\alpha} = N^2 + (\tilde{\xi}\lambda/2)^2$$
$$= (\tilde{\xi}\gamma^5\lambda/2)^2 + (\tilde{\xi}\lambda/2)^2.^8 \tag{61}$$

Further

$$e^A_{(1)}e_{(2)A} = e^\alpha_{(1)}e_{(2)\alpha} = 0, \tag{62}$$

which also follows from Eq. (57). Therefore whenever $\tilde{\xi}\lambda$ and $\tilde{\xi}\gamma^5\lambda$ are not both zero $e^\alpha_{(1)}$ and $e^\alpha_{(2)}$ are linearly independent orthogonal spacelike SO (3, 1) vectors.

Let

$$e^\alpha_{(4)} = -\tfrac{1}{4}\tilde{\psi}\Gamma^4\Gamma^\alpha\psi \tag{63}$$

and

$$e^\alpha_{(3)} = -\tfrac{1}{4}\tilde{\psi}\Gamma^4\Gamma^\alpha\Gamma^7\psi. \tag{64}$$

Since

$$\Gamma^4\Gamma^\alpha = -\tilde{\Gamma}^\alpha\Gamma^4, \tag{65}$$

$$\Gamma^4 = \tilde{M}\Gamma^4 M \tag{66}$$

under SO (3, 1). Thus $e^\alpha_{(3)}$ and $e^\alpha_{(4)}$ transform as vectors under SO (3, 1). Under a reflection about the x^α axis,

$$g^{\alpha\alpha}\Gamma^4 = \widetilde{\Gamma^\alpha\Gamma^7}\Gamma^4\Gamma^\alpha\Gamma^7 \text{ (no sum on } \alpha), \tag{67}$$

so that they transform as vectors under spatial reflections, but pick up a minus sign under time reversal.

Using Eqs. (44) and (47) we find that

$$e^\alpha_{(4)} = \tfrac{1}{4}(n^\alpha + l^\alpha) \tag{68}$$

and

$$e^\alpha_{(3)} = \tfrac{1}{4}(n^\alpha - l^\alpha), \tag{69}$$

where

$$n^\alpha = -\tilde{\lambda}\gamma^4\gamma^\alpha\lambda \tag{70}$$

and

$$l^\alpha = -\tilde{\xi}\gamma^\alpha\gamma^4\xi. \tag{71}$$

n^α and l^α transform as vectors under SO (3, 1) since $\gamma^4\gamma^\alpha = -\tilde{\gamma}^\alpha\gamma^4 \Rightarrow \gamma^4 = \tilde{S}\gamma^4 S$ under the SO (3, 1) transformation matrix S of Eqs. (52)–(54). Moreover, n^α and l^α are both null vectors, independent of the values of λ and ξ. Consider for example n^α; we have that $n^j n_j = \tilde{\lambda}\gamma^4\gamma^j\lambda\tilde{\lambda}\gamma^4\gamma_j\lambda = \lambda_p\gamma^{j4}_{pq}\lambda_q\lambda_r\gamma^{j4}_{rs}\lambda_s$. Substituting Eq. (42) for $\gamma^{j4}_{pq}\gamma^{j4}_{rs}$ yields $n^j n_j = (\tilde{\lambda}\lambda)^2 + 2(\tilde{\lambda}\gamma^5\lambda)^2 = (\tilde{\lambda}\lambda)^2$, since $\tilde{\lambda}\gamma^5\lambda = 0$ due to $\tilde{\gamma}^5 = -\gamma^5$. We also have $n^4 = -\tilde{\lambda}\gamma^4\gamma^4\lambda = \tilde{\lambda}\lambda$, so that $n^\alpha n_\alpha = (\tilde{\lambda}\lambda)^2 - (\tilde{\lambda}\lambda)^2 = 0$.

The evaluation of $l^\alpha l_\alpha$ is similar. Thus

$$n^\alpha n_\alpha = 0 = l^\alpha l_\alpha. \tag{72}$$

Further, using Eq. (42) once again, one finds that

$$n^\alpha l_\alpha = -2\{(\tilde{\xi}\lambda)^2 + (\tilde{\xi}\gamma^5\lambda)^2\}. \tag{73}$$

Therefore when $n^\alpha l_\alpha \neq 0$, $e^\alpha_{(4)}$ is the sum of two linearly independent null vectors, and is timelike (because $n^4 > 0$ and $l^4 > 0$ for $\lambda \neq 0 \neq \xi$); moreover, $e^\alpha_{(3)}$ is spacelike and orthogonal to $e^\alpha_{(4)}$:

$$e^\alpha_{(3)}e_{(4)\alpha} = 0. \tag{74}$$

$e^\alpha_{(3)}$ and $e^\alpha_{(4)}$ are orthogonal to both $e^\alpha_{(1)}$ and $e^\alpha_{(2)}$ because

$$n_\alpha e^\alpha_{(1)} = 0 = n_\alpha e^\alpha_{(2)} = l_\alpha e^\alpha_{(1)} = l_\alpha e^\alpha_{(2)}. \tag{75}$$

One may verify $n_\alpha e^\alpha_{(1)} = 0$, for instance, by contracting Eq. (42) with γ^{54}_{st} and using the resulting identity along with Eqs. (51), (59), and (70) to evaluate $n_\alpha e^\alpha_{(1)}$.

Summarizing: we have constructed four vectors $e^\alpha_{(\mu)}$ from the spinor ψ, which when ψ is nondegenerate (to be discussed below) are mutually orthogonal: $e^\alpha_{(\mu)} e_{(\nu)\alpha} = -\frac{1}{8} n^\alpha l_\alpha \eta_{(\mu)(\nu)}$, where $\eta_{(\mu)(\nu)} =$ diagonal $(1, 1, 1, -1) = \eta^{(\mu)(\nu)}$; the $e^\alpha_{(i)}$ are spacelike while $e^\alpha_{(4)}$ is timelike.

$\tilde{\xi}\lambda$ and $\tilde{\xi}\gamma^5\lambda$ are simultaneously zero whenever ψ is restricted to one of the subspaces of S_8 by the SO (3, 1) invariant equations

$$\epsilon\xi = -(\text{constant})\lambda \tag{76}$$

or

$$\gamma^4\xi = -(\text{constant})\lambda \tag{77}$$

(because then $\tilde{\xi}\lambda$ and $\tilde{\xi}\gamma^5\lambda =$ one of $\{\tilde{\lambda}\epsilon\lambda, \tilde{\lambda}\gamma^4\lambda\}\cdot$constant and each of these expressions vanishes identically due to the skew-symmetry of ϵ and γ^4). When the constant is one, Eq. (76) implies that $n^\alpha = l^\alpha = 2e^\alpha_{(1)}$ and $e^\alpha_{(2)} = 0$, while Eq. (77) implies that $n^\alpha = l^\alpha = 2e^\alpha_{(2)}$ and $e^\alpha_{(1)} = 0$. In both of these cases ψ no longer determines the four linearly independent vectors of a tetrad but instead defines a unique null vector. ψ is referred to as degenerate or *null* when its components are related by any combination of Eq. (76) and Eq. (77).

8. SPINORS AND PARTICLES; A GEOMETRICAL INTERPRETATION OF THE DIRAC EQUATION

At any point P in Minkowski spacetime M_4 we can construct an (in general unnormalized) orthogonal tetrad, denoted by $e^\alpha_{(\mu)}$, from a nondegenerate spinor $\psi\epsilon S_8(P)$. This tetrad comprises a basis of the tangent space at P, $T_P(M_4)$. Through the point P in M_4 we consider a timelike curve $x^\alpha = X^\alpha(s)$ (which can represent the world line of an electron) and define $\dot{X}^\alpha = (dX^\alpha/ds)\epsilon T_P(M_4)$ (evaluation at P is implicit). Here s is the arc length (proper time) along the curve, where $ds^2 = -g_{\alpha\beta}dX^\alpha dX^\beta$. We are especially interested in those tetrads for which $e^\alpha_{(4)}$ and $\dot{X}^\alpha$ are parallel. We can pick out a family of spinors $\psi\epsilon S_8(P)$ for which $e^\alpha_{(4)}$ and $\dot{X}^\alpha$ are parallel as follows. $\dot{X}^\alpha \propto e^\alpha_{(4)} \Rightarrow e^\alpha_{(1)}\dot{X}_\alpha = 0 = e^\alpha_{(2)}\dot{X}_\alpha$, and $e^\alpha_{(4)}\dot{X}_\alpha \neq 0$. Consider a rest frame in which $\dot{X}^\alpha = \delta^\alpha_{(4)}$. We require that $e^4_{(1)} = 0 = e^4_{(2)}$ in this frame and not both of $\tilde{\xi}\lambda$ and $\frac{1}{2}\tilde{\xi}\gamma^5\lambda = N$ be zero. But $e^4_{(1)} = \frac{1}{2}\tilde{\xi}\epsilon\lambda$ and $e^4_{(2)} = \frac{1}{2}\tilde{\xi}\gamma^4\lambda$, and if both of these expressions are to vanish while also ensuring that either $\tilde{\xi}\lambda$ or N is nonzero, then we must have either

$$\xi = -\gamma^5\lambda = \epsilon\gamma^\alpha\dot{X}_\alpha\lambda \tag{78}$$

or

$$\xi = \lambda = \gamma^4\gamma^\alpha\dot{X}_\alpha\lambda \tag{79}$$

in this rest frame, up to a scaling factor.

The question of whether Eq. (78) or (79) should be utilized in our construction is answered on purely physical grounds. If one is modeling a particle possessing an intrinsic magnetic moment, whose electric moment vanishes in the rest frame, Eq. (78) should be chosen. On the other hand, a model of a particle possessing only an electric moment in its rest frame should incorporate Eq. (79) as the spinor constraint. This can be seen as follows. The electromagnetic moments $\mu_{\alpha\beta} = -\mu_{\beta\alpha}$ of an "arbitrary particle" are linear combinations of the components of a spin tensor $\Sigma^{\alpha\beta} = -\Sigma^{\beta\alpha}$, namely, $\mu_{\alpha\beta} = \frac{1}{2}(I_\alpha{}^\mu\Sigma_{\mu\beta} - I_\beta{}^\mu\Sigma_{\mu\alpha})$. For an electron, and we assume also for an elementary "electric moment particle," $I_\alpha{}^\mu$ assume the simple form $I_\alpha{}^\mu \propto \delta_\alpha{}^\mu$. For an electron the intrinsic electric moment should vanish in the rest frame, thus $\Sigma^\alpha{}_\beta\dot{X}^\beta = 0$. (This is the well-known Frenkel condition.) For the "electric moment particle" the intrinsic magnetic moment should vanish in the rest frame, $\epsilon^{\alpha\beta\mu\nu}\Sigma_{\beta\mu}\dot{X}_\nu = 0$. Now by Eqs. (30) and (51), $\Sigma^{\alpha\beta} = \frac{1}{2}\tilde{\xi}\gamma^{\alpha\beta}\lambda$. If we evaluate $\Sigma^{\alpha\beta}$ using Eq. (78), supposing that we are in a rest frame, we find that $\Sigma^{j4} = 0$, $\Sigma^{jk} \neq 0$; alternatively, if we use Eq. (79) we find that $\Sigma^{j4} \neq 0$, $\Sigma^{jk} = 0$. We shall restrict our attention in this section to the electron and shall work with Eq. (78) as the spinor constraint. As shown above this implies the Frenkel condition

$$\Sigma^\alpha{}_\beta\dot{X}^\beta = 0. \tag{80}$$

Expressed covariantly Eq. (78) is $(\Gamma_\alpha X^\alpha - \Gamma^7)\psi = 0$; however, this equation is invariant under rotations in the x^5-x^6 plane (such a rotation is generated by M^{56} and $[M^{56}, \Gamma_\alpha\dot{X}^\alpha - \Gamma^7] = 0$) whereas $e^\alpha_{(1)}$ and $e^\alpha_{(2)}$ are not $[e^\alpha_{(1)} - ie^\alpha_{(2)} \to \exp(i\omega_{56})(e^\alpha_{(1)} - ie^\alpha_{(2)})$ when $\psi \to \psi' = \exp(\omega_{56}M^{56})\psi]$. Actually Eq. (78) also possesses invariance under these rotations [given in this case by the chiral transformation $\exp(-\frac{1}{2}\omega_{56}\gamma^5)$] and so is invariant under an operation which is not a symmetry transformation of the tetrad. This undesirable feature of the constraint equation may be eliminated by recasting Eq. (78) in a more symmetrical form. Equation (78) implies that $\epsilon\xi = -\gamma_\alpha\dot{X}^\alpha\lambda$ and $\lambda = \gamma_\alpha\dot{X}^\alpha\epsilon\xi$, so that we may reformulate Eq. (78) in terms of $\lambda + i\epsilon\xi$ as

$$(i\gamma_\alpha\dot{X}^\alpha + 1)\phi = 0, \tag{81}$$

where

$$\phi = \tfrac{1}{2}(\lambda + i\epsilon\xi). \tag{82}$$

Equation (81) does not possess the unwanted symmetry and so is adopted as the spinor constraint equation. It should be emphasized that Eq. (82) is not manifestly covariant under O (3, 1) because the decomposition of ψ into $\psi = \binom{\lambda}{\xi}$ is preserved only under SO (3, 1), not O (3, 1). However, this equation is invariant under SO (3, 1) transformations since the matrix ϵ transforms as

$$\begin{aligned}\epsilon \to \epsilon' &= \tilde{S}\epsilon S \\ &= \epsilon,\end{aligned} \tag{83}$$

under SO (3, 1), where S is given by Eq. (52).

In terms of ϕ and $\bar{\phi}$, where

$$\bar{\phi} = \phi^\dagger\gamma^4, \tag{84}$$

$e^\alpha_{(4)}$ and $N = m^{56}$ can be simply expressed as

$$e^\alpha_{(4)} = -\bar{\phi}\gamma^\alpha\phi, \tag{85}$$

and

$$N = -i\bar{\phi}\phi, \tag{86}$$

where we have used Eqs. (68), (70), and (71) in arriving at Eq. (85).

Let us suppose that Eq. (81) holds and evaluate $e^{\alpha}_{(4)}$. Equation (81) implies that $\xi = \epsilon\gamma_\alpha \dot{X}^\alpha\lambda$ and $\lambda = \gamma_\alpha \dot{X}^\alpha\epsilon\xi$; upon substitution of these expressions into Eqs. (70) and (71) and then using Eq. (68) one finds that

$$e^{\alpha}_{(4)} = -\dot{X}^\alpha \dot{X}_\beta e^{\beta}_{(4)}, \tag{87}$$

as desired. Therefore when Eq. (81) holds $e^{\alpha}_{(4)}$ is parallel to $\dot{X}^\alpha$.

Equation (81) implies that $\bar{\xi}\lambda = 0$, but $N \neq 0$ for $\psi \neq 0$. Therefore by Eq. (73)

$$n^\alpha l_\alpha = -8N^2, \tag{88}$$

and by Eq. (61),

$$e^{\alpha}_{(1)} e_{(1)\alpha} = N^2 = e^{\alpha}_{(2)} e_{(2)\alpha}, \tag{89}$$

when ϕ satisfies Eq. (81). Also by Eq. (88),

$$e^{\alpha}_{(4)} e_{(4)\alpha} = -N^2 \tag{90}$$

and

$$e^{\alpha}_{(3)} e_{(3)\alpha} = N^2. \tag{91}$$

Hence $-N^{-1}e^{\alpha}_{(\mu)}$ is an *orthonormal* tetrad when ϕ satisfies Eq. (91). We note that $-N^{-1}e^{\alpha}_{(3,4)}[-\det(g_{\mu\nu})]^{1/2}$ and $-N^{-1}e^{\alpha}_{(1,2)}$ transform as vectors under *all* O (3, 1) transformations. Equations (57) and (81) imply that

$$\Sigma^{\alpha\beta} = N^{-1}(e^{\alpha}_{(1)} e^{\beta}_{(2)} - e^{\beta}_{(1)} e^{\alpha}_{(2)}). \tag{92}$$

Therefore when Eq. (81) holds $e^{\alpha}_{(3)}$ is normal to the two-plane determined by the spin tensor. Thus $e^{\alpha}_{(3)}$ may be identified with the Pauli-Lubanski spin vector.

If one considers the Dirac equation for a free electron, $(i\gamma^\alpha p_\alpha + m)\phi = 0$, then one would also arrive at Eq. (81) when the electron is in an eigenstate of momentum whose value is $p^\alpha = m\dot{X}^\alpha$. Therefore the Dirac equation not only puts the momentum of the electron on mass shell but also ensures that ψ (constructed from ϕ) is nondegenerate and aligns the tetrad constructed from ψ so that $e^{\alpha}_{(4)}$ is parallel to the momentum vector p^α.

The current vector density j^α in the Dirac theory is

$$j^\alpha = -\bar{\phi}\gamma^\alpha\phi = e^{\alpha}_{(4)}, \tag{93}$$

which by Eq. (68) is seen to be proportional to the sum of n^α and l^α. Thus the electron current naturally decomposes into the sum of two null currents. Further Eq. (82) may be interpreted as the statement that a physical electron state ϕ is the superposition of the amplitudes for the electron to be in the massless states λ and ξ. [Of course the particle that we have been calling an "electron" could actually be any spin-$\frac{1}{2}$ massive fermion satisfying Eq. (80), since the interactions that would allow us to further classify particles have not yet been introduced.]

A result analogous to Eq. (81) and (93) is true for wavefunctions $2\chi = \lambda + i\gamma^4\xi$ which describe particles satisfying $\epsilon^{\alpha\beta\mu\nu}\Sigma_{\beta\mu}\dot{X}_\nu = 0$. (It is an open question whether these particles are magnetic monopoles.) This suggests that nondegenerate spinors correspond to particles. However, not all of these particles are necessarily realized in nature, since an arbitrary spinor contains, in general, both "electron" and "magnetic monopole" parts.

The two Lorentz scalars $\bar{\xi}\lambda$ and $\bar{\xi}\gamma^5\lambda$ partially classify particles: $\bar{\xi}\lambda$ vanishes when one describes "electrons" using Eq. (78), while $\bar{\xi}\gamma^5\lambda$ vanishes when one describes "magnetic monopoles" using Eq. (79).

Equation (80) actually follows from the following two identities, which are valid for any ψ,

$$\Sigma^{\alpha\beta} n_\beta = -\tfrac{1}{2}\bar{\xi}\lambda n^\alpha, \tag{94}$$

and

$$\Sigma^{\alpha\beta} l_\beta = \tfrac{1}{2}\bar{\xi}\lambda l^\alpha, \tag{95}$$

and may be derived with the help of Eq. (42). Since Eq. (81) ensures that $\bar{\xi}\lambda = 0$ and $N \neq 0$, Eq. (80) follows.

Similarly one can show that, for any ψ,

$$\bar{\xi}\gamma^5 S^{\alpha\beta}\lambda n_\beta = Nn^\alpha, \tag{96}$$

and

$$\bar{\xi}\gamma^5 S^{\alpha\beta}\lambda l_\beta = -Nl^\alpha. \tag{97}$$

Since

$$\gamma^5 = -(1/4!)\epsilon_{\alpha\beta\mu\nu}\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu, \tag{98}$$

and

$$\gamma^5 S^{\alpha\beta} = -\tfrac{1}{2}g^{\alpha\mu}g^{\beta\nu}\epsilon_{\mu\nu\lambda\sigma}S^{\lambda\sigma}, \tag{99}$$

the analog of the Frenkel condition for elementary "electric dipole" particles follows when $N = 0$ and $\bar{\xi}\lambda \neq 0$.

We shall not address the question of electromagnetic duality, but clearly the formalism does not prefer "electrons" over "magnetic monopoles," or even "electrons" over, say, particles described by $\frac{1}{2}\lambda + \frac{i}{2}(\epsilon\cos\theta + \gamma^4\sin\theta)\xi = \frac{1}{2}\lambda + \frac{i}{2}(\cos\theta + \gamma^5\sin\theta)\epsilon\xi$. A complete discussion of this symmetry is probably best deferred until it can be discussed within a theory of interacting particles.

9. CONCLUSION

We may partition S_8 into rays (or "fibers") and write S_8 as $S_8 = \mathbb{P}_7(\mathbb{R})\times\mathbb{R}^* + \{0\}$, where $\mathbb{P}_7$ is a fibering of S_8 (a seven-dimensional projective space) by the equivalence relation $\psi \sim \psi'$ iff $\psi = \theta\psi'$, $\theta\in\mathbb{R}^*$. By scalar multiplication in the fibers we realize $\mathbb{R}^*$. Therefore $(S_8, \omega, \overline{\mathrm{SO}(3,3)}\times\mathbb{R}^*)$ satisfies requirements (i)-(iv) for $\sqrt{M_4}$. Furthermore the apparent correspondence between (at least some) spinors and particles leads one to hope that a physical theory cast on $\sqrt{M_4}$ will give rise to an operational definition of spacetime (i.e., M_4) via particles interacting with particles, thereby fulfilling the objective stated in the introduction.

A nondegenerate spinor defines an orthogonal tetrad; thus the set of nondegenerate spinors determines the space of orthogonal frames over M_4. This space is isomorphic to $\overline{\mathrm{SO}(3,1)}\times\mathbb{R}^*$. From our initial *ansatz* of $G = \overline{\mathrm{SO}(3,3)}\times\mathbb{R}^*$ we have arrived at SO (3,1)$\times\mathbb{R}^*$ by breaking the SO (3, 3) symmetry down to SO (3, 1). An interesting question is whether there is any physical significance to this symmetry breaking, and if the symmetry can be approximately restored in some (perhaps, high-temperature) limit.

[1]A. Einstein, *The Meaning of Relativity* (Princeton U. P. Princeton, New Jersey, 1956), p. 2.

[2]R. Penrose, "Twistor Theory: Its Aims and Achievements," in *Quantum Gravity*, edited by C.J. Isham, R. Penrose, and D.W. Sciama (Clarendon, Oxford, England, 1975), p. 268.

[3]B. Kostant, *Lecture Notes in Mathematics*, 170 (Springer, Berlin, 1970). J.-M. Souriau, *Structure des Systemes Dynamiques* (Dunod, Paris, 1970).

[4]Greek letters, upper case Latin, and the lower case Latin letters $\{h, i, j, k, l\}$ are spacetime indices. For pedagogical reasons Greek indices run from 1 to 4, although x^4, say, is real. Upper case Latin indices run from 1 to 6; the above specified lower case Latin indices take values from 1 to 3. A lower case Latin index with a dot above it takes values $\dot{1} = 4$, $\dot{2} = 5$, and $\dot{3} = 6$. An index in parentheses is a vector label. The lower case Latin letters $a, b, c, d = 1,...,8$ and $p, q, r, s, t = 1,...,4$ are reserved for spinor indices. The metric tensor on M_6 has components $g_{AB} = \mathrm{diag}\,(1, 1, 1, -1, -1, -1)$ and the metric tensor on M_4 has components $g_{\alpha\beta} = \mathrm{diag}\,(1, 1, 1, -1)$. ϵ_{ABCDRS} is the completely antisymmetric Levi-Civita tensor density of weight -1 in six dimensions; $\epsilon_{123456} = +1$.

[5]R. Brauer and H. Weyl, Am. J. Math. **57**, 425 (1935).

[6]Therefore one may define $\theta = \frac{1}{2}\Omega_{ab}\psi^a d\psi^b$ and $\omega = d\theta = \frac{1}{2}\Omega_{ab} d\psi^a \wedge d\psi^b$. With Ω given by Eq. (48), $\omega/2\pi$ defines an integral de Rham cohomology class.

[7]P.A.M. Dirac, J. Math. Phys. **4**, 901 (1963).

[8]It is important to note that Eq. (61), which ensures that $e^{\alpha}_{(1)}$ and $e^{\alpha}_{(2)}$ are both spacelike for all values of $\tilde{\xi}\lambda$, $\tilde{\xi}\gamma^5\lambda$, $\tilde{\xi}\gamma^{A5}\lambda = \tilde{\xi}\epsilon\lambda$, and $\tilde{\xi}\gamma^{A6}\lambda = \tilde{\xi}\gamma^4\lambda$, is only possible when the metric is $g_{AB} = \mathrm{diag}(1, 1, 1, -1, -1, -1)$, or equivalently, when the group is SO (3, 3). An attempt at a similar construction based on the group SO (4, 2), for example, cannot succeed.

On choosing the points in product integration [a)]

Ian H. Sloan [b)]
Department of Physics and Astronomy & Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

(Received 27 May 1979; accepted for publication 28 December 1979)

A product-integration rule for the integral $\int_a^b k(t)f(t)\,dt$ is a rule of the form $\Sigma_{i=1}^n w_i f(t_i)$, with the weights $w_1,\dots,w_n$ chosen so that the rule is exact if f is any linear combination of a chosen set of functions $\phi_1,\dots,\phi_n$. For some choices of $\{\phi_j\}$, including the polynomial case, the points $\{t_i\}$ need to be carefully chosen if reliable results are to be obtained. In this paper known convergence results for the polynomial case with well-chosen points are summarized and illustrated, and extended to some nonpolynomial cases, including one proposed by Y.E. Kim for use in solving the three-body Faddeev equations. The convergence theorems yield practical prescriptions for choosing the points $\{t_i\}$.

1. INTRODUCTION

Product integration[1,2] is a simple technique for handling integrals of the form

$$I(f) = \int_a^b k(t)f(t)\,dt\,, \tag{1.1}$$

where k is a singular function and f is smooth. The method has proved useful[3–8] in the integral equation formulation of the three-body problem and is likely to find many other applications in the future.

The aim of this paper is to make the method more useful in practice, by providing theoretically well-based prescriptions for choosing the quadrature points.

A product-integration rule for $I(f)$ is an expression of the form

$$I_n(f) = \sum_{i=1}^n w_i f(t_i)\,, \tag{1.2}$$

where $t_1,\dots,t_n$ are a set of distinct points in $[a,b]$, and $w_1,\dots,w_n$ are suitable weights. Note that the singular function k does not appear explicitly but rather has been incorporated into the weights. The weights are determined by requiring that the rule be exact if f is any linear combination of a chosen linearly independent set $\phi_1,\dots,\phi_n$. Thus the weights satisfy a system of n linear equations,

$$\sum_{i=1}^n \phi_j(t_i)w_i = \int_a^b k(t)\phi_j(t)\,dt\,, \quad j=1,\dots,n\,. \tag{1.3}$$

We shall assume that the $n\times n$ matrix $\{\phi_j(t_i)\}$ is nonsingular, so that the weights exist and are unique. (Note that t_i, w_i, and ϕ_i may all depend on n, but the dependence on n has been suppressed to simplify the notation.)

In the early formulations,[1,2] the interval $[a,b]$ was usually restricted to be finite, the points were usually taken to be equally spaced, and the functions ϕ_j were taken to be polynomials of degree $j-1$ [in which case the matrix $\{\phi_j(t_i)\}$ is automatically nonsingular]. However, many other choices are possible. Intuitively, the main requirement for the functions ϕ_j is that a suitable linear combination should be capable of accurately representing the function f. Much less obvious—and this is the central concern of this paper—is the corresponding requirement for the points t_i.

Many particular examples of polynomial product-integration methods have appeared in the literature,[9–15] often with special choices for the function k.

The application of product-integration to the numerical solution of integral equations of the second kind appears to have been first suggested by Young.[16] In that application, the function k in (1.1) is the kernel of the integral equation (with one of its variables suppressed), or perhaps just a singular factor of the kernel, and f is the solution of the integral equation, multiplied by any remaining factor of the kernel. In the applications to the three-body problem the functions ϕ_j have been variously taken to be polynomials,[6] piecewise polynomials,[5,7] or other functions.[3,4]

A convergence theorem for the numerical solution of integral equations of the second kind by product integration with piecewise polynomials has been obtained by Atkinson,[17] and one for the polynomial case by Sloan.[18] In the present work convergence questions are only considered for integrals, not integral equations (though the two questions are, of course, closely related).

How should the points t_i in the product-integration rule (1.2) be chosen? First, we may dispose of the case in which the functions ϕ_j are piecewise polynomials—to simplify the discussion, let us consider specifically the piecewise-linear case. Such rules arise in a natural way if the interval $[a,b]$ is divided into subintervals $[t_{i-1}, t_i]$, with f approximated on each subinterval by linear interpolation. Here the choice of the points t_i is not at all a delicate matter—the points may, if desired, be taken equally spaced, or may be concentrated more densely in regions where f varies most rapidly. It is easily proved (see the following section) that the approximate integral converges to the exact result as $n\to\infty$, provided the length of the longest subinterval converges to zero.

If, on the other hand, the function ϕ_j is taken to be a single polynomial of degree $j-1$ over the whole interval, then a proper choice of the quadrature points becomes very important. Most strikingly, if the points are taken to be

a) Work supported by the U.S. Department of Energy and the University of Maryland Computer Science Center.
b) Permanent address: Department of Applied Mathematics, University of New South Wales, Sydney, N.S.W. 2033, Australia.

equally spaced, then the results can be disastrously bad, even if f is a smooth function. (An example is given in Sec. 3.) Nevertheless, known convergence theorems for the polynomial case[10-12,14,19] tell us that there are some very good choices of points for this case. The theorems are summarized in Sec. 4. One good choice of points for the polynomial case, if the interval is taken as $[-1,1]$, is the set of Gauss quadrature points; another is

$$t_i = -\cos\left(\frac{2i-1}{2n}\pi\right), \quad i=1,...,n,$$

and another is

$$t_i = -\cos\left(\frac{i-1}{n-1}\pi\right), \quad i=1,...,n.$$

Numerical examples for these sets of points are given in Sec. 5.

If the choice of points is important for the polynomial case, then there are surely many nonpolynomial cases in which the choice of points is just as important. For some cases of this kind it is possible to obtain useful results by adapting the polynomial results.

An interesting example of this kind is a product-integration rule proposed by Kim[3] and used[4] in the numerical solution of the two-dimensional Faddeev[20] integral equations for the case of local two-nucleon interactions. In this application the integrals are from 0 to ∞, so that (1.1) becomes

$$I(f) = \int_0^\infty k(t)f(t)\,dt.$$

The functions ϕ_j in the product-integration method were taken by Kim to be

$$\phi_j(t) = \frac{1}{t+\alpha}\left(\frac{t}{t+\alpha}\right)^{j-1}, \quad j\geqslant 1,$$

where α is a suitable parameter. (The equations of Ref. 3 may be recovered by setting $t=p^2$.) How should the points t_i be chosen? In Sec. 6 we show that one choice with excellent convergence properties, in both theory and practice, is

$$t_i = \alpha\tan^2\left(\frac{2i-1}{4n}\pi\right), \quad i=1,...,n.$$

Finally, in Sec. 7 we consider a generalization of the previous paragraph: for an arbitrary $[a,b]$ (finite or infinite), we consider product integration, with the function ϕ_j taken to be

$$\phi_j(t) = g(t)p_{j-1}(h(t)), \quad j\geqslant 1,$$

where g and h are suitable real-valued functions which are supposed given, and p_{j-1} is a given polynomial of degree $j-1$. A convergence result for this case is established by transforming the problem into one of polynomial product integration.

2. PRODUCT INTEGRATION AS AN INTERPOLATORY METHOD

An alternative approach to the product-integration rule (1.2) is to derive it by replacing f in (1.1) by a suitable interpolating function, and then performing the integration exactly. Here we establish the equivalence of the two approaches and use the second approach to prove convergence in a simple case.

For a given n, let us suppose that the functions $\phi_1,...,\phi_n$ and the points $t_1,...,t_n$ are given, and that the $n\times n$ matrix $\{\phi_j(t_i)\}$ is nonsingular. Then if f is a continuous function on $[a,b]$, let L_n^f denote the unique linear combination of $\phi_1,...,\phi_n$ that interpolates f at $t_1,...,t_n$, i.e., that satisfies

$$L_n^f(t_i) = f(t_i), \quad i=1,...,n. \tag{2.1}$$

Then

$$L_n^f(t) = \sum_{j=1}^n a_j\phi_j(t), \tag{2.2}$$

where the coefficients a_j satisfy the linear equations

$$\sum_{j=1}^n \phi_j(t_i)a_j = f(t_i), \quad i=1,...,n,$$

the matrix of which, by assumption, is nonsingular.

Using (2.1), we may write the product-integration rule (1.2) as

$$I_n(f) = \sum_{i=1}^n w_i L_n^f(t_i) = \int_a^b k(t)L_n^f(t)\,dt, \tag{2.3}$$

with the last step following from the fact that the rule is exact for every linear combination of $\phi_1,...,\phi_n$, and therefore in particular for the linear combination L_n^f. Equation (2.3) tells us that the product-integration rule $I_n(f)$ may be obtained by replacing f in (1.1) by the interpolating function L_n^f and then integrating exactly.

In some cases (2.3) may be used to obtain a simple convergence proof, for if L_n^f converges uniformly to f, and if k is absolutely integrable, then the convergence of $I_n(f)$ to $I(f)$ follows immediately from (2.3) and (1.1). This method of proof is particularly convenient for piecewise-polynomial cases. For example, in the piecewise-linear case (in which each ϕ_j is continuous on the finite interval $[a,b]$ and linear on $[t_{i-1}, t_i]$, for $i=2,...,n$, with $a=t_1<t_2\cdots<t_n=b$), it is very clear the L_n^f converges uniformly to f for all continuous functions f, provided only that

$$\lim_{n\to\infty}\max_{2\leqslant i\leqslant n}|t_i - t_{i-1}| = 0.$$

For the polynomial case, on the other hand, it is *not* true that L_n^f converges uniformly to f for all continuous functions f:

TABLE I. Approximate integrals $I_n(f)$ for $f(t)=(1+25t^2)^{-1}$, $k(t)\equiv 1$, and equally spaced points.

n	$I_n(f)$	$\Sigma_{i=1}^n \lvert w_i\rvert$
6	0.46	2
11	0.93	6.13
16	0.83	1.67×10^1
21	−5.37	1.09×10^3
26	−5.40	3.54×10^3
31	153.8	4.24×10^5
exact	0.55	

for it is known[21] that no matter how the points $t_1,\ldots,t_n$ are chosen, there always exists a continuous function f such that L_n^f does not converge uniformly to f. Thus the polynomial case requires different techniques[10-12,19] if we are to prove the convergence of $I_n(f)$ to $I(f)$ for all continuous functions f.

3. THE POLYNOMIAL CASE WITH A BAD SET OF POINTS

We assume for the present that the interval is $[-1,1]$ and that the function ϕ_j is a polynomial of degree $j-1$. In principle ϕ_j can be any polynomial of degree $j-1$, but in practice it is usually wise to avoid the obvious choice

$$\phi_j(t) = t^{j-1}, \quad j \geqslant 1,$$

because it can lead to a badly conditioned matrix $\{\phi_j(t_i)\}_{j,i=1}^{n}$. A better choice for ϕ_j is the Legendre polynomial,

$$\phi_j(t) = P_{j-1}(t), \quad j \geqslant 1,$$

or the Chebyshev polynomial of the first kind,

$$\phi_j(t) = T_{j-1}(t), \quad j \geqslant 1, \tag{3.1}$$

defined by $T_{j-1}(\cos\theta) = \cos(j-1)\theta$. The Chebyshev polynomial basis has the advantage that for this case a technology already exists,[22] at least for many of the most common singular functions k, for evaluating recursively the integrals that are required on the right-hand side of (1.3). The latter basis has been used for the numerical calculations of this paper.

In this section we assume that the points t_i are equally spaced over the interval $[-1,1]$, i.e.,

$$t_i = -1 + 2(i-1)/(n-1), \quad i = 1, \ldots, n. \tag{3.2}$$

To show how bad this choice can be, it is sufficient to take k to be simply $k(t) \equiv 1$, so that the integral we are evaluating is merely

$$I(f) = \int_{-1}^{1} f(t)\, dt.$$

The results obtained by applying the equally-spaced quadrature rule to the function

$$f(t) = (1 + 25t^2)^{-1} \tag{3.3}$$

are shown for a number of values of n in Table I. Clearly, the approximate integrals $I_n(f)$ are spectacularly bad: they show no sign of convergence as n increases and sometimes even have the wrong sign.

A related problem, for all but the smallest values of n, is that some of the weights are negative. In the last column of Table I we show the sum of the absolute values of the weights and see that it grows without bound as n increases; yet the algebraic sum of the weights is just

$$\sum_{i=1}^{n} w_i = \int_{-1}^{1} dt = 2,$$

since the rule (1.2) is exact if f is the polynomial $f(t) \equiv 1$. Thus for the larger values of n we see that the weights have mixed signs and large absolute values, properties that lead to a serious loss of significance in $I_n(f)$ through cancellation. [However, the negative signs found for some values of $I_n(f)$ in Table I are considered to be genuine, and not to be caused by cancellation errors.]

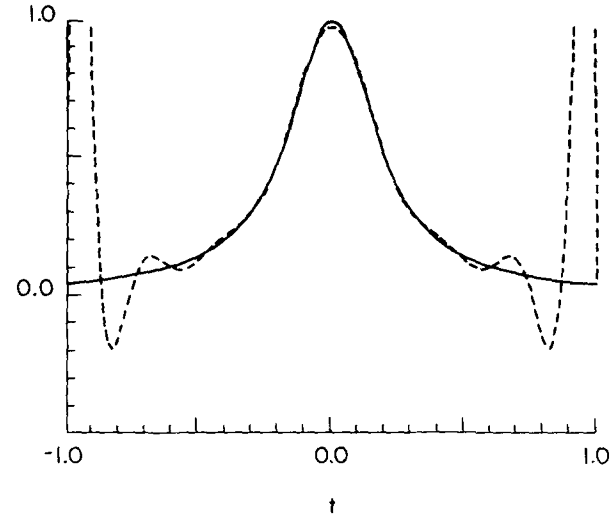

FIG. 1. The function $f(t) = (1 + 25t^2)^{-1}$ (solid curve), and its interpolating polynomial (dashed curve) for equally spaced points with $n = 16$.

One way of understanding the poor behavior of the integration rule in this case is to recall the interpretation of the previous section, that $I_n(f)$ results from integrating exactly the interpolating polynomial L_n^f. In Fig. 1 we show, for $n = 16$, the interpolating polynomial L_n^f for the function defined by (3.3). After seeing the wild oscillations of the interpolating polynomial near the ends of the interval, we are perhaps less surprised at the poor results in Table I. The function $f(t) = (1 + 25t^2)^{-1}$ is in fact a famous example for demonstrating the hazards of equally-spaced polynomial interpolation, apparently first introduced by Runge.

The integration rules discussed above are in fact well known in numerical analysis: they are the so-called Newton–Cotes rules.[23] It is known that the Newton–Cotes rules can diverge even for functions f that are infinitely differentiable on $[-1,1]$; even if they converge (as they do in fact if f is analytic in a sufficiently large region containing $[-1,1]$), they are beset, when n is large, with numerical instability problems arising from the occurrence of negative weights. For those reasons, the use of high-order Newton–Cotes rules is generally not recommended.[23] Equally, in product integration with polynomials (of which, after all, the Newton–Cotes rules are special cases), the use of equally spaced points should be avoided.

4. THE POLYNOMIAL CASE: GOOD CHOICES OF POINTS

If the points in the polynomial case should not be taken to be equally spaced, how then should they be chosen? If k is a classical non-negative weight function, then one answer[23] is to take the points to be the Gaussian points appropriate to the weight function k, since the product-integration rule is then simply a Gaussian rule. However, here we suppose that the Gaussian points are not known or are not suitable and we seek a set of points that does not depend on k. It is not obvious, perhaps, that a good set of points can exist independently of k.

Yet there do indeed exist some good choices of points for the polynomial case. The simplest, perhaps, is

$$t_i = -\cos\left(\frac{2i-1}{2n}\pi\right), \quad i=1,...,n, \tag{4.1}$$

or the closely related set that includes the end points,

$$t_i = -\cos\left(\frac{i-1}{n-1}\pi\right), \quad i=1,...,n. \tag{4.2}$$

For both of these choices it is known[10,11] that the quadrature rule $I_n(f)$ converges to the exact result for *every* continuous function f, if k satisfies the condition

$$\int_{-1}^{1} |k(t)|^p \, dt < \infty \tag{4.3}$$

for some $p > 1$. This condition is marginally stronger than the requirement that k be absolutely integrable, but in practice most integrable functions k also satisfy (4.3) for some $p > 1$. [For example, the barely integrable function $k(t) = |t|^{-0.99}$ satisfies (4.3) for all p in $1 < p < 1/0.99$.]

Another good set of points is the set of quadrature points for ordinary Gauss–Legendre quadrature, i.e.,

$$t_i = \xi_{ni}, \quad i=1,...,n, \tag{4.4}$$

where $P_n(\xi_{ni}) = 0$. For this set it is known[12,19] that $I_n(f)$ converges to the exact result for every continuous function f if k satisfies

$$\int_{-1}^{1} |k(t)(1-t^2)^{-1/4}|^p \, dt < \infty, \tag{4.5}$$

for some $p > 1$. The only difference from the condition (4.3) is that (4.5) is slightly more restrictive in the singularities of $k(t)$ that it allows at the ends of the interval.

Similar results are also known[12,19] for the more general case of product-integration rules based on the points

$$t_i = \xi_{ni}^{(\alpha,\beta)}, \quad i=1,...,n, \tag{4.6}$$

where the numbers $\xi_{ni}^{(\alpha,\beta)}$ are the zeros of the Jacobi polynomial $P_n^{(\alpha,\beta)}$, $\alpha, \beta > -1$. The previous choices of points (4.1) and (4.4) may be obtained by setting $\alpha = \beta = -\frac{1}{2}$ and $\alpha = \beta = 0$, respectively, and probably represent the most important special cases. In the general case the condition on k becomes[12]

$$\int_{-1}^{1} |k(t)(1-t)^{-\max[(2\alpha+1)/4,0]}(1+t)^{-\max[(2\beta+1)/4,0]}|^p \, dt < \infty, \tag{4.7}$$

for some $p > 1$.

For all of the above choices of points it is also known,[10-12] provided k satisfies the appropriate condition above, that the sum of the absolute values of the weights has a limit,

$$\lim_{n\to\infty} \sum_{i=1}^{n} |w_i| = \int_{-1}^{1} |k(t)| \, dt, \tag{4.8}$$

whereas it is easily seen that the algebraic sum of the weights is just

$$\sum_{i=1}^{n} w_i = \int_{-1}^{1} k(t) \, dt. \tag{4.9}$$

Thus the phenomenon seen in Table I, of a sum $\Sigma|w_i|$ that grows without bound, cannot now occur. Moreover, in the important special case in which $k(t)$ is non-negative, we deduce from (4.8) and (4.9) that

$$\lim_{n\to\infty} \sum_{i=1}^{n} |w_i| = \sum_{i=1}^{n} w_i;$$

thus the weights in this case have the property of "asymptotic positivity."

A rigorous bound on the error in the quadrature rule $I_n(f)$ is[10]

$$|I_n(f) - I(f)| \leqslant \left(\sum_{i=1}^{n} |w_i| + \int_{-1}^{1} |k(t)| \, dt\right) E_{n-1}(f), \tag{4.10}$$

where

$$E_{n-1}(f) = \min_{p\in\mathscr{P}_{n-1}} \max_{|t|\leqslant 1} |f(t) - p(t)|,$$

$\mathscr{P}_{n-1}$ being the set of polynomials of degree $\leqslant n-1$. The rate at which $E_{n-1}(f)$ converges to zero depends on the smoothness of f; one useful result[24] for a function f with s continuous derivatives is

$$E_{n-1}(f) \leqslant \left(\frac{\pi}{2}\right)^s \frac{\max|f^{(s)}(t)|}{n(n-1)\cdots(n-s+1)} = O(n^{-s}).$$

So far in this section we have assumed the function k to be at least integrable, and have thereby excluded the important case of the Cauchy principal-value integral,

$$k(t) = P(t-\lambda)^{-1}, \quad -1 < \lambda < 1. \tag{4.11}$$

Nevertheless, even in this case the convergence of $I_n(f)$ to $I(f)$ can be assured, provided we impose slightly stronger restrictions on f. In fact, if the points are given by (4.6) [or by the special cases (4.1) or (4.4)], then it is known[14,25] that $I_n(f)$ converges to $I(f)$ if f is Hölder continuous of order μ for some $\mu > 0$, i.e., if there exist numbers $M > 0$ and $\mu > 0$ such that

$$|f(t) - f(s)| \leqslant M|t-s|^{\mu}.$$

It appears that the corresponding result for the case of the points (4.2) has not yet been proved, though its validity can hardly be doubted. The Hölder continuity condition is in practice satisfied by most of the continuous functions that arise naturally.

Also of some interest for the applications to scattering theory is the case of the delta function

$$k(t) = \delta(t-\lambda), \quad -1 < \lambda < 1.$$

In the case $I(f) = f(\lambda)$ and $I_n(f) = L_n^f(\lambda)$, so that the question of the convergence of $I_n(f)$ to $I(f)$ reduces merely to the question of the pointwise convergence of the interpolating polynomial L_n^f at $t = \lambda$. For all of the choices of points discussed in this section, a sufficient condition for convergence is that discussed in the previous paragraph, namely that f be Hölder continuous of order μ for some $\mu > 0$. If the points are zeros of Jacobi polynomials, then the result follows directly from Theorem 14.4 of Ref. 26. An almost identical argument to that in Ref. 26 holds also for the points (4.2).

TABLE II. Quadrature errors for $f(t) = (1+25t^2)^{-1}$, $k(t)\equiv 1$.

n	$t_i = -\cos\left(\frac{2i-1}{2n}\pi\right)$	$t_i = -\cos\left(\frac{i-1}{n-1}\pi\right)$	$t_i = \xi_{ni}$
4	−0.21(0)	−0.30(0)	−0.18(0)
8	−0.50(−1)	−0.73(−1)	−0.41(−1)
16	−0.21(−2)	−0.32(−2)	−0.18(−2)
24	−0.88(−4)	−0.14(−3)	−0.74(−4)
32	−0.35(−5)	−0.56(−5)	−0.31(−5)

5. NUMERICAL RESULTS FOR THE POLYNOMIAL CASE

How good is the convergence for the polynomial case in practice? If the points are well chosen, and if f is reasonably smooth, then the following examples show that the convergence can be quite satisfactory.

First, in Table II we give results for the case already considered in Table I, of $k(t)\equiv 1, f(t) = (1+25t^2)^{-1}$. The results are given for the three sets of points (4.1), (4.2), and (4.4). Because $k(t)\equiv 1$, the product-integration rule for each of these sets of points reduces to a known quadrature rule[23]: they are, respectively, a quadrature rule of Fejér, the Clenshaw–Curtis rule, and the ordinary Gauss–Legendre rule. The weights in each rule are known to be positive.[23] The observed convergence rate in all three cases is satisfactory, and quite similar. The Gauss rule is the most accurate, as one might expect with $k(x)\equiv 1$, but the Fejér rule is a close second, with errors larger by less than 20%.

The vastly improved behavior compared with that seen in Table I can be understood through Fig. 2, where we again show the $n = 16$ interpolating polynomial for the function $f(t) = (1+25t^2)^{-1}$, this time with the points given by (4.1). Note that the interpolating polynomial no longer has the wild oscillations seen in Fig. 1 near the ends of the interval.

Next, in Table III we consider an example with a very singular function k, namely $k(t) = |t-0.8|^{-3/4}$, and $f(t) = (1.2-t)^{-1}$, for which the exact value is 15.695... . The integrals on the right-hand side of (1.3), with $\phi_j = T_{j-1}$, were evaluated by the recursive technique of Ref. 22. The three choices of points in the table are the same as in Table II. Evidently, the rate of convergence is very satisfactory in all three cases. The Fejér and Gauss–Legendre points (i.e., the first and third cases) here give comparable accuracy; again they give slightly better accuracy than the choice that includes the end points.

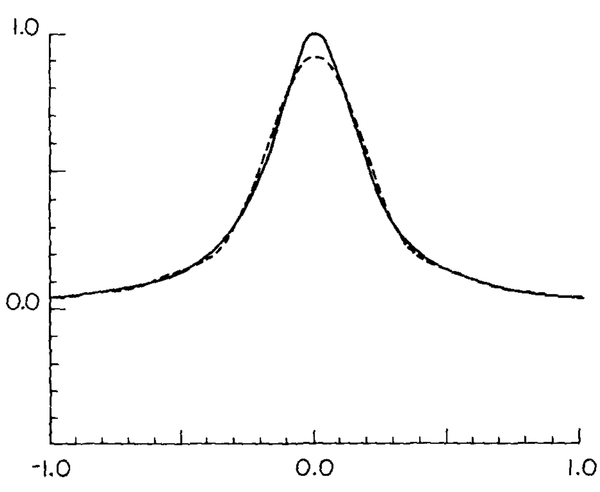

FIG. 2. The function $f(t) = (1+25t^2)^{-1}$ (solid curve), and its interpolating polynomial (dashed curve) for the points $t_i = -\cos[(2i-1)\pi/2n]$ with $n = 16$.

In Table IV we show the ratio $\Sigma|w_i|/\int|k(t)|\,dt$ for the function k of the preceding paragraph, namely $k(t) = |t-0.8|^{-3/4}$. The three choices of points are as in Table III. Because k satisfies the conditions (4.3) and (4.5) (for $1<p<4/3$) we recall from (4.8) that the ratio for each choice of points necessarily converges to 1 as $n\to\infty$. Obviously the behavior in the table is consistent with that, though the convergence is far from monotonic. What is more important in practice is that the ratio never departs very far from the limiting value 1, so that the loss of significance due to negative weights is entirely negligible for all values of n—a situation in marked contrast to that seen for the equally-spaced case in Table I.

Other numerical examples for the polynomial case are given in Refs. 10 and 11.

6. KIM'S PRODUCT-INTEGRATION RULE

Kim,[3] in an interesting application of product integration to the three-body problem, proposed that infinite integrals of the form

$$I = \int_0^\infty k(t)f(t)\,dt, \tag{6.1}$$

where f is smooth, be evaluated by product integration, with the functions ϕ_j of Eq. (1.3) taken to be

$$\phi_j(t) = \frac{1}{t+\alpha}\left(\frac{t}{t+\alpha}\right)^{j-1}, \quad j\geqslant 1, \tag{6.2}$$

where $\alpha>0$ is a suitable scaling parameter. [Obviously that procedure is sensible only if $f(t) = O(t^{-1})$ as $t\to\infty$ and if $\int_0^\infty k(t)(t+\alpha)^{-1}\,dt<\infty$; both properties hold in Kim's application.]

How should the product-integration points t_i be chosen in this case? We may obtain rigorous convergence results for at least some special choices of points by transforming the problem into one of polynomial product integration. The first step is to rewrite the integral (6.1) as

$$I = \int_0^\infty K(t)F(t)\,dt, \tag{6.3}$$

TABLE III. Quadrature errors for $f(t) = (1.2 - t)^{-1}$, $k(t) = |t - 0.8|^{-3/4}$.

n	$t_i = -\cos\left(\frac{2i-1}{2n}\pi\right)$	$t_i = -\cos\left(\frac{i-1}{n-1}\pi\right)$	$t_i = \xi_{ni}$
4	0.15(1)	0.30(1)	0.46(0)
8	−0.44(−1)	−0.21(0)	0.74(−2)
16	0.46(−3)	−0.29(−3)	0.49(−3)
24	0.42(−5)	0.58(−5)	0.24(−5)
32	0.47(−8)	0.42(−7)	−0.64(−8)

where

$$K(t) = k(t)(t + \alpha)^{-1},$$

$$F(t) = f(t)(t + \alpha).$$

Kim's product-integration rule can be transformed in a similar way: it can be expressed as

$$I_n = \sum_{i=1}^{n} W_i F(t_i),$$

with the weights W_i determined by requiring the rule to equal the integral in (6.3) if F is any linear combination of $\Phi_1,...,\Phi_n$, where

$$\Phi_j(t) = \phi_j(t)(t+\alpha) = [t/(t+\alpha)]^{j-1}, \quad j \geqslant 1.$$

Now we observe that Φ_j is a polynomial of degree $j-1$ in the variable

$$x = t/(t + \alpha), \quad 0 \leqslant x < 1,$$

or equivalently, in the variable

$$z = 2x - 1 = (t - \alpha)/(t + \alpha), \quad -1 \leqslant z < 1. \tag{6.4}$$

Written in terms of z, the integral (6.3) becomes

$$I = \int_{-1}^{1} \tilde{K}(z)\tilde{F}(z)\,dz, \tag{6.5}$$

where

$$\tilde{K}(z) = 2\alpha(1 - z)^{-2} K[\alpha(1 + z)/(1 - z)],$$

$$\tilde{F}(z) = F[\alpha(1 + z)/(1 - z)].$$

The above product-integration rule can also be transformed in a similar way: regarded as a product-integration rule for (6.5) it can be written as

$$I_n = \sum_{i=1}^{n} W_i \tilde{F}(z_i),$$

where the weights W_i are determined by requiring the rule to be exact if $\tilde{F}$ is any linear combination of $\tilde{\Phi}_1,...,\tilde{\Phi}_n$, where

$$\tilde{\Phi}_j(z) = \Phi_j\left(\alpha\frac{1+z}{1-z}\right) = \left(\frac{1+z}{2}\right)^{j-1}.$$

Since $\tilde{\Phi}_j$ is a polynomial of degree $j-1$, we may simply say that the weights W_i are determined by the requirement that the rule be exact if $\tilde{F}$ is any polynomial of degree $\leqslant n-1$.

Since the problem is now reduced to one of polynomial product-integration, we may use the theoretical results of Sec. 4. To avoid uninteresting complications, we consider only a single choice of product-integration points for (6.5), namely the classical Chebyshev set (4.1), i.e.,

$$z_i = -\cos\left(\frac{2i-1}{2n}\pi\right), \quad i = 1, ..., n. \tag{6.6}$$

In terms of the original variable t, related to z by (6.4), the expression for the points becomes

$$t_i = \alpha\tan^2\left(\frac{2i-1}{4n}\pi\right), \quad i = 1, ..., n. \tag{6.7}$$

With the points given by (6.6), the condition to be satisfied by $\tilde{K}$ is [see (4.3)]

$$\int_{-1}^{1} |\tilde{K}(z)|^p\,dz < \infty,$$

for some $p > 1$, or, equivalently,

$$\int_0^\infty \left|\frac{k(t)}{t+\alpha}\right|^p (t+\alpha)^{2(p-1)}\,dt < \infty, \tag{6.8}$$

for some $p > 1$. A sufficient condition for (6.8) to be satisfied is that

$$\int_0^T |k(t)|^p\,dt < \infty,$$

for some finite upper limit T and some $p > 1$, together with

$$|k(t)| \leqslant Bt^{-\epsilon}$$

for $B, \epsilon > 0$ and all $t \geqslant T$.

Under the condition (6.8), we are assured by the theorems summarized in Sec. 4 that Kim's product-integration

TABLE IV. $\Sigma|w_i|/\int|k|$ for $k(t) = |t - 0.8|^{-3/4}$.

n	$t_i = -\cos\left(\frac{2i-1}{2n}\pi\right)$	$t_i = -\cos\left(\frac{i-1}{n-1}\pi\right)$	$t_i = \xi_{ni}$
4	1.02	1	1
8	1	1.21	1
16	1	1.02	1.10
24	1.09	1.13	1.02
32	1	1.16	1

TABLE V. Kim's method for the integral (6.9), using the quadrature points (6.7).

n	Error	$\Sigma\lvert w_i\rvert/\int\lvert k\rvert$
2	−0.25(−2)	1
4	−0.50(−3)	1
6	0.74(−7)	1
8	−0.68(−6)	1
10	0.58(−8)	1
12	−0.17(−8)	1

rule with the points (6.7) converges to the exact result (6.1) if the function F is continuous, or, equivalently, if $f(t)(t+\alpha)$ is a continuous function on $[0,\infty)$ and has a finite limit as $t\to\infty$. (Of course the convergence is much faster if that function is not merely continuous but also smooth.)

It also follows from the results in Sec. 4 that the convergence property is valid even if k contains a principal-value or delta function singularity, provided $f(t)(t+\alpha)$ is not merely continuous but also Hölder continuous of order μ for some $\mu>0$.

For a numerical test of Kim's product-integration rule with the points (6.7), we consider

$$I=\int_0^\infty |t-1|^{1/2}(t+1)^{-5/2}\ \frac{t}{3t+1}\,dt, \tag{6.9}$$

the exact value of which is $0.145\cdots$. We choose

$$k(t)=|t-1|^{1/2}(t+1)^{-3/2},$$

$$f(t)=t/(t+1)(3t+1),$$

so that the conditions on k and f are satisfied, and take $\alpha=1$. Note that $f(t)(t+1)$ is a smooth function, so that we would expect the convergence to be rapid.

The results obtained for this example are shown in Table V. Note that the convergence is indeed rapid, and also that the weights all turn out to be positive for the values of n considered in the table.

A final comment about Kim's product-integration method concerns numerical stability: whether used with the points (6.7) or any other set, the basis (6.2) can lead to a very poorly conditioned matrix $\{\phi_j(t_i)\}$. A mathematically equivalent basis set with much better stability properties is

$$\phi_j(t)=\frac{1}{t+\alpha}\,T_{j-1}\left(\frac{t-\alpha}{t+\alpha}\right),\quad j\geqslant 1.$$

7. A GENERALIZATION

The argument of the previous section is here generalized to integrals of the form

$$I=\int_a^b k(t)f(t)\,dt, \tag{7.1}$$

where a and b may be finite or infinite, and the basis functions ϕ_j are of the form

$$\phi_j(t)=g(t)\,p_{j-1}(h(t)), \tag{7.2}$$

with g being a given positive function on (a,b), p_{j-1} a polynomial of degree $j-1$, and h a differentiable, monotonically increasing function that maps (a,b) onto $(-1,1)$.

By analogy with the argument of the previous section, the product-integration points are taken to be

$$t_i=h^{-1}\left[-\cos\left(\frac{2i-1}{2n}\pi\right)\right],\quad i=1,\dots,n, \tag{7.3}$$

where h^{-1} denotes the inverse function of h. The condition to be satisfied by k is

$$\int_a^b |k(t)g(t)|^p[h'(t)]^{-(p-1)}\,dt<\infty, \tag{7.4}$$

for some $p>1$.

If (7.4) is satisfied, then it follows, as in the previous section, that the product-integration rule (1.2) converges to the exact result as $n\to\infty$ if $f(t)/g(t)$ is continuous on (a,b) and if it also has finite limits at a and b. On the other hand, if (7.4) is not satisfied because of principal-value or delta-function singularities in k, then the convergence still holds if $f(t)/g(t)$ is not merely continuous but also Hölder continuous of order μ for some $\mu>0$.

As a simple example, let us take the interval $[a,b]$ to be $[0,\pi]$ and the functions ϕ_j to be

$$\phi_j(t)=\cos(j-1)t,\quad j\geqslant 1.$$

Then $\phi_j(t)$ is of the form (7.2), with $g(t)=1$ and $h(t)=-\cos t$. The above results then tell us that if the points are given by

$$t_i=[(2i-1)/2n]\pi,\quad i=1,\dots,n,$$

and if k satisfies

$$\int_0^\pi |k(t)|^p(\sin t)^{-(p-1)}\,dt<\infty, \tag{7.5}$$

for some $p>1$, then the product-integration rule (1.2) converges to the exact result for all functions f that are continuous on $[0,\pi]$; or, if (7.5) is not satisfied because of the occurrence of principal-value or delta-function singularities, then the product-integration rule converges for all functions f that are Hölder continuous of order μ for some $\mu>0$.

The condition (7.5) is in fact equivalent to an apparently simpler condition,

$$\int_0^\pi |k(t)|^{p'}\,dt<\infty,$$

for some $p'>1$. The equivalence may be proved by changing the integration variable in (7.5) to $x=\cos t$, and then using a result proved in the appendix of Ref. 10.

[1]R.E. Beard, J. Inst. Actu. **73**, 356 (1947).
[2]A. Young, Proc. R. Soc. Lond. **A224**, 552 (1954).
[3]Y.E. Kim, J. Math. Phys. **10**, 1491 (1969).
[4]Y.E. Kim, Phys. Lett. **29B**, 411 (1969); Y.E. Kim and A. Tubis, Phys. Rev. C **1**, 1627 (1970).
[5]F. Sohre and H. Ziegelmann, Phys. Lett. **34B**, 579 (1971).
[6]P. Doleschall, Nucl. Phys. **A201**, 264 (1973); **A220**, 491 (1974).
[7]N.M. Larson and J.H. Hetherington, Phys. Rev. C **9**, 699 (1974).
[8]A review of three-nucleon applications has been given by Y.E. Kim and A. Tubis, Ann. Rev. Nucl. Sci. **24**, 69 (1974).
[9]R. Kussmaul, Computing **9**, 159 (1972); R. Piessens and M. Branders BIT **13**, 443 (1973); N.S. Bakhvalov and L.G. Vasil'eva, U.S.S.R. Computational Math. and Math. Phys. **8**, No.1, 241 (1968); D. Elliott and D.F. Paget, BIT **16**, 32 (1976).
[10]I.H. Sloan, BIT **18**, 91 (1978).
[11]I.H. Sloan and W.E. Smith, Numerische Math. **30**, 415 (1978); *ibid.* (to be published).
[12]W.E. Smith and I.H. Sloan, SIAM J. Numer. Anal. (to be published).
[13]D.F. Paget and D. Elliott, Numerische Math. **19**, 373 (1972); D.B. Hunter, Numerische Math. **19**, 419 (1972).
[14]D. Elliott and D.F. Paget, Numerische Math. **23**, 311 (1975); **25**, 287 (1976).
[15]D. Elliott and D.F. Paget, Math. Comp. **33**, 301 (1979).
[16]A. Young, Proc. R. Soc. London A **224**, 561 (1954).
[17]K.E. Atkinson, SIAM J. Numer. Anal. **4**, 337 (1967).
[18]I.H. Sloan, J. Integral Equations (to be published).
[19]D. Elliott and D.F. Paget, BIT **18**, 137 (1978).
[20]L.D. Faddeev, *Mathematical Aspects of the Three-Body Problem in the Quantum Scattering Theory* (Daniel Davey, Hartford, CT., 1965).
[21]P.J. Davis, *Interpolation and Approximation* (Blaisdell, Waltham, MA, 1963), pp. 78–79.
[22]R. Piessens and M. Branders, J. Computational Phys. **21**, 178 (1976).
[23]P.J. Davis and P. Rabinowitz, *Methods of Numerical Integration* (Academic, New York, 1975).
[24]E.W. Cheney, *Introduction to Approximation Theory* (McGraw-Hill, New York, 1966).
[25]The results of Ref. 14 were established for a slightly different $k(t)$, namely (4.11) multiplied by the Jacobi weight function $(1-t)^\alpha(1+t)^\beta$. However, the corresponding result for (4.11) itself can be deduced by writing $(t-\lambda)^{-1}=(1-t)^\alpha(1+t)^\beta(1-\lambda)^{-\alpha}(1+\lambda)^{-\beta}(t-\lambda)^{-1}+$ remainder. The first term can then be handled by the result in Ref. 14, whereas the remainder, being no longer singular at $t=\lambda$, satisfies the condition (4.7).
[26]G. Szegö, *Orthogonal Polynomials*, 3rd ed. (American Mathematical Society Colloquium Publications, Vol. 23, Providence, Rhode Island, 1967).

Decaying states in the rigged Hilbert space formulation of quantum mechanics

A. Bohm
Center for Particle Theory, Department of Physics, University of Texas, Austin, Texas 78712

(Received 6 February 1979; accepted for publication 6 April 1979)

Within the rigged Hilbert space formulation of quantum mechanics idealized resonances (without background) are described by generalized eigenvectors of an essentially self-adjoint Hamiltonian with complex eigenvalue and a Breit–Wigner energy distribution. This establishes the link between the S matrix description of resonances by a pole and the usual description of states by vectors, overcomes theoretical problems connected with the deviation from exponential law and simplifies the calculation of the decay rate formula.

I. INTRODUCTION

The rigged Hilbert space (RHS) $\Phi \subset \mathscr{H} \subset \Phi^{\times}$, introduced into physics around 1965,[1] has already displayed so many features which make it ideally suited for the description of quantum mechanics[2] that one wonders why it has not yet been more generally accepted by physicists. More recently, around 1975, a new attribute of the RHS was uncovered[3]: It occurred that decay phenomena are most naturally described[3] using generalized eigenvectors of the self-adjoint energy operator with complex eigenvalue.[4] Such a description would establish the link between the S-matrix description of a resonance state as a pole and the usual description of states as vectors in a linear space; this is probably the reason why the complex energy eigenvectors have immediately caught the fancy of physicists working on decaying systems.[5,6] However, there are two points which were the actual motivation for the introduction of these generalized eigenvectors[7] and which have not been adequately mentioned in Ref. 5: (1) Their justification from the physical production process of decaying states, and (2) their application in the derivation of the decay rate formula.

In distinction to von Neumann's formulation of quantum mechanics, which is based on the postulated one-to-one correspondence between (pure) physical states and rays of the Hilbert space, we will use the RHS formulation of quantum mechanics,[2,7] according to which the physical states are elements of the dense nuclear subspace $\Phi \subset \mathscr{H}$. Though this change makes mathematically an enormous difference, physically one can not really discriminate between these two formulations. Both are idealizations, though the RHS formulation gives a description of physical states which is closer to the experimental situation. The use of the RHS in von Neumann's formulation may lead to a deeper insight, but the full advantage of the RHS can only be realized in the RHS formulation which eliminates many mathematical complications of von Neumann's formulation, as it uses only RHS-continuous operators for the observables. The RHS has in $\Phi^{\times}$ (space of continuous *antilinear* functionals) elements that describe idealized scattering states and idealized decaying states. These are generalized eigenvectors[2] $|\omega\rangle$ of the essentially self-adjoint (e.s.) Hamiltonian H:

$$\langle H\phi\,|\omega\rangle = \langle \phi\,|H^{\times}|\omega\rangle = \omega\langle \phi\,|\omega\rangle. \tag{1}$$

They occur in the nuclear spectral theorem[5] for the operator H:

$$(\psi,\phi) = \int_{\Lambda} \mu\, dE\, \langle \psi|E\rangle\, \langle E\,|\tilde{\phi}\,\rangle, \tag{2}$$

where Λ is the spectrum of H (which for simplicity is assumed to be absolutely continuous and μ is taken to be one).

If Φ is such that $\langle \psi|E\rangle$ and $\langle E\,|\tilde{\phi}\,\rangle$ are restrictions of analytic functions $\langle \psi|\omega\rangle$, $\langle \omega^*|\tilde{\phi}\,\rangle$ in a domain that includes Λ one can deform the path of integration from the real axis Λ to a curve $\mathscr{C}$ in the complex energy plane

$$(\psi,\phi) = \int_{\mathscr{C}} d\omega\, \langle \psi|\omega\rangle\, \langle \omega^*|\tilde{\phi}\,\rangle.$$

If the function $\langle \omega^*|\tilde{\psi}\rangle = \overline{\langle \psi|\omega\rangle}$ has a pole at (* and $\overline{}$ denote complex conjugation) $z_R = E_R - \frac{i}{2}\Gamma$, i.e., the function $\langle \psi|\omega\rangle$ has a pole at z_R^* between the old path Λ and the new path $\mathscr{C}$ one has to add to the above integral along the path $\mathscr{C}$ the term

$$\int_{\mathfrak{S}} dz \frac{\psi_{-1}}{z - z_R^*}\, \langle z^*|\tilde{\phi}\,\rangle,$$

where $\mathfrak{S}$ is the circle around z_R^* and ψ_{-1} is the residuum of

$$\langle \psi|z\rangle = \frac{\psi_{-1}}{z - z_R^*} + \psi_0 + \psi_1 (z - z_R^*) + \cdots.$$

With suitably assumed properties of the function $\langle z^*|\tilde{\phi}\,\rangle$ at the infinite semi-circle the path for this integral can be taken along the real line from $-\infty$ to $+\infty$. So one can write for the physical state vector $\psi \in \Phi$ (considered as a functional on $\Phi \cap \mathscr{H}_+$, where $\mathscr{H}_+$ is the space of Hardy class functions with respect to the upper half-plane)

$$\langle \psi| = \int_{\mathscr{C}} d\omega\, \langle \psi|\omega\rangle \langle \omega^*| + \int_{-\infty}^{+\infty} dE \frac{\psi_{-1}}{E - z_R^*} \langle E\,|. \tag{3}$$

It has been proposed[3,7] to use the second term in (3) for the description of decaying states.

Before we define these idealized decaying state vectors from the physical motivation and show their usefulness for the calculation of decay rates we want to make some remarks regarding our notation. Physical observables and transformations are represented by bounded and continuous operators, A, in ϕ. Let $A^{\times}$ denote its conjugate operator $A^{\times}$ [defined as in (1)] in $\Phi^{\times}$ (which is also continuous), $\bar{A}$ its closure in $\mathscr{H}$ (which is in general not bounded), and $A^{\dagger}$ its adjoint (restriction of the Hilbert space adjoint to Φ), then to each physical quantity corresponds the triplet of operators $A \subset \bar{A}$

 0022-2488/80/051040-04$01.00

$\subset A^{+\times}$ in the Gelfand triplet of space $\Phi \subset \mathcal{H} \subset \Phi^\times$. To simplify the notation and to exhibit the correlate to formal scattering theory we will denote this triplet just by A, the precise meaning follows from the space in which it acts.

II. PHYSICAL MOTIVATION FOR THE MATHEMATICAL FORM OF THE DECAYING STATE VECTOR

Unstable systems are prepared by scattering experiments in which the delay time is large. The intermediate quasistationary system, when it is considered as an isolated decaying system ignoring the mode of formation, must therefore have the properties of a resonance which means that the state vector ϕ^R, which is to represent it, must have a Breit–Wigner energy distribution, i.e, if we write

$$\langle \phi^R | = \int dE \langle E | f^*(E - E_R), \tag{4}$$

where $|E\rangle$ are the generalized eigenvectors of the essentially self-adjoint H with E belonging to the spectrum of $\bar{H}$, then

$$|f(E - E_R)|^2 = \frac{1}{\pi} \frac{\Gamma/2}{(E_R - E)^2 + (\Gamma/2)^2}, \tag{5}$$

with (E_R, Γ) the resonance parameters. Consequently

$$f(E - E_R) = \left(\frac{\Gamma}{2\pi}\right)^{1/2} \frac{1}{z_R - E} \quad z_R = E_R - i\tfrac{\Gamma}{2}. \tag{6}$$

We take the "scalar product" of (4) with a $\varphi \in \Phi \subset \mathcal{H} \subset \Phi^\times$, which has the property that $\langle E|\varphi\rangle$ is the limit of a function $\langle z|\varphi\rangle$ analytic in the upper half-plane (precisely, the value of the functional $\langle \phi^R |$ at $\varphi \in \Phi \cap \mathcal{H}_+$). Then

$$\begin{aligned}\langle \phi^R | \varphi\rangle &= \left(\frac{\Gamma}{2\pi}\right)^{1/2} \int dE \langle E|\varphi\rangle \frac{1}{z_R^* - E} \\ &= \left(\frac{\Gamma}{2\pi}\right)^{1/2} (-2\pi i) \langle z_R^* | \varphi\rangle.\end{aligned} \tag{7}$$

In the last equality we have used Titchmarsh theorem[8] after extending the integration over $-\infty < E < +\infty$.

In the same way one calculates

$$\langle \phi^R | H | \varphi\rangle = z_R^* \langle z_R^* | \varphi\rangle (-i\sqrt{2\pi\Gamma}), \tag{8}$$

and

$$\langle \phi^R | e^{iHt} | \varphi\rangle = e^{iE_Rt} e^{-(\Gamma/2)t} \langle \phi^R | \varphi\rangle. \tag{9}$$

Omitting again the arbitrary $\varphi \in \Phi \cap \mathcal{H}_+$, we have in

$$\begin{aligned}\langle \phi^R | &= -i\sqrt{2\pi\Gamma}\, \langle z_R^* | \\ &= -i\sqrt{2\pi i}\, \frac{1}{2\pi i} \int dE \langle E | \frac{1}{E - z_R^*},\end{aligned} \tag{10}$$

a vector which is (a) normalized (element of $\mathcal{H}_+$), (b) not in the domain of $\bar{H}$, (c) a generalized eigenvector of H with eigenvalues $z_R = E_R - i(\Gamma/2)$, and (d) a decaying state vector.[9] Properties (c) and (d) are given by (8) and (9), respectively. Properties (a) and (b) can be seen as follows:

$$\begin{aligned}\langle \phi^R | \phi^R\rangle &= \frac{\Gamma}{2\pi} \int dE dE' \langle E | E'\rangle \frac{1}{(E - E_R)^2 + (\Gamma/2)^2}, \\ &= \frac{1}{\pi} \int dx \frac{1}{x^2 + 1} = 1,\end{aligned}$$

$$\langle \phi^R | H^2 | \phi^R\rangle = \frac{\Gamma}{2\pi} \int dE \frac{E^2}{(E - E_R)^2 + (\Gamma/2)^2} \to \infty.$$

ϕ^R is certainly not in the space of physical states Φ. This we should have expected, as in all experimental situations a resonance is accompanied by a background and the Breit–Wigner amplitude without background is just an idealization. Thus, even for the simplest possiblity an element of Φ will always contain a background term represented by the first integral in (3). Nevertheless, it is often useful and sufficiently accurate to employ a description in which one isolates the intermediate quasistationary system, ignores its mode of formation, describes it by a Breit–Wigner amplitude, and represents it by an idealized decaying state vector ϕ^R.

III. APPLICATION OF THE COMPLEX ENERGY EIGENVECTORS

To demonstrate the usefulness of describing decaying states by (10) we will calculate the decay rate. To do this we will have to make use of some results of scattering theory. Digressing from the subject of this paper we briefly comment on the possibilities of their justification and introduce the notation. The subject of this paper will then continue with the application of the decaying states in the general formula (11) below.

It is usually assumed that $\bar{H} = \bar{H}_0 + \bar{V}$ where $\bar{H}_0$ and $\bar{H}$ are the self-adjoint free and interaction Hamiltonians, respectively. For simplicity we assume that the spectra of H_0 and H are absolutely continuous and identical (no bound states) and ignore degeneracy. As it is not known how to construct the topology in Φ we will have to assume that also $H = H_0 + V$ with H, H_0, V e.s. continuous operators in Φ. We will further assume that V is such that the wave operators $\Omega^{\pm}$ exist and that the T matrix $\langle \lambda | V | E\rangle$ exists, where $|\lambda\rangle$ are the generalized eigenvectors of the spectrum of H_0. Then the Lippmann–Schwinger equation is given in the form

$$\Omega_+^{+\times} = \int_{\Lambda_0} d\lambda \left(|\lambda\rangle \langle \lambda | - R_{\lambda - i\epsilon}^{0+\times} V^\times \Omega_+^{+\times} |\lambda\rangle \langle \lambda | \right),$$

where $R_z^0 = (\bar{H}_0 - z)^{-1}$ and Λ_0 is the spectrum of H_0. $\Omega_+^{+\times}$ relates the generalized eigenvector $|\lambda\rangle$ of H_0 to the generalized eigenvector $|E\rangle$ of H: $|E\rangle = \Omega_+^{+\times} |\lambda\rangle$. The transition rate from a state $\phi(t) = e^{-iHt}\phi$ into a subspace $\Pi\Phi$ is given by (omitting all other quantum numbers but the energy)

$$\begin{aligned}\dot{\mathcal{P}}(t) &= \frac{d}{dt} (\mathrm{Tr}\Pi |\phi(t)\rangle \langle \phi(t)|) \\ &= -i \int dE_b \int dE dE' e^{-i(E - E')t} \\ &\quad \times \langle b | V | E\rangle \langle E' | V | b\rangle \langle E | \phi\rangle \langle \phi | E'\rangle \\ &\quad \times \left(\frac{1}{E' - E_b - i\epsilon} - \frac{1}{E - E_b + i\epsilon} \right),\end{aligned} \tag{11}$$

where $\langle b | V | E\rangle$ is the T matrix, $|b\rangle$ is a basis of eigenvectors of H_0 for the subspace $\Pi\Phi$, and the integration $\int dE_b$ runs over this subspace.

To specify this general expression for the transition rate to the decay rate we choose for the state ϕ the decaying state ϕ^R. ϕ^R is, of course, not exactly a physically preparable state

ψ, but ϕ^R is the resonance ingredient with the exponential time development of such a physically preparable state. Of the generalized eigenvector expansion (3) of ψ only the ϕ^R component is significant in (8). Then, inserting (10) into (11) one obtains for the decay rate

$$\dot{\mathscr{P}}(t) = -i\int dE_b \int dE dE' \langle b|V|E\rangle \langle E'|V|b\rangle \times \frac{\Gamma}{2\pi}\frac{e^{-iEt}}{E-[E_R+i(\Gamma/2)]}\frac{e^{iE't}}{E'-[E_R-i(\Gamma/2)]} \times\left(\frac{1}{E'-(E_b+i\epsilon)} - \frac{1}{E-(E_b-i\epsilon)}\right) = \dot{\mathscr{P}}_1 + \dot{\mathscr{P}}_2. \tag{12}$$

The integration can now be carried out for $\epsilon > 0$ using again the Titchmarsh theorem[8] and assuming that $\langle E|V|b\rangle$ is such that its conditions are fulfilled. For the first term $\dot{\mathscr{P}}_1$ of (12) one integrates first over E' then takes the complex conjugate, integrates over E and takes again the complex conjugate. The result is

$$\dot{\mathscr{P}}_1(t) = 2\pi\Gamma i\int dE_b \langle b|V|E_R - i\Gamma/2\rangle \langle E_b + i\epsilon|V|b\rangle \times\frac{e^{-iE_Rt}e^{-\Gamma t/2}e^{iE_bt}e^{-\epsilon t}}{E_b - E_R + i(\epsilon+\Gamma/2)}, \tag{13}$$

where we have used $\langle b|V|E+i\Gamma/2\rangle = -\langle b|V|E - i\Gamma/2\rangle$ for the T matrix, which is an immediate consequence of the well-known symmetry relation for the S matrix.[8] For the initial decay rate $\dot{\mathscr{P}}(0)$ one obtains from this for $\epsilon\to\Gamma/2\to 0$

$$\dot{\mathscr{P}}_1(0) = 2\pi\Gamma i\int dE_b \langle b|V|E_R - i\Gamma/2\rangle \times\langle E_b + i\Gamma/2|V|b\rangle\frac{1}{E_b - E_R + i0}. \tag{14}$$

A similar expression is obtained for $\dot{\mathscr{P}}_2(0)$ with $+i0$ replaced by $-i0$. If one then uses the well-known relation between distributions

$$\frac{1}{E_b - E_R + i0} - \frac{1}{E_b - E_R - i0} = -2\pi i\delta(E_b - E_R), \tag{15}$$

and re-inserts (10) one obtains the well-known expression for the initial decay rate

$$\dot{\mathscr{P}}(0) = 2\pi\int dE_b \langle b|V|\phi^R\rangle\langle\phi^R|V|b\rangle\delta(E_b - E_R). \tag{16}$$

If the T matrix is a slowly varying function of the complex energy z, $\langle E_b + i\epsilon|V|b\rangle \approx \langle E_R + i\frac{\Gamma}{2}|V|b\rangle$ for $E_b \approx E_R$, then one can use (13) also for Γ which are larger than the energy resolution of the detector, $\Delta E_b \ll \Gamma$, and obtains ($\epsilon\to +0$):

$$\dot{\mathscr{P}}(0) = 2\pi\Gamma i\int dE_b\left(\frac{\langle b|V|E_R - i\Gamma/2\rangle\langle E_b + i\epsilon|V|b\rangle}{E_b - E_R + i(\epsilon+\Gamma/2)} - \frac{\langle b|V|E_b - i\epsilon\rangle\langle E_R + i\Gamma/2|V|b\rangle}{E_b - E_R - i(\epsilon+\Gamma/2)}\right)$$
$$= 2\pi\Gamma i\int dE_b \langle b|V|E_R - i\Gamma/2\rangle \times\left\langle E_R + i\frac{\Gamma}{2}\middle|V\middle|b\right\rangle\left(\frac{-i\Gamma}{(E_b - E_R)^2 + (\Gamma/2)^2}\right)$$
$$= 2\pi\int dE_b \langle b|V|\phi^R\rangle\langle\phi^R|V|b\rangle \times\left(\frac{\Gamma/2}{(E_b - E_R)^2 + (\Gamma/2)^2}\right), \tag{17}$$

where the last factor is the well-known natural line width. If one compares these simple straightforward calculations with the conventional procedure[10] one will appreciate the usefulness of our new vectors (10).

IV. CONCLUDING REMARKS

The questions we have addressed ourselves to is the description of a decaying state with the resonance parameter (E_R, Γ); our suggestion is (10). We have not discussed the question how one calculates the position of the resonance. That is a completely different problem whose answer depends upon the particular property of the energy operator H, and may be connected with the choice of the topology in Φ, a problem which is so far completely unresolved. For discussions in mathematics connected with the spectral analysis of self-adjoint operators $\bar{H}$ one likes to choose the topology of Φ such that $\Phi^\times$ contains only those generalized eigenvectors whose eigenvalues correspond to the spectrum of $\bar{H}$.[11] For physical systems whose algebra of observables is given by the enveloping algebra of a semisimple group, it was suggested[12] to define the topology (which is then nuclear[12]) by the countable number of scalar products $(\phi,\psi)_n = (\phi,(\Delta+1)^n\psi)$. Here Δ is the Nelson operator of the group, which is often related to the Hamiltonian of that system. These two remarks suggest to construct the topology in Φ for a physical system with energy operator H such that $\Phi^\times$ contains only those generalized eigenvectors with complex eigenvalues $E_R^{(n)} - i[\Gamma^{(n)}/2]$ that correspond to the parameters $(E_R^{(n)}, \Gamma^{(n)})$ for resonances of that particular physical system and does not contain $\langle\phi^R|$'s given by (10) for any other value of z_R. How this can be done and how this is connected with the construction using the Nelson operator is a problem that has not yet been treated.

ACKNOWLEDGMENTS

I am particularly grateful to B. Nagel, who taught me the properties of the generalized eigenvectors with complex eigenvalue many years ago and with whom I discussed the vectors defined by (10); to K. Napiorkowski with whom I discussed (10) and who referred me to the work of Baumgärtel; and to H. Baumgärtel for his extensive correspondence on the subject of this paper over the past two years. I also gratefully acknowledge discussions of this subject with L. Horwitz, E.C.G. Sudarshan, I. Prigogine, and V. Gorini in Fall 1976, and with W.C. Schieve, A. Grecos, V. Gorini, and G. Parravicini more recently.

[1] J.E. Roberts, J. Math. Phys. 7, 1097 (1966); A. Böhm, Boulder Lectures in Theoretical Physics 9A, 255 (1966).
[2] A. Böhm, The Rigged Hilbert Space and Quantum Mechanics, Springer Lecture Note 78 (1978); J.E. Roberts, Commun. Math. Phys. 3, 98 (1966);

J.P. Antoine, J. Math. Phys. **10**, 53 (1969); O. Melsheimer, J. Math. Phys. **15**, 902 (1974); G. Lassner, Wiss. Z. Karl Marx Univ. Leipzig Math.-Naturwiss. R. **22** (1973) H.Z.; A. Böhm, *Lectures at the Istanbul Summer Institute* 1970, edited by A.O. Barut (Reidel, Dordrecht, 1973), p. 197.

[3]Two distinct suggestions have been made, the first (which constructs a RHS that contains exactly those complex energy eigenvectors which are related to the discrete eigenvalues of $H_0 = H - V$) by H. Baumgärtel in Math. Nachr. **75**, 133 (1976) and the second (which defines complex energy eigenvectors that have a Breit–Wigner energy distribution) by A. Böhm in a manuscript distributed to colleagues in 1976 which has appeared as Chapter XXI of Ref. 7.

[4]That such eigenvectors exist was known to physicists for a long time from their careless manipulations with generators of noncompact subgroups in representations of Lie algebras. The first detailed rigorous discussion of such vectors of SU(1,1) was given by G. Lindblad and B. Nagel, Ann. Inst. Henri Poincaré XIII **27** (1970). See also G.J. Iverson, Phys. Lett. **26B**, 229 (1968). Usually one tries to exclude from $\Phi^\times$ all generalized eigenvectors with eigenvalues that do not belong to the spectrum by an appropriate choice of the nuclear topology of ϕ; see Ref. 11.

[5]G. Parravicini, V. Gorini, and E.C.G. Sudarshan, Resonances, "Scattering Theory and Rigged Hilbert Space," J. Math. Phys. (unpublished); E.C.G. Sudarshan, C.B. Chiu, and V. Gorini, Phys. Rev. D **18**, 2914 (1978).

[6]The association of complex energies with resonances is of course almost as old as quantum mechanics, and other recent suggestions include the use of nonself-adjoint Hamiltonian operators. See, e.g., J.S. Howland, J. Math. Anal. Appl. **50**, 415 (1975); B. Simon, Ann. Math. **97**, 247 (1973); J. Aguilar and J.M. Combes, Comm. Math. Phys. **22**, 280 (1971); L.P. Horwitz and I. Sigal, Helv. Phys. Acta (to be published). T.K. Bailey and W.C. Schieve, Nuovo Cimento, Dec. 1978; C. George. F. Henin, F. Mayne, and I. Prigogine, Hadron Journal. **1** (1978); A.P. Grecos and I. Prigogine, Proc. Nat. Acad. Sci., USA **69**, 1629 (1972); A. Grossmann, J. Math. Phys. **5**, 1025 (1964).

[7]A. Böhm, *Quantum Mechanics* (Springer, New York, 1979).

[8]See, e.g., H.M. Nussenzveig, *Causality and Dispersion Relations* (Academic, New York, 1972), Eqs. (1.6.6) and (2.8.16).

[9]F. Fonda, G.C. Ghirardi, and A. Rimini, Rep. Prog. Phys. **41**, 587 (1978), describe the well-known deviations from the exponential decay law which however do not apply for ϕ^R because ϕ^R is not in the domain of the Hilbert space operator $\bar{H}$ (small times) and ϕ^R was defined in the approximation $E_R/\Gamma \to \infty$ with the lower bound of the spectrum going to $-\infty$ (large times).

[10]M.L. Goldberger and K.M. Watson, *Collision Theory* (Wiley, New York, 1964).

[11]K. Napiorkowski, Bull. Polish Acad. Sci. **22**, 1215 (1974); **23**, 251 (1975). See also W. Faris, Helv. Phys. Acta **44**, 930 (1971) and D. Babbitt, Rep. Math. Phys. **3**, 37 (1972), where this problem has been discussed not for the RHS but for a "scale of Hilbert spaces."

[12]A. Böhm, J. Math. Phys. **8**, 1557 (1967), Appendix B.

Resonances in the Klein–Gordon theory of the relativistic Stark effect

S. Graffi,[a)] V. Grecchi[a),b)]
Centre de Physique Théorique II, CNRS Marseille, France
M. Maioli
Istituto Matematico, Università di Modena, 41100 Modena, Italy

(Received 18 October 1979; accepted for publication 30 November 1979)

Existence of resonances is proved for the time-independent Klein–Gordon equation describing the interaction of a charged particle with an external uniform field of small strength F in addition to the Coulomb attraction. It is further shown that the resonances reduce to the exactly known bound states of the problem as $F \to 0$, and to the resonances of the nonrelativistic Stark effect as $c \to \infty$.

I. INTRODUCTION

This paper represents a further contribution to the current rapidly developing rigorous theory of resonances in Schrödinger operators under the action of a constant electric field.[1-11]

As a first step towards the understanding of the resonance phenomena in the relativistic Stark effect, at least for two-body systems, and of the corresponding nonrelativistic limit, we consider here a spinless case, i.e., the Klein–Gordon theory for a particle in an electrostatic potential due to the Coulomb attraction by a fixed charge Z and to an external uniform electric field of strength F. Furthermore, this model is also of some direct physical interest, since it is believed to describe the scalar mesic atoms (see, for example, Ref. 12).

The time-independent Klein–Gordon equation describing the interaction of a relativistic particle of rest mass m and charge e with an electromagnetic field of potential $(\mathbf{A},\Phi)$ is given by (see, for example, Ref. 22).

$$(E - e\Phi)^2\psi = (-ic\nabla - e\mathbf{A})^2\psi + m^2c^4\psi, \tag{1.1}$$

where E is the energy of the particle, and $\hbar = 1$. In the present case $\mathbf{A} = 0$, $\Phi = -e(Z/r) + eFx$; furthermore, defining as usual $W = E - mc^2$, Eq. (1.1) can be rewritten as

$$H\psi = W\psi, \tag{1.2}$$

where

$$H = -\frac{1}{2m}\Delta - \frac{e^2}{2mc^2}\Phi^2 + e\Phi + \frac{eW}{mc^2}\Phi - \frac{W^2}{2mc^2}. \tag{1.3}$$

These formulas clearly exhibit a major difficulty of the problem when written in this form, i.e., the constrained nature of its spectral problem, namely, one has to find those values of W for which the point W itself belongs to the spectrum of H. In other words, Eqs. (1.2) and (1.3) give rise to an implicit spectral problem. This difficulty does not occur if one considers the Hamiltonian theory of the problem, instead of the Klein–Gordon one, as done by Weder[14] and Herbst[9] for the pure Coulomb case ($F = 0$). However, the pure Coulomb case is the only one in which, as is well known, the implicit spectral problem can be exactly solved, and in the nonrelativistic case a similar difficulty, arising when the appropriate Schrödinger operator is realized in squared parabolic coordinates, has been overcome to yield a first proof of the existence of the resonances.[4] It is therefore conceivable that, at least for small values of the field strength F, the arguments of Ref. 4 can be generalized to provide, through the dilation analyticity techniques, an existence proof of resonances also in this context.

The purpose of this paper is to present such a generalization, to be given in the following way: In Sec. II we discuss the realization of the partial differential expression H, for real values of W, as an essentially self-adjoint operator in $L^2(\mathbb{R}^3)$; in addition, the nonself-adjoint operator $H(\theta)$ associated with H through the dilation analyticity technique is realized, and a generalized strong convergence as $\mathrm{Im}\theta \to 0$ is proved. In Sec. III, the spectral properties of H and $H(\theta)$ are examined, both in the standard sense as well as in the implicit one. In particular, the essential spectrum of $H(\theta)$ is determined. In Sec. IV the realization of $H(\theta)$ in squared parabolic coordinates is introduced, which is such that $H(\theta)$ is represented, in any invariant subspace of magnetic quantum number $m = 0,1,\ldots$, by a slightly generalized two-dimensional anharmonic oscillator to which the Simon key results[15] are easy to extend. In this way, we show for F small, through an implicit function argument, the existence of complex implicit eigenvalues, independent of θ, which reduce to the pure Coulomb bound states as $F \to 0$ and to the nonrelativistic Stark effect resonances as $c \to \infty$. In Sec. V these eigenvalues are shown to be second sheet poles of the scalar products of the appropriate resolvent operator on some suitable dense set of states, so that they represent resonances according to the standard notion of this concept.[16]

Finally, the adaptment of Simon's results needed in Sec. IV and V is presented in the Appendix.

II. ESSENTIAL SELF-ADJOINTNESS AND COMPLEX SCALING

Our first aim in this section is to realize, when $W \in \mathbb{R}$ and $\Phi = e[-(Z/r) + Fx]$, the partial differential expression (1)–(3) as an essentially self-adjoint operator in $L^2(\mathbb{R}^3)$. Assuming from now on $m = \frac{1}{2}$, $e^2 = 1$, the explicit form of H is

[a)]Permanent address: Istituto Matematico, Università di Modena, 41100 Modena, Italy.
[b)]Partially suported by I.N.F.N., Sezione di Bologna.

$$H = -\Delta - Z\left(1+\frac{2W}{c^2}\right)\frac{1}{r} - \frac{Z^2}{c^2}\frac{1}{r^2} + F\left(1+\frac{2W}{c^2}\right)x - \frac{F^2}{c^2}x^2 + \frac{2ZF}{c^2}\frac{x}{r} - \frac{W^2}{c^2}. \quad (2.1)$$

Since we intend to adapt to the present situation the Faris–Lavine[17] version of the Nelson[18] commutator theorem, we begin by introducing an auxiliary operator, related to the self-adjoint realization of $H + (2F^2/c^2)x^2$.

More precisely (see Ref. 19, Sec. VI.4.3) let N be the semibounded self-adjoint operator in $L^2(\mathbb{R}^3)$ defined by the partial differential expression

$$N(\alpha,\beta,\gamma) = -\Delta - \frac{\alpha}{r^2} - \frac{\beta}{r} + \gamma x^2, \quad \alpha<\tfrac{1}{4},\ \beta\in\mathbb{R},\ \gamma>0, \quad (2.2)$$

on the domain [from now on, we shall take $L^2 = L^2(\mathbb{R}^3)$ unless otherwise stated]

$$D(N) = \{u\in L^2 | u' \equiv \operatorname{grad} u \in (L^2)^3 | \|xu\| < \infty \ | \Delta u \text{ exists} | N(\cdot)u\in L^2\}, \quad (2.3)$$

where all differentiations are taken in the generalized sense.

Let in addition H_0 be the semibounded self-adjoint operator in L^2 defined by $N(\alpha,\beta,0)$ on the domain

$$D(H_0) = \{u\in L^2 | u' = \operatorname{grad} u \in (L^2)^3 | \Delta u \text{ exists} | N(\cdot,0)u\in L^2\}. \quad (2.4)$$

We recall that the implicit spectral problem for H_0, i.e., Eq. (1.2) for $F=0$, can be exactly solved (See Ref. 22, Sec. 42), the implicit energy eigenvalues being given by

$$W^0_{n,l} = \tfrac{1}{2}c^2\left[\left(1+\frac{Z^2}{c^2}\lambda^{-2}\right)^{-1/2} - 1\right],$$

$$\lambda = n + \tfrac{1}{2} + \left[(l+\tfrac{1}{2})^2 - \frac{Z^2}{c^2}\right]^{1/2}, \quad n,l = 0,1,\cdots. \quad (2.5)$$

Our purpose is to prove that H as given by Eq. (2.1) generates an essentially self-adjoint operator in L^2 when defined on $D(H_0)\cap D(x^2)$, $D(x^2)$ being the domain of the maximal multiplication operator by x^2 in L^2.

Let us begin by the following quadratic estimate:

Lemma 2.1: Let $u\in D(H_0)\cap D(x^2)$. Then there are positive constants a and b such that

$$\|H_0u\|^2 + \|\gamma x^2 u\|^2 \leqslant a\|(H_0+\gamma x^2)u\|^2 + b\,\|u\|^2. \quad (2.6)$$

Proof: Let $u\in D(H_0)\cap D(x^2)$. Then, following the argument of Ref. 15, Lemma II.1.1, as quadratic forms on $D(H_0)\cap D(x^2)\otimes D(H_0)\cap D(x^2)$ we can write

$$\begin{aligned}(H_0+\gamma x^2)^2 &= H_0^2 + \gamma^2x^4 + \gamma H_0x^2 + \gamma x^2H_0\\ &= H_0^2 + \gamma^2x^4 + \gamma p^2x^2 + \gamma x^2p^2 - 2\gamma x^2 \times\left(\frac{\alpha}{r^2}+\frac{\beta}{r}\right)\\ &= H_0^2 + \gamma^2x^4 + \gamma[p_x,[p_x,x^2]] + 2\gamma\,\mathbf{p}\cdot x^2\mathbf{p} - 2\gamma x^2\left(\frac{\alpha}{r^2}+\frac{\beta}{r}\right)\\ &\geqslant H_0^2 + \gamma^2x^4 - 2\gamma - 2\gamma x^2\left(\frac{\alpha}{r^2}+\frac{\beta}{r}\right).\end{aligned}$$

[It is easy to check that all commutations appearing above make sense when considered as quadratic forms on $D(H_0)\cap D(x^2)\otimes D(H_0)\cap D(x^2)$.] Now choose two positive constants $k<1$ and b such that $k\gamma^2x^4 - 2\gamma x^2[(\alpha/r^2)+(\beta/r)] - 2\gamma + b\geqslant 0$ for all x. Then we have $(H_0+\gamma x^2)^2 + b \geqslant (1-k)\times(H_0^2+\gamma^2x^4)$, whence the result with $a=(1-k)^{-1}$.

The above estimate implies $u\notin D(N)$ if $u\notin D(H_0)\cap D(x^2)$. Since $D(N)\supset D(H_0)\cap D(x^2)$, and N is self-adjoint, we have the following:

Corollary 2.1: Let N be defined by Eqs. (2.2) and (2.3). Then $D(N) = D(H_0)\cap D(x^2)$.

Lemma 2.2: Let H_1 be defined as an operator in L^2 by

$$D(H_1) = D(H_0)\cap D(x^2), \quad H_1 = N - 2\gamma x^2 + \delta x, \quad \delta\in\mathbb{R}. \quad (2.7)$$

Then H_1 is essentially self-adjoint.

Proof: H_1 is trivially symmetric. Hence, by the Faris–Lavine variant[17] of the Nelson[18] commutator theorem, it is enough to prove on $D(N)\otimes D(N)$ the quadratic form estimate

$$\pm i[H_1,N] \leqslant kN \quad (2.8)$$

for some $k>0$.

A simple computation shows that, as quadratic forms on $D(N)\otimes D(N)$:

$$\pm i[H_1,N] = \pm 2i[p_x^2,\gamma x^2] \mp i[p_x^2,\delta_x]$$

so that, when $u\in D(N)$;

$$\begin{aligned}&\pm i\langle u,[H_1,N]u\rangle\\ &\quad= \pm 4\gamma[\langle p_xu,xu\rangle + \langle xu,p_xu\rangle] \pm 2\delta\langle p_xu,u\rangle\\ &\quad\leqslant 4\gamma\langle -\Delta u + x^2u,u\rangle + 2\delta\langle -\Delta u,u\rangle + k_1\langle u,u\rangle\end{aligned}$$

from some $k_1>0$. Now (see again Ref. 19, Sec. VI.4.3) the quadratic form domain $Q(N)$ of N is given by

$$Q(-\Delta)\cap Q(x^2) \equiv H^{2,1}(\mathbb{R}^3)\cap Q(x^2),$$

$H^{2,1}$ being the usual Sobolev space. Since $Q(N)\supset D(N)$, we have $u\in H^{2,1}(\mathbb{R}^3)$, so that, as is well known (see, for example, Ref. 20), there are $a<1$ and $b>0$ such that the quadratic form estimate $(\alpha/r^2) + (\beta/r) < a(-\Delta) + b$ holds. Hence, as quadratic forms on $D(N)\otimes D(N)$ $-\Delta \leqslant (1-a)^{-1}H_0 + b$, [and thus $4\gamma\langle -\Delta u + x^2u,u\rangle \leqslant 4\gamma(1-a)^{-1}\langle H_0u,u\rangle + 4\gamma\langle x^2u,u\rangle + 4\gamma b\langle u,u\rangle$, $2\delta\langle -\Delta u,u\rangle \leqslant 2\delta(1-a)^{-1}\times\langle H_0u,u\rangle + 2\delta b\langle u,u\rangle$], whence, adding a positive constant to N, we get the existence of $k>0$ such that

$$\pm i[H_1,N] \leqslant kN,$$

and this proves the lemma.

The realization of H as an essentially self-adjoint operator is now an immediate consequence of this lemma.

Theorem 2.1: Let $W\in\mathbb{R}$, and $(Z^2/c^2)<\frac{1}{4}$. Then the operator H defined by Eq. (2.1) on $D(N) = D(H_0)\cap D(x^2)$ is essentially self-adjoint in L^2.

Proof: It is enough to take $\alpha = Z^2/c^2$, $\beta = Z[1+(2W/c^2)]$, $\gamma = F^2/c^2$, $\delta = F[1+(2W/c^2)]$, and to remark that the maximal multiplication operator by x/r in L^2 is bounded. Hence, the result follows from Lemma 2.2 and the Rellich–Kato theorem.

Remark: Returning to the usual units, we see that H is essentially self-adjoint for all $Z<c/2 = 137/2$, in our units, exactly as in the $F=0$ case.

Let us now turn to the realization of the complex scaled operator associated with H by the dilation analyticity technique. Let $U(\theta)$, $\theta\in\mathbb{R}$, be the unitary dilation group in L^2: $(U(\theta)f)(x) = e^{(3/2)\theta}f(e^{\theta}x)$, $f\in L^2$, $x\in\mathbb{R}^3$.

Then it is easily seen that, on $U(\theta)D(N)$,

$$\begin{aligned}U(\theta)HU(\theta)^{-1} &= e^{-2\theta}\Big(-\Delta - \frac{\alpha}{r^2} - \frac{\beta}{r}e^{+\theta} - \gamma e^{4\theta}x^2 + \delta e^{3\theta}x \\ &\quad + \eta e^{2\theta}\frac{x}{r} - \frac{W^2}{c^2}e^{2\theta},\end{aligned} \tag{2.9}$$

$$\alpha = \frac{Z^2}{c^2},\quad \beta = Z\Big(1+\frac{2W}{c^2}\Big),\quad \gamma = \frac{F^2}{c^2},$$

$$\eta = \frac{2ZF}{c^2},\quad \delta = F\Big(1+\frac{2W}{c^2}\Big).$$

In particular, on $U(\theta)D(H_0)$,

$$U(\theta)H_0U(\theta)^{-1} = e^{-2\theta}\Big(-\Delta - \frac{\alpha}{r^2} - \beta\frac{e^{\theta}}{r} - \frac{W^2}{c^2}e^{2\theta}\Big). \tag{2.10}$$

We further define

$$H(\theta) = U(\theta)HU(\theta)^{-1},\quad H_0(\theta) = U(\theta)H_0(\theta)U(\theta)^{-1},$$

$$\tilde H(\theta) = e^{2\theta}H(\theta),\ \tilde H_0(\theta) = e^{2\theta}H_0(\theta). \tag{2.11}$$

To get well defined operators out of the above differential expressions also for $\mathrm{Im}\theta \neq 0$, let us first recall that (see, for example, Ref. 15) the operator T in L^2 defined by $D(T) = D(-\Delta)\cap D(x^2)$, $T = -\Delta - \gamma x^2$, $0 < \arg\gamma < 2\pi$, is strictly m sectorial, with quadratic form domain $Q(T) = Q(-\Delta)\cap Q(x^2)$. The uncertainty principle inequality $\alpha\langle u, r^{-2}u\rangle < \langle -\Delta u, u\rangle$, $\alpha < \frac{1}{4}$, $u\in Q(-\Delta)$, shows that $\langle Tu,u\rangle - \alpha\langle u, r^{-2}u\rangle$ is a closed strictly sectorial quadratic form if $u\in Q(-\Delta)$ (see again Ref. 19, Sec. 1-4-3). Let T_1 be the unique strictly m-sectorial operator associated with this quadratic form. To determine its domain, let us prove a further quadratic estimate:

Lemma 2.3: Let $u\in D(N)$, $\alpha < \frac{1}{8}$, $\mathrm{Im}\gamma\neq 0$. Then there are positive constants $a < 1 - |\mathrm{Re}\gamma/\gamma|$ and b such that

$$\Big\|\Big(-\Delta - \frac{\alpha}{r^2}\Big)u\Big\|^2 + |\gamma|^2\|x^2u\|^2 \leqslant a\Big\|\Big(-\Delta - \frac{\alpha}{r^2} - \gamma x^2\Big)u\Big\|^2 + b\|u\|^2. \tag{2.12}$$

Proof: Again, we follow an argument of Ref. 15 (Lemma II.9.1).

As quadratic forms on $D(N)\otimes D(N)$ we can write

$$\begin{aligned}&\Big(-\Delta - \frac{\alpha}{r^2} - \gamma x^2\Big)\Big(-\Delta - \frac{\alpha}{r^2} - \bar\gamma x^2\Big)\\
&= \Big(-\Delta - \frac{\alpha}{r^2}\Big)^2 + |\gamma|^2x^4 - \mathrm{Re}\gamma\Big[x^2\Big(-\Delta - \frac{\alpha}{r^2}\Big) + \Big(-\Delta - \frac{\alpha}{r^2}\Big)x^2\Big] \pm i|\mathrm{Im}\gamma|\Big[-\Delta - \frac{\alpha}{r^2}, x^2\Big]\\
&= \Big|\frac{\mathrm{Re}\gamma}{\gamma}\Big|\Big[\Big(-\Delta - \frac{\alpha}{r^2}\Big) \mp |\gamma|x^2\Big]^2 + \Big(1 - \Big|\frac{\mathrm{Re}\gamma}{\gamma}\Big|\Big)\Big[\Big(-\Delta - \frac{\alpha}{r^2}\Big)^2 + |\gamma|^2x^4\Big] \mp 2|\mathrm{Im}\gamma|(p_xx + xp_x)\\
&\geqslant (\text{for some } R>0)(a+R)\Big[\Big(-\Delta - \frac{\alpha}{r^2}\Big)^2 + |\gamma|^2x^4\Big] - 2|\mathrm{Im}\gamma|(-\Delta + x^2) + 2|\mathrm{Im}\gamma|(p_x \pm x)^2\\
&\geqslant (a+R)\Big[\Big(-\Delta - \frac{\alpha}{r^2}\Big)^2 + |\gamma|^2x^4\Big] - 2|\mathrm{Im}\gamma|\Big(-\Delta - \frac{\alpha}{r^2} + x^2\Big) - 2|\mathrm{Im}\gamma|\frac{\alpha}{r^2}\\
&\geqslant (\text{by the uncertainty principle lemma})\\
&\geqslant (a+R)\Big[\Big(-\Delta - \frac{\alpha}{r^2}\Big)^2 + |\gamma|^2x^4\Big] - 4|\mathrm{Im}\gamma|\Big(-\Delta - \frac{\alpha}{r^2}\Big) - 2|\mathrm{Im}\gamma|x^2\\
&\geqslant a\Big[\Big(-\Delta - \frac{\alpha}{r^2}\Big)^2 + |\gamma|^2x^4\Big] - b + R\Big(-\Delta - \frac{\alpha}{r^2}\Big)^2 - 4|\mathrm{Im}\gamma|\Big(-\Delta - \frac{\alpha}{r^2}\Big) + \frac{b}{2} + R\ |\gamma|^2x^4 - 2|\mathrm{Im}\gamma|x^2 + \frac{b}{2}.\end{aligned}$$

A suitable choice of b makes both $R[-\Delta - (\alpha/r^2)]^2 - 4|\mathrm{Im}\gamma|[-\Delta - (\alpha/r^2)] + (b/2)$ and $R|\gamma|^2x^4 - 2|\mathrm{Im}\gamma|x^2 + (b/2)$ positive, and this proves the lemma.

To define $H(\theta)$, we need a further preliminary result.

Lemma 2.4: Let $\alpha < \frac{1}{8}$, $0 < \arg\gamma < 2\pi$, $|\beta| < 1/\sqrt{8}$, $\delta\in C$. Then the maximal multiplication operator by $\beta/r + \delta x$ is relatively bounded with respect to T_1, with relative bound smaller than one.

Proof: By lemma 2.3, we have $D(T_1) = D(H_0)\cap D(x^2) = D(N)$. Now β/r and δx are both relatively bounded with respect to T_1 by the closed graph theorem, since $D(1/r) = Q(1/r^2)\supset Q(-\Delta) = Q(H_0)\supset Q(T_1)\supset D(T_1)$, $D(x)\supset D(x^2)\supset D(T_1)$. Now x is relatively bounded, with relative bound zero, with respect to x^2, and thus by Eq. (2.12) is relatively bounded with respect to T_1 with relative bound zero. Furthermore, applying twice the uncertainty principle inequality we see that $|\beta|^2/r^2$ is relatively form bounded with respect to $-\Delta - (\alpha/r^2)$ with relative bound smaller than 1. Since $-\Delta - (\alpha/r^2)$ is self-adjoint and bounded below for $\alpha < \frac{1}{4}$, $|\beta|/r$ is also relatively bounded with respect to $-\Delta - (\alpha/r^2)$ with relative bound smaller than one. Hence, by Eq. (2.12) it is relatively bounded with respect to T_1 with relative bound smaller than one. Hence, for $u\in D(T_1)$ we have

$$\begin{aligned}\Big\|\frac{\beta}{r}u + \delta xu\Big\| &\leqslant \Big\|\frac{\beta}{r}u\Big\| + \|\delta xu\|\\ &\leqslant a\|T_1u\| + b\|u\| + a'\|T_1u\| + b'\|u\|\\ &\leqslant (a+a')\|T_1u\| + (b+b')\|u\|,\end{aligned}$$

with $a < 1$. Since we can take a' as small as we like, the Lemma is proved.

Theorem 2.2: Let $(Z^2/c^2) < \frac{1}{8}$, $|Z[1 + (2W/c^2)]e^{\theta}|$

$< 1/\sqrt{8}$. Then $H_0(\theta)$ and $\tilde{H}_0(\theta)$ defined on $D(H_0)$ by Eq. (2.11) are a holomorphic family of operators of type A (in W for fixed θ, and in θ for fixed W). The implicit eigenvalues of $H_0(\theta)$ are given by Eq. (2.5).

Proof: We have just seen that β/r is relatively bounded with relative bound smaller than 1 with respect to the self-adjoint operator $-\Delta - (\alpha/r^2)$, $\alpha < \frac{1}{8}$, where $|\beta| < 1/\sqrt{8}$. Hence, the first assertion follows by the standard criterion (see, for example, Ref. 19). The second assertion follows from the usual dilation analyticity arguments (see, for example Ref. 13) which make the eigenvalues of $H_0(\theta)$, and hence also its implicit ones, independent of θ. This proves the Theorem.

Theorem 2.3: Let

$$\frac{Z^2}{c^2} < \frac{1}{8}, \quad \left|Z\left(1+\frac{2W}{c^2}\right)e^{\theta}\right| < \frac{1}{\sqrt{8}}, \quad 0 < \mathrm{Im}\theta < \frac{\pi}{2}. \tag{2.13}$$

Then $H(\theta)$ [and $\tilde{H}(\theta)$] defined by Eq. (2.9) on $D(H_0)\cap D(x^2)$ is a holomorphic family of operators of type A (in θ for fixed W and in W for fixed θ).

Proof: By Lemma 2.4, the operator $T_1 - (\beta/r) + \delta x + (\eta x/r)$ defined on $D(H_0)\cap D(x^2)$ is closed and has a nonempty resolvent set. Then it is enough to take $\alpha = Z^2/c^2$, $\beta = Z[1+(2W/c^2)]e^{\theta}$, $\gamma = (F^2/c^2)e^{4\theta}$, $\delta = F[1+(2W/c^2)]e^{3\theta}$, $\eta = (2ZF/c^2)e^{2\theta}$, and to remark that not only the domain is independent of W and θ but also that $H(\theta)u$ is of course a vector valued holomorphic function of θ and W when $u\in D(H_0)\cap D(x^2)$. Hence, $H(\theta)$ is a holomorphic family of type A by definition,[13] and the theorem is proved.

Remark: From now on, in writing $H(\theta)$ we shall always mean that the conditions (2.13) are satisfied.

As in the nonrelativistic case, the lack of dilation analyticity for θ real makes the limit $\mathrm{Im}\theta\to 0$, which provides the connection between H and $H(\theta)$, a delicate one. We solve this problem along the lines of Herbst's treatment,[9] actually proving in this particular case a slightly stronger statement.

Theorem 2.4: $H(\theta)$ converges strongly in the generalized sense to $H(\theta\in\mathbb{R})$ as $\mathrm{Im}\theta\to 0$.

Proof: The assertion is of course equivalent to the strong resolvent convergence of $\tilde{H}(\theta)$ to $\tilde{H}$, where $\tilde{H}$ stands for $\tilde{H}(\theta)$, $\theta\in\mathbb{R}$. To see this, first remark that the union $\cup$ of the numerical ranges of $\tilde{H}(\theta)$, $0\leqslant \mathrm{Im}\theta < \epsilon < \pi/2$, is not the whole complex plane, as it is easy to check. Hence we can take a point $\lambda\in C$ such that $d(\lambda,\theta) > \eta > 0$, where $d = \mathrm{dist}(\lambda, U)$, and η does not depend on θ. Since $\|[\tilde{H}(\theta)-\lambda]^{-1}\| \leqslant d(\lambda,\theta)^{-1}$, $0\leqslant \mathrm{Im}\theta < \epsilon < \pi/2$, the resolvent is uniformly bounded for $\mathrm{Im}\theta \geqslant 0$ for some $\lambda\in C$. Then the strong resolvent convergence follows by a direct application of theorem VIII.I.5 of Ref. 19, because one has $\|[\tilde{H}(\theta)-\tilde{H}]u\|\to 0$ as $\mathrm{Im}\theta\to 0$ when $u\in D(H_0)\cap D(x^2)$, which is a core of $\tilde{H}$. This proves the theorem.

III. ESSENTIAL SPECTRUM OF $H(\theta)$

In this section we intend to obtain some information on the spectral properties of H and $H(\theta)$, both in the standard sense as well as in the implicit one.

Let T be the essentially self-adjoint operator in L^2 defined as

$$D(T) = D(-\Delta)\cap D(x^2), \quad T = -\Delta - \gamma x^2 + \delta x,$$
$$\gamma > 0, \quad \delta\in\mathbb{R} \tag{3.1}$$

and $T(\theta)$ the m-sectorial operator in L^2 defined as

$$D(T(\theta)) = D(-\Delta)\cap D(x^2),$$
$$T(\theta) = -e^{-2\theta}\Delta - \gamma x^2 e^{2\theta} + \delta x e^{\theta},$$
$$0 < \mathrm{Im}\theta < \pi/2. \tag{3.2}$$

It is well known that $\sigma(T) = \sigma_{\mathrm{ess}}(T) = \mathbb{R}$. The spectrum of $T(\theta)$ is easy to determine.

Lemma 3.1: $\sigma(T(\theta)) \equiv \sigma_{\mathrm{ess}}(T(\theta))$ consists of an infinite family of parallel half-lines emanating from the points $-i\gamma^{1/2}(2n+1) - (\delta^2/4\gamma)$, $n = 0,1,\cdots$, and forming an angle $-2\,\mathrm{Im}\theta$ with the positive real axis.

Proof: We can of course realize $T(\theta)$ as a tensor product

$$T(\theta) = -e^{-2\theta}(p_y^2 + p_z^2)\otimes I$$
$$+ I\otimes(-e^{-2\theta}p_x^2 - \gamma x^2 e^{2\theta} + \delta e^{\theta}x) \tag{3.3}$$

(the meaning of the symbols being obvious). The strict sectoriality of the operators appearing in Eq. (3.3) allows the application of a well known result (see, for example, Ref. 20) stating that the spectrum of the tensor product is the set theoretic sum of the spectra of the separated operators. Now, as is well known, $e^{-2\theta}p_x^2 - \gamma e^{2\theta}x^2 + \delta e^{\theta}x$ defined on $D(p_x^2)\cap D(x^2)$ has a purely discrete spectrum consisting precisely of the simple eigenvalues $-i\gamma^{1/2}(2n+1) - (\delta^2/4\gamma)$, $n = 0,1,\cdots$, so that by the well known nature of the spectrum of $-e^{-2\theta}(p_y^2 + p_z^2)$ the lemma is proved.

Theorem 3.1: $\sigma_{\mathrm{ess}}(H(\theta))$ consists of an infinite family of parallel half-lines emanating from the points $-i(F/c)(2n+1) - (c^2/4)[1+(2W/c^2)]^2 - (W^2/c^2)$, $n = 0,1,2,\cdots$, and forming an angle $-2\,\mathrm{Im}\theta$ with the positive real axis.

Proof: By a well known criterion, it is enough to prove that $[T(\theta)-\lambda]^{-1} - [H(\alpha,\beta,\gamma,\delta,\eta;\theta)-\lambda]^{-1}$ is compact for some $\lambda\in C$, and then to set $\alpha,\beta,\gamma,\delta,\eta$ to their physical values. In this proof $H(\theta)$ is intended of course without the factor $-(W^2/c^2)$. We first remark that the set of the $\lambda\in C$ for which both $[T(\theta)-\lambda]^{-1}$ and $[H(\theta)-\lambda]^{-1}$ exist as bounded operator in L^2 is not empty by the strict sectoriality. Then since $Q(T(\theta)) = Q(-\Delta)\cap Q(x^2) = Q(H(\theta))$, as quadratic forms on $L^2\otimes L^2$, we can write (see Ref. 19, Theorem VII.4.3)

$$[T(\theta)-\lambda]^{-1} - [H(\theta)-\lambda]^{-1}$$
$$= [T(\theta)-\lambda]^{-1}[H(\theta)-T(\theta)][H(\theta)-\lambda]^{-1}$$
$$= [T(\theta)-\lambda]^{-1}\left[-e^{-2\theta}\frac{\alpha}{r^2} - e^{-\theta}\frac{\beta}{r} + \eta\frac{x}{r}\right]$$
$$\times[H(\theta)-\lambda]^{-1}. \tag{3.4}$$

Let us prove the compactness of each term separately. First recall that r^{-1} is compact as a map from $H^{2,2}(\mathbb{R}^3) = D(-\Delta)$ to L^2 and hence, *a fortiori*, as a map from $H^{2,2}\cap D(x^2) = R([T(\theta)-\lambda]^{-1}) = D(T(\theta))$ to L^2. Hence, $r^{-1}[T(\bar{\theta})-\bar{\lambda}]^{-1}$ is compact, and so is its adjoint $[T(\theta)-\lambda]^{-1}r^{-1}$. The middle term is thus the form of a compact operator because $[H(\theta)-\lambda]^{-1}$ is bounded. Apart from multiplicative constants, the third term can be written as

$[T(\theta)-\lambda]^{-1}r^{-1}\cdot x[H(\theta)-\lambda]^{-1}$. As above, $[T(\theta)-\lambda]^{-1}r^{-1}$ is compact, and hence it is enough to have $x[H(\theta)-\lambda]^{-1}$ bounded, which is true by the closed graph theorem since $D(x)\supset D(x^2)\supset D(H(\theta))$. As for the first term, again apart from multiplicative constants, it can be written as $[T(\theta)-\lambda]^{-1}r^{-1}\cdot r^{-1}[H(\theta)-\lambda]^{-1}$. Now the first factor is compact, and the second is bounded again by the closed graph theorem, because $D(r^{-1})\supset D(H(\theta))$ (see the proof of Lemma 2.4). To sum up, the rhs of Eq. (3.4) is the form of a compact operator, and putting $\alpha=Z^2/c^2, \beta=Z[1+(2W/c^2)]$, $\gamma=F^2/c^2, \delta=F[1+(2W/c^2)], \eta=2ZF/c^2$ the theorem is proved.

According to physical intuition the spectrum of H, both in the standard sense as well as in the implicit one, should cover the whole real axis.

In this case the holomorphic dependence of $H(\theta)$ on θ and the known nature of its essential spectrum allow one to prove this assertion.

Theorem 3.2: Let $W\in\mathbb{R}$. Then $\sigma(H(W))=\mathbb{R}$.

Proof: Let (a,b) be an open, bounded interval having empty intersection with $\sigma(H(W))$. Let us prove that $(a,b)=\phi$. To see this, consider the function $f_\psi(\lambda)=\langle\psi,[H(W)-\lambda]^{-1}\psi\rangle$, analytic for $\mathrm{Im}\lambda>0$, ψ being any dilation analytic vector for $|\mathrm{Im}\theta|<(\pi/2)$. By our assumption all functions $f_\psi(\lambda)$ have uniform analytic continuation to the whole complex λ plane cut along $(-\infty,a]$ and $[b,+\infty)$. On the other hand, since ψ is dilation analytic, standard arguments (see, for example, Ref. 13) and Theorem 2.4 show that $\langle\psi,[H(W)-\lambda]^{-1}\psi\rangle$ $=\langle\psi(\bar\theta),[H(W,\theta)-\lambda]^{-1}\psi(\theta)\rangle$, with $\psi(\theta)=U(\theta)\psi$. Since the dilation analytic vectors are dense, by Theorem 3.1 there is at least a ψ for which $f_\psi(\lambda)$ has a singularity when λ is a point of $\sigma_{\text{ess}}(H(\theta))$, which lies in the lower half-plane $\mathrm{Im}\lambda<0$. This contradicts the hypothesis, and the theorem is proved.

As an immediate consequence we have the following:

Corollary 3.1: Let $W\in\mathbb{R}$. Then the implicit spectrum of $H(W)$ covers the whole real axis.

Proof: By the above result, given W the spectrum of $H(W)$ is $\mathbb{R}$, so that in particular it contains the point W itself.

IV. COMPLEX ENERGY EIGENVALUES OF *H(θ)*

Our purpose in this section is to prove the existence of a discrete spectrum of $H(\theta)$ in the energy variable W, to be identified with the resonances in Sec. V. To this end, it is convenient, as in the nonrelativistic case, to realize $H(\theta)$ in squared parabolic coordinates, in order to exploit its invariance with respect to any subspace corresponding to a constant value of the projection of the angular momentum along the x axis.

Introducing the squared parabolic coordinates u, v, ϕ defined by

$$\begin{aligned} u&=(r+x)^{1/2}, \quad x=\tfrac{1}{2}(u^2-v^2),\\ v&=(r-x)^{1/2}, \quad y=uv\sin\phi,\\ \phi&=\arctan\left(\frac{y}{z}\right), \quad z=uv\cos\phi,\\ r&=\sqrt{x^2+y^2+z^2}=\tfrac{1}{2}(u^2+v^2), \end{aligned} \tag{4.1}$$

after separation of the angular part and a unitary transformation one finds (see Ref. 4 for details; compare also with Ref. 7) that the eigenvalue equations $H_0(\theta)\psi=W\psi$ and $H(\theta)=W\psi$ in L^2 [$H(\theta)$ and $H_0(\theta)$ being defined as in Sec. II] are, respectively, equivalent to the following infinite sets of implicit eigenvalue equations in

$L^2(\mathbb{R}^2_{++})=L^2(\mathbb{R}_+\times\mathbb{R}_+)$:

$$H^0_m(\theta)\psi=2Z\psi, \quad m=0,1,\cdots, \tag{4.2}$$

$$H_m(\theta)\psi=2Z\psi, \quad m=0,1,\cdots. \tag{4.3}$$

here $H^0_m(\theta)$ is defined as an operator in $L^2(\mathbb{R}^2_{++})$ by

$$\begin{aligned} &D(H^0_m(\theta))\\ &=\{u\in L^2(\mathbb{R}^2_{++})|ru=\tfrac{1}{2}(u^2+v^2)u\in L^2(\mathbb{R}^2_{++})|u'\equiv\mathrm{grad}u\\ &\in(L^2(\mathbb{R}^2_{++}))^2|\Delta u \text{ exists } |H^0_m(\theta)u\in L^2(\mathbb{R}^2_{++})\} \end{aligned}$$

(differentiations in the generalized sense) with

$$\begin{aligned} H^0_m(\theta)=&-e^{-\theta}\left[\frac{\partial^2}{\partial u^2}+\frac{\partial^2}{\partial v^2}+(m^2-\tfrac{1}{4})\right.\\ &\left.\times\left(\frac{1}{u^2}+\frac{1}{v^2}\right)-\frac{4\alpha}{u^2+v^2}\right]\\ &-e^{\theta}\left(\frac{W^2}{c^2}+W\right)(u^2+v^2) \end{aligned} \tag{4.4}$$

and $H_m(\theta)$ by $D(H_m(\theta))=D(H^0_m(\theta))\cap D((u^2-v^2)^2\times(u^2+v^2))$ with $H_m(\theta)=H^0_m(\theta)+\eta e^\theta(u^2-v^2)+\tfrac{1}{2}e^{2\theta}\delta(u^4-v^4)-\tfrac{1}{4}e^{3\theta}\gamma(u^2-v^2)^2(u^2+v^2)$.

The coefficients $\alpha,\gamma,\delta,\eta$ are given by Eq. (2.9).

Under the same conditions on Z,c,W, and θ of Theorems 2.2 and 2.3, $H^0_m(\theta)-2Z$ and $H_m(\theta)-2Z$ will of course be holomorphic families (of type A) of operators in $L^2(\mathbb{R}^2_{++})$, in W for fixed θ, and in θ for fixed W.

For the sake of simplicity from now on by the generalized spectrum of $H_m(\theta)$ corresponding to the point $\lambda\in c$ we shall mean (compare also Ref. 19, Theorem VII.6) the set of all points $W\in c$ such that λ belongs to the spectrum of $H_m(\theta)$, i.e., the set $\{\lambda\in b(H_m(\theta))|W\in c\}$. An analogous definition for $H^0_m(\theta)$ applies. The implicit W spectrum of $H(\theta)$ is then given by the union over $m\geqslant 0$ of the generalized spectrum of $H_m(\theta)$ corresponding to the point $\lambda=2Z$. Then we have the following:

Theorem 4.1: Let $H^0_m(\theta)$ be defined as a holomorphic family of type A of operators in $L^2(\mathbb{R}^2_{++})$ as above, and let $0<\arg[(W^2/c^2)+W]e^{2\theta}<2\pi$. Then $H^0_m(\theta)$ has a compact resolvent and discrete generalized spectrum. The union over $m\geqslant 0$ of the generalized spectra of $H^0_m(\theta)$ corresponding to the point $\lambda=2Z$ coincides with the unperturbed implicit energy eigenvalues $W^0_{m,l}$ given by Eq. (2.5).

Proof: See the Appendix.

Theorem 4.2: Let $H_m(\theta)$ be defined as a holomorphic family of type A of operators in $L^2(\mathbb{R}^2_{++})$ as above, and let $\mathrm{Re}[-e^{2\theta}[(W^2/c^2)+W]>0$. Then we have the following: (a) $H_m(\theta)$ has compact resolvents; (b) $H_m(\theta)$ has discrete generalized spectrum; (c) any eigenvalue

$\lambda(\alpha,[W^2/c^2)+W],\eta,\delta,\gamma)$ of $H_m(\theta)$ enjoys the Symanzik scaling property

$$\lambda\left(\alpha,\left(\frac{W^2}{c^2}+W\right),\eta,\delta,\gamma\right)$$
$$=\omega^{-1}\lambda\left(\alpha,\omega^2\left(\frac{W^2}{c^2}+W\right),\omega^2\eta,\omega^3\delta,\omega^4\gamma\right) \qquad (4.5)$$

for each $\omega>0$, and elsewhere by analytic continuation. In particular, by setting $\omega=e^{\theta}$ we see that the eigenvalues of $H_m(\theta)$ do not depend on θ. (d) $H_m(\theta)$ converges in the norm resolvent sense to $H_m^0(\theta)$ as $F\to 0$, uniformly on compacts in θ and W. (e) The eigenvalues $\lambda(\cdot,\gamma)$ of $H_m(\theta)$ have an asymptotic expansion to all orders in powers of γ near $\gamma=0$, given by the Rayleigh–Schrödinger perturbation series.

Proof: See the Appendix.

By Eq. (4.3) and Theorem 4.2(a), the energy eigenvalues of the problem are implicitly defined by all equations of the type

$$\lambda(W,F)=2Z, \qquad (4.6)$$

if we require the condition $\mathrm{Re}[-e^{2\theta}[(W^2/c^2)+W]]>0$. Here $\lambda(W,F)$ denotes an arbitrary eigenvalue of $H_m(\theta)$.

As in the nonrelativistic case, we can prove the existence of the implicit function $W=W(F)$, independent of θ, for small values of the field strength F.

Let us first remark that by Theorem 4.2(d) any eigenvalue $\lambda=\lambda(\cdot,F)$ of $H_m(\theta)$ is continuous as $F\to 0$, i.e., its limit exists and is equal to an eigenvalue λ_0 of $H_m^0(\theta)$, for which the equation

$$\lambda_0(W)=2Z \qquad (4.7)$$

can be explicitly solved to yield an unperturbed energy eigenvalue. Then we have the following:

Theorem 4.3: Let $H_m(\theta)$ and $H_m^0(\theta)$ fulfill the conditions of Theorem 4.1 and 4.2, and $\theta<\pi/4$ be fixed. Let in addition $\lambda=\lambda(W,F)$ be an eigenvalue of $H_m(\theta)$, $\lambda_0=\lim_{F\to 0}\lambda$ an eigenvalue of $H_m^0(\theta)$, and W_0 an unperturbed energy eigenvalue determined by λ_0, i.e., Eq. (4.7) holds for $W=W_0$, $\lambda=\lambda_0$. Then there are $\delta>0$ and $\epsilon>0$ such that for $|W-W_0|<\delta$, $0<F<\epsilon$, Eq. (4.6) with the initial condition $\lambda_0(W_0)=2Z$ implicitly defines a function $W=W(F)$, independent of θ, such that $\lim_{F\to 0}W(F)=W_0$.

Proof: First remark that for δ small enough we have $|W|/c^2<\frac{1}{2}$, $(\arg(-W))<(\pi/2)-2\,\mathrm{Im}\theta$, so that the condition $\mathrm{Re}[-e^{2\theta}[(W^2/c^2)+W]]>0$ is satisfied.

Given Eqs. (4.6) and (4.7) for $W=W_0$, the implicit function theorem tells us that the implicit function, which does not depend on θ by the scaling property (4.5) and the fact that also the initial condition does not depend on θ (Theorem 4.1), exists when λ and $\partial\lambda/\partial W$ are analytic in W near W_0 for all $F>0$, $\partial\lambda/\partial W|_{W=W_0}\neq 0$, λ uniformly continuous as $F\to 0$. The anlyticity of λ near W_0 is true because $H_m(\theta)$ is a holomorphic family in W, and the uniform continuity as $F\to 0$ follows from Theorem 4.2(d). To verify the remaining condition, remark that since $H_m(\theta)$ is a holomorphic family of type A in W at fixed θ for any $F>0$, the eigenvalue $\lambda(W,F)$ can be expanded in Taylor series of powers of $(W-W_0)$ near W_0, the nth derivative at W_0 coinciding with the nth coefficient of the Rayleigh–Schrödinger perturbation expansion (times $n!$). Hence, we immediately have for $F>0$:

$$\left.\frac{\partial\lambda}{\partial W}\right|_{W=W_0}=\left\langle\psi(W_0,F,\bar\theta),\left[-e^{\theta}\left(1+\frac{2W_0}{c^2}\right)(u^2+v^2)-\frac{4Z}{c^2}+\frac{e^{2\theta}F}{c^2}(u^4-v^4)\right]\psi(W_0,F,\theta)\right\rangle, \qquad (4.8)$$

$\psi(W_0,F,\theta)$ being the eigenvector corresponding to $\lambda_0(W_0,F)$. Analogously, for $F=0$, we get

$$\left.\frac{\partial\lambda_0(W)}{\partial W}\right|_{W=W_0}=-\left(1+\frac{2W_0}{c^2}\right)\langle\psi(W_0,\bar\theta),e^{\theta}(u^2+v^2)\psi(W_0,\theta)\rangle-\frac{4Z}{c^2}\langle\psi(W_0,\bar\theta),\psi(W_0,\theta)\rangle<0 \qquad (4.9)$$

strictly.

Now the norm resolvent convergence of $H_m(\theta)$ to $H_m^0(\theta)$, uniform in compacts in W, implies that the scalar product appearing in the rhs of Eq. (4.8) is continuous at $F=0$, i.e., its limit exists and is equal to the scalar product in the rhs of Eq. (4.9). This is because

$$\psi(W_0,F,\theta)\underset{F\to 0}{\to}\psi(W_0,\theta),$$

and

$$(u^4-v^4)\psi(W_0,F,\theta)\underset{F\to 0}{\to}(u^4-v^4)\psi(W_0,\theta)$$

by the known exponential fall off of the eigenfunctions at infinity. Hence,

$$\left.\frac{\partial\lambda(W,F)}{\partial W}\right|_{W=W_0}\underset{F\to 0}{\to}\left.\frac{\partial\lambda_0(W)}{\partial W}\right|_{W=W_0}\neq 0.$$

Hence, by continuity, there is $\epsilon>0$ such that $\partial\lambda(W,F)/\partial W|_{W=W_0}\neq 0$ for all F in $0\leqslant F<\epsilon$, and the theorem is proved.

The realization of $H(\theta)$ in squared parabolic coordinates is most convenient also for the determination of the nonrelativistic limit $c\to\infty$. To this end, we briefly recall some results of Ref. 4.

Let $0<\mathrm{Im}\theta<\pi/3$, and $A_m(F,\theta)$, $m=0,1,\cdots$ be the differential operator defined in $L^2(\mathbb{R}_+)$ by

$$D(A_m)=\{f\in L^2(\mathbb{R}_+)|f\in H^{2,2}(\mathbb{R}_+)|f(0)=0|u^4f\in L^2(\mathbb{R}_+)\}, \qquad (4.10)$$

$$A_m=e^{-\theta}\left[-\frac{d^2}{du^2}+(m^2-\tfrac{1}{4})\frac{1}{u^2}\right]-We^{\theta}u^2+\tfrac{1}{2}Fe^{2\theta}u^4.$$

It is proved in Ref. 4 that if the differential expression

$$K_m(\theta)=e^{-\theta}\left[-\frac{\partial^2}{\partial u^2}-\frac{\partial^2}{\partial v^2}+(m^2-\tfrac{1}{4})\left(\frac{1}{u^2}+\frac{1}{v^2}\right)\right]-We^{\theta}(u^2+v^2)+\tfrac{1}{2}Fe^{2\theta}(u^4-v^4)$$

is realized as an operator in $L^2(\mathbb{R}^2_{++})$ by the tensor product

$$K_m(\theta)=A_m(F,\theta)\otimes I+I\otimes A_m(e^{-i\pi}F,\theta), \qquad (4.11)$$

the following properties hold:

(i) $K_m(\theta)$ is a holomorphic family of type A of operators in $L^2(\mathbb{R}^2_{++})$, in θ for any fixed W and in W for any fixed θ.

(ii) $K_m(\theta)$ has a discrete spectrum and discrete generalized spectrum.

(iii) All points W of the generalized spectrum of $K_m(\theta)$ corresponding to the point $2Z$ are resonances of the nonrelativistic essentially self-adjoint Stark effect Hamiltonian $K = -\Delta - (Z/r) + Fx$, $D(K) = C_0^\infty(\mathbb{R}^3)$, and conversely.[7]

(iv) The implicit spectral problem corresponding to the point $2Z$ can be solved, at least for a finite number of eigenvalues of $K_m(\theta)$, when F is small enough, and the resulting resonances converge to the Hydrogen bound states as $F\to 0$.

We now have the following:

Theorem 4.4: Let $0 < \operatorname{Im}\theta < \pi/3$, $\operatorname{Re}[-e^{2\theta}[(W^2/c^2) + W]] > 0$. Then $H_m(\theta)$ converges in the norm resolvent sense to $K_m(\theta)$ as $c\to\infty$, uniformly on compacts in W, θ, and F.

Proof: See the Appendix.

An immediate convergence is as follows:

Corollary 4.1: Let $\lambda(W,F,c)$ be an eigenvalue of $H_m(\theta)$, and $\mu(W,F) = \lim_{c\to\infty} \lambda(W,F,c)$ an eigenvalue of $K_m(\theta)$. Let F be so small that both implicit equations $\lambda(W,F,c) = 2Z$ and $\mu(W,F) = 2Z$ have solutions. Let $W(F,c)$ and $W(F)$ be the corresponding generalized eigenvalues. Then $\lim_{c\to\infty} W(F,c) = W(F)$.

By Theorem 5.3 below, when F is small enough, the implicit eigenvalues $W(F)$ of $H_m(\theta)$ are actually resonances of H. Hence we can conclude the following:

Theorem 4.5: Let $W(F)$ be a resonance of H. Then there are values of F so small that $W(F,c)$ converges to a resonance of the nonrelativistic Stark effect as $c\to\infty$.

V. RESONANCES

We have so far shown the existence of solutions of the implicit spectral problem $H(W,\theta)\psi = W\psi$, $0 < \operatorname{Im}\theta < \pi/2$. These implicit eigenvalues are independent of θ, so that they could be already interpreted as resonances (see Simon[16]). However, the implicit nature of the spectral problem and the lack of dilation analyticity for $\theta\in\mathbb{R}$ make by no means obvious the fact that these eigenvalues are resonances also according the more usual notion of second sheet poles of the scalar products of the resolvent operator taken on some dense set of vectors. (It is well known that this assertion is true for the standard dilation analytic problems; see, for example Ref. 16).

In the present situation, given the implicit nature of the spectral problem, we have to show not only that, when ψ belongs to a suitable dense set in L^2, the function $\langle\psi,[H(W)-\lambda]^{-1}\psi\rangle$, $W\in\mathbb{R}$, which is *a priori* an analytic function of λ for $\operatorname{Im}\lambda > 0$, has a meromorphic continuation to the second sheet $\operatorname{Im}\lambda\leqslant 0$, but also that the second sheet poles coincide with the solutions of the implicit spectral problem for $H(\theta)$ found in the former section. This represents of course a translation into the implicit language of the usual notion of resonance, and hence we go on to prove the above assertions in order to conclude that the W eigenvalues of $H(\theta)$ are resonances of the problem.

Let us first formulate the above remarks in a more mathematical language. Consider the operator $H(\theta) = H(W,F,\theta)$ defined in Sec. II, with $0 < \operatorname{Im}\theta < \pi/4$, $\operatorname{Re}W < 0$, $\operatorname{Im}W\leqslant 0$, $\operatorname{Re}[-e^{2\theta}W[1 + (W/c^2)]] > 0$. By Theorem 4.2, we know that $H(W,F,\theta)$ has a discrete spectrum with this choice of the parameters, independent of θ.

Definition 5.1: By resonances of the self-adjoint operator H defined in Sec. II we mean all eigenvalues $\Lambda = \Lambda(W,F)$ of $H(W,F,\theta)$, $0 < \operatorname{Im}\theta < \pi/4$, $\operatorname{Re}[-e^{2\theta}W[1 + (W/c^2)]] > 0$, $\operatorname{Re}W < 0$, $\operatorname{Im}W\leqslant 0$, satisfying the condition

$$\Lambda(W,F) = W. \tag{5.1}$$

Remark: (a) The resonances as defined above can be characterized also by the implicit functions $W = W(F)$ defined by Eq. (5.1). (b) The implicit W eigenvalues defined by $\lambda(W,F) = 2Z$, λ an eigenvalue of $H_m(W,F,\theta)$, satisfy of course Eq. (5.1), and conversely. Then we have the following

Theorem 5.1: Let $\psi\in L^2$ be a dilation analytic vector for $|\operatorname{Im}\theta| < \pi/4$. Then the function

$$f_\psi(W,F,\lambda) = \langle\psi,[H(W,F)-\lambda]^{-1}\psi\rangle, \tag{5.2}$$

originally defined as an analytic function of λ in the upper half-plane $\operatorname{Im}\lambda > 0$ when $W\in\mathbb{R}$, $F\in\mathbb{R}$, has a meromorphic continuation to the lower half-plane $\operatorname{Im}\lambda\leqslant 0$. For any fixed λ in the lower half-plane, $f_\psi(W,F,\lambda)$ is a meromorphic function of W when $\operatorname{Re}W < 0$, $\operatorname{Im}W\leqslant 0$, $\operatorname{Re}[-W - (W^2/c^2)] < 0$. The resonances of H in the sense of Definition 5.1 are second sheet (i.e., $\operatorname{Im}\lambda = \operatorname{Im}W\leqslant 0$) poles of $f_\psi(W,F,\lambda)$ such that $\Lambda(W,F) = W$, where $\Lambda(W,F)$ is a pole of f_ψ for $\operatorname{Im}\lambda\leqslant 0$.

Proof: By Theorem 2.4 and the standard dilation analyticity arguments (see, for example, Ref. 21), if $W\in\mathbb{R}$, ψ is a dilation analytic vector for $|\operatorname{Im}\theta| < \pi/4$, and $\psi(\theta) = U(\theta)\psi$, we have

$$f_\psi(W,F,\lambda) = \langle\psi(\bar\theta),[H(W,F,\theta)-\lambda]^{-1}\psi(\theta)\rangle, \tag{5.3}$$

whence the analyticity properties in W and λ by the properties of the operator families $H_m(W,F,\theta)$ stated in Theorem 4.2. Furthermore, the resonances in the sense of Definition 5.1 are by direct inspection second sheet poles of $f_\psi(W,F,\lambda)$ satisfying the constraint $\Lambda(W,F) = W$, and the theorem is proved.

To complete the argument showing that the implicit W eigenvalues of $H(W,F,\theta)$ whose existence has been shown in Sec. IV are resonances of H according to Definition 5.1, it remains to be proved that, at least for F small, one has $\operatorname{Im}W(F)\leqslant 0$. To this end, let us first prove a preliminary proposition.

Lemma 5.1: Let $\Lambda(W,F)$ be an eigenvalue of $H(\theta)$ satisfying the constraint $\Lambda(W,F) = W$, $W = W(F)$, the implicit function thus defined, and W_0 an unperturbed energy eigenvalue [i.e., an implicit eigenvalue of $H_0(\theta)$] such that $\lim_{F\to 0} W(F) = W_0$, and let $g(W,F) = \Lambda(W,F) - W$. Then there are $\delta > 0$ and $\epsilon > 0$ such that $\partial g/\partial W \neq 0$ for all $|W - W_0| < \delta$, when $0 < F < \epsilon$.

Proof: We know that $H(\theta)$ is for any fixed θ a holomorphic family of type A in W when $|Z[1 + (2W/c^2)]e^\theta| < 1/\sqrt{8}$. Hence, proceeding as in the proof of Theorem 4.3, we can conclude that $\partial\Lambda/\partial W|_{W=W_0}$ coincides with the coefficient of $(W - W_0)$ in the Rayleigh–Schrödinger expansion of $\Lambda(W,F)$ of initial point W_0. Hence, we have

$$\frac{\partial g}{\partial W}\bigg|_{W=W_0} = \frac{\partial \Lambda(W,F)}{\partial W}\bigg|_{W=W_0} - 1$$
$$= -\frac{2}{c^2}\langle \psi(W_0,\bar\theta), V\psi(W_0,\theta)\rangle - 1,$$

where $V = (-Fxe^{\theta} + (Z/r)e^{-\theta} + W_0)$, and $\psi(W,\theta)$ denotes an eigenvector corresponding to $\Lambda(W,F)$. The same argument, applied now to $H_0(\theta)$, yields

$$\frac{\partial \Lambda_0}{\partial W}\bigg|_{W=W_0} - 1$$
$$= -1 - \frac{2}{c^2}\left\langle \psi_0(W_0,\bar\theta), \frac{Z}{r}e^{-\theta}\psi_0(W_0,\theta)\right\rangle$$
$$-2W_0\langle \psi_0(W_0,\bar\theta), \psi_0(W_0,\theta)\rangle$$
$$< -\frac{2}{c^2}W_0 - 1 < -\frac{2}{c^2}\left(W_0 + \frac{c^2}{2}\right) < 0,$$

where $\Lambda_0 = \lim_{F\to 0}\Lambda(W,F)$, and $\psi_0(W,\theta)$ denotes the eigenvector corresponding to Λ_0. Now, as in Theorem 4.3, $\partial\Lambda(W,F)/\partial W \to_{F\to 0} \partial\Lambda_0/\partial W$. Hence by continuity, there is $\delta > 0$ such that $\partial g(W,F)/\partial W \neq 0$ for $|W - W_0| < \delta$, and this proves the Lemma.

We are now in position to prove that for F small enough the imaginary part of the implicit eigenvalues found in Sec. IV has the correct sign.

Theorem 5.2: Let $\Lambda(W,F)$ be an eigenvalue of $H(\theta)$, $0 < \text{Im}\theta < \pi/4$, satisfying the constraint $\Lambda(W,F) = W$. Then there are values of F so small that if $W(F)$ denotes the corresponding implicit function, $W(F) \to W_0$ as $F \to 0$, W_0 an implicit eigenvalue of $H_0(\theta)$, one has $\text{Im}W(F) \leqslant 0$.

Proof: Consider first the implicit eigenvalue problem for $H(W,F,\theta)$ defined as $\text{Re}\Lambda(W,F) = W$, $\Lambda(W,F)$ as above. This problem has a solution under the present assumptions, because $\partial\,\text{Re}\Lambda(W,F)/\partial W - 1 \neq 0$ for W real, $|W - W_0| < \delta$, since $\partial\,\text{Re}\Lambda(W,F)/\partial W = \text{Re}\partial\Lambda(W,F)/\partial W$ for W real, and $\text{Re}\partial\Lambda(W,F)/\partial W - 1 \neq 0$ by Lemma 5.1. Denote by $W_1 = W_1(F)$ the solution, of course real, of $\text{Re}\Lambda(W,F) = W$. Now $H(W)$ is essentially self-adjoint for W real, and since by Theorem 2.4 and the usual dilation analyticity arguments we have

$$\langle \Phi, [H(W) - \lambda]^{-1}\Phi\rangle = \langle \Phi(\bar\theta), [H(W,\theta) - \lambda]^{-1}\Phi(\theta)\rangle$$

for a suitable dense set of dilation analytic vectors $\{\Phi\}$, it must be $\text{Im}\Lambda(W_1(F),F) \leqslant 0$.

Let us now relate $\Lambda_1 \equiv \Lambda_1(W_1(F),F)$ to $\Lambda \equiv \Lambda(W(F),F) \equiv W(F)$. As in Lemma 5.1 and Theorem 4.3, the analyticity of $H(W,\theta)$ allows one to write, through first-order perturbation theory,

$$\Lambda - \Lambda_1 = -\frac{(W - W_1)}{c^2}C_1(W_1,F) + O(|W - W_1|^2)$$
$$= \frac{(\Lambda - \text{Re}\Lambda_1)}{c^2}C_1(W_1,F) + O(\Lambda - \text{Re}\Lambda_1)^2, \quad (5.4)$$

where

$$C_1(W_1,F) = 2\langle \psi(W_1,F,\bar\theta), (-Fx\,e^{\theta} + (Z/r)e^{-\theta} + W_1)\psi(W_1,F,\theta)\rangle > 2W_0$$

for F small by the same continuity argument of Lemma 5.1 and Theorem 4.3. Taking the real part of Eq. (5.4) we get

$$\text{Re}(\Lambda - \Lambda_1)\left(1 + \frac{C_1}{c^2}\right) + O\,(\text{Re}(\Lambda - \Lambda_1)) = O(\text{Im}\Lambda)^2.$$

For F small, we have $1 + (C_1/c^2) + O(\text{Re}(\Lambda - \Lambda_1)) > 1 + 2W_0/c^2 > 0$, so that $\text{Re}(\Lambda - \Lambda_1) = O(\text{Im}\Lambda)^2$.

Taking the imaginary part of Eq. (5.4) we get

$$\text{Im}\Lambda\left[1 + \frac{C_1}{c^2} + O(\text{Im}\Lambda)\right] = \text{Im}\Lambda_1,$$

so that $\text{Im}\Lambda/\text{Im}\Lambda_1 > 0$ for F small if $\text{Im}\Lambda_1 \neq 0$. When $\text{Im}\Lambda_1 = 0$ we have $\Lambda = \Lambda_1$ so that we can conclude that there are values of F so small that $\text{Im}\Lambda \leqslant 0$ in any case. The Theorem is proved.

Remark: By Remark (b) after Definition 5.1, the implicit eigenvalues of $H_m(\theta)$ of Sec. IV satisfy the conditions of Theorem 5.2. Hence, they are resonances of $H(W)$ according to Definition 5.1.

ACKNOWLEDGMENTS

It is a pleasure to thank Professor A. Grossmann for his interest in this work. Two of us (S.G. and V.G.) are grateful to C.P.T., C.N.R.S., Marseille, for hospitality.

APPENDIX

Consider in $L^2(\mathbb{R}^2_{++}) = L^2(\mathbb{R}_+ \times \mathbb{R}_+)$ the operator defined as follows:

$$S_m = -\Delta + (m^2 - \tfrac{1}{4})\left(\frac{1}{x^2} + \frac{1}{y^2}\right) - \frac{4\alpha}{x^2 + y^2} + \gamma'(x^2 + y^2),$$

$$D(S_m) = \{u \in L^2(\mathbb{R}^2_{++}) | u' \equiv \text{grad}\,u \in (L^2(\mathbb{R}^2_{++}))^2 | \Delta u \text{ exists} | \left(-\Delta + (m^2 - \tfrac{1}{4})\left(\frac{1}{x^2} + \frac{1}{y^2}\right) - \frac{\alpha}{x^2 + y^2}\right) \times u \in L^2(\mathbb{R}^2_{++})\} \cap D(x^2 + y^2), \quad (A1)$$

$$m = 0, 1, \cdots.$$

(Differentiations in the generalized sense). Then we have the following:

Lemma A.1: Let $\alpha < 1/16$. Then S_m is a holomorphic family of type A of compact resolvent operators when $|\arg\gamma'| < \pi$.

Proof: Consider in $L^2(\mathbb{R}^4)$ the operator B defined by

$$B = -\Delta_1 - \Delta_2 + \gamma'(x^2 + y^2), \quad x^2 = x_1^2 + x_2^2,$$
$$y^2 = y_1^2 + y_2^2, \quad -\Delta_1 = -\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2},$$
$$-\Delta_2 = -\frac{\partial^2}{\partial y_1^2} - \frac{\partial^2}{\partial y_2^2},$$
$$D(B) = H^{2,2}(\mathbb{R}^4) \cap D(x^2 + y^2).$$

It is well known that B is a holomorphic family (of type A) of strictly m-sectorial compact resolvent operators for $|\arg\gamma'| < \pi$, with quadratic form domain given by

$$Q(B) = H^{2,1}(\mathbb{R}^4) \cap L^{2,1}(\mathbb{R}^4) = Q(-\Delta_1 - \Delta_2) \cap Q(x^2 + y^2).$$

By the uncertainty principle lemma, the maximal multiplication operator by $4\alpha/(x^2 + y^2)$ is relatively form bounded with relative bound smaller than one with respect to B if

$\alpha < 1/16$. Then we can define $C = B - 4\alpha/(x^2+y^2)$, $Q(C) = Q(B)$, as the form sum of B and $4\alpha/(x^2+y^2)$, and C is again strictly m sectorial with compact resolvents. Through a quadratic estimate analogous to Lemma 2.1, one easily proves that

$$D(C) = D(x^2+y^2) \cap \{u \in L^2(\mathbb{R}^4) | u' = \text{grad} u \in (L^2(\mathbb{R}^4))^4 | (\Delta_1 + \Delta_2)u \text{ exists} | Bu - 4\alpha/(x^2+y^2)u \in L^2(\mathbb{R}^4)\},$$

which does not depend on γ'. Then the Lemma is proved just by remarking that S_m is nothing else than the "radial part" of C with respect to the change of variables

$$(x_1,x_2) \to (x,\phi_1), \quad \phi_1 = \arctan \frac{x_2}{x_1},$$

$$(y_1,y_2) \to (y,\phi_2), \quad \phi_2 = \arctan \frac{y_2}{y_1},$$

i.e., the restriction of C to the invariant subspaces spanned by the functions of the form $f(x,y)\, e^{\pm im(\phi_1 + \phi_2)}$.

Proof of Theorem 4.1: By Eqs. (4.4) and (A1), we have $H^0_m(\theta) = e^{-\theta} S_m$ if we set $\gamma' = -e^{2\theta}[W + (W^2/c^2)]$. Hence the discreteness of the generalized spectrum follows by a well known result (see Ref. 19, Theorem VII.1.10). Its explicit form for $\lambda = 2Z$ is of course due to the exact solvability of the problem, known without the separation into squared parabolic coordinates.

Consider now in $L^2(\mathbb{R}^2_{++})$ the operator A_m defined as follows:

$$A_m = T_m + i\,\text{Im}\gamma'(x^2+y^2) + \beta' V,$$

$$T_m = S_m(\text{Re}\gamma'), \quad V = (x^2-y^2)^2(x^2+y^2),$$

$$\beta' > 0, \quad \text{Re}\gamma' > 0,$$

$$D(A_m) = D(S_m) \cap D((x^2-y^2)^2(x^2+y^2)) = D(S_m) \cap D(V).$$

Let us prove the following quadratic estimate (compare with Lemma II.9.1 of Ref. 15).

Lemma A.2: Let $u \in D(S_m) \cap D(V)$. Then there are $a > 0$, $b > 0$, independent of β' and of γ', when γ' ranges on the compacts of the half-plane $\text{Re}\gamma' > 0$, such that

$$a[\|T_m u\|^2 + |\text{Im}\gamma'|^2 \|(x^2+y^2)u\|^2 + \beta'^2 \|Vu\|^2] \leqslant \|A_m u\|^2 + b\,\|u\|^2. \tag{A2}$$

Proof: As quadratic forms on $D(S_m) \cap D(V) \otimes D(S_m) \cap D(V)$ we have

$$\begin{aligned} A^*_m A_m &= T^2_m + |\text{Im}\gamma'|^2 (x^2+y^2)^2 + \beta'^2 V^2 \pm i\,\text{Im}\gamma'[T_m,(x^2+y^2)] + \beta'(T_m V + V T_m) \\ &= T^2_m + |\text{Im}\gamma'|^2 (x^2+y^2)^2 + \beta'^2 V^2 \pm i\,\text{Im}\gamma'[T_m,(x^2+y^2)] + \beta'(-\Delta V + V(-\Delta)) \\ &\quad + 2\beta'\left[(m^2 - \tfrac14)\left(\frac{1}{x^2} + \frac{1}{y^2}\right) - \frac{4\alpha}{x^2+y^2} + \text{Re}\gamma'(x^2+y^2)\right] V. \end{aligned}$$

Now

$$\begin{aligned} \pm i\,\text{Im}\gamma'[T_m,(x^2+y^2)] &= \pm i\,\text{Im}\gamma'[-\Delta,(x^2+y^2)] \geqslant \pm |\text{Im}\gamma'| \{-i\nabla\cdot(2x,2y) + (2x,2y)\cdot - i\nabla\} \\ &= |\text{Im}\gamma'| \{-i\nabla \pm (2x,2y)\}^2 - |\text{Im}\gamma'|[-\Delta + 4x^2 + 4y^2] \geqslant -|\text{Im}\gamma'|(-\Delta + 4x^2 + 4y^2). \end{aligned}$$

Furthermore

$$\begin{aligned} \beta'[-\Delta V + V(-\Delta)] &= \beta'[-i\nabla,[-i\nabla,(x^2-y^2)^2(x^2+y^2)]] + 2\beta' i\nabla(x^2-y^2)^2(x^2-y^2) i\nabla \\ &= \beta'(-28x^4 - 28y^4 + 24x^2y^2) + 2\beta' i\,\nabla(x^2-y^2)^2(x^2-y^2) i\nabla, \end{aligned}$$

$$\begin{aligned} &2\beta'\left[(m^2 - \tfrac14)\left(\frac{1}{x^2} + \frac{1}{y^2}\right) - \frac{4\alpha}{x^2+y^2} + \text{Re}\gamma'(x^2+y^2)\right](x^2-y^2)^2(x^2+y^2) \\ &\quad = 2\beta'(m^2 - \tfrac14)(x^2-y^2)^2 \frac{(x^2+y^2)^2}{x^2y^2} + 2\beta'\,\text{Re}\gamma'(x^2+y^2)^2(x^2-y^2)^2 - 8\alpha\beta'(x^2-y^2)^2, \end{aligned}$$

whence

$$\begin{aligned} A^*A &\geqslant T^2_m + |\text{Im}\gamma'|^2(x^2+y^2)^2 + \beta'^2 V^2 - |\text{Im}\gamma'|(-\Delta + 4x^2 + 4y^2) \\ &\quad + \beta'[24x^2y^2 - 28(x^4+y^4)] + 2\beta' i\,\nabla(x^2-y^2)^2(x^2+y^2) i\nabla \\ &\quad + 2\beta'(m^2 - \tfrac14)(x^2-y^2)^2 \frac{(x^2+y^2)^2}{x^2y^2} + 2\beta'\,\text{Re}\gamma'(x^2+y^2)^2(x^2-y^2)^2 - 8\alpha\beta'(x^2-y^2)^2 \\ &\geqslant T^2_m + |\text{Im}\gamma'|^2(x^2+y^2)^2 + \beta'^2(x^2-y^2)^4(x^2+y^2)^2 - |\text{Im}\gamma'|(-\Delta + 4x^2 + 4y^2) - 28\beta'(x^4+y^4) \\ &\quad - \tfrac12\beta'(x^2-y^2)^2 \frac{(x^2+y^2)^2}{x^2y^2} + 2\beta'\,\text{Re}\gamma'(x^2+y^2)^2(x^2-y^2) - 8\alpha\beta'(x^2-y^2)^2 \\ &\geqslant (\text{for some } a < 1 \text{ and some } R, 0 < R < 1-a) \\ &\geqslant a[T^2_m + |\text{Im}\gamma'|^2(x^2+y^2)^2 + \beta'^2(x^2-y^2)^4(x^2+y^2)^2] + R\,[T^2_m + |\text{Im}\gamma'|^2(x^2+y^2)^2 + \beta'^2(x^2-y^2)^4(x^2+y^2)^2] \\ &\quad - |\text{Im}\gamma'|(-\Delta + 4x^2 + 4y^2) - \tfrac12\beta'(x^2-y^2)^2 \frac{(x^2+y^2)^2}{x^2y^2} - 28\beta'(x^4+y^4) - 8\alpha\beta'(x^2-y^2)^2 \\ &\quad + 2\beta'\,\text{Re}\gamma'(x^2+y^2)^2(x^2-y^2)^2 \\ &\geqslant a[T^2_m + |\text{Im}\gamma'|^2(x^2+y^2)^2 + \beta'^2(x^2-y^2)^4(x^2+y^2)^2] - b + \left[R T^2_m - |\text{Im}\gamma'|(-\Delta + 4x^2 + 4y^2) + \frac{b}{4}\right] \\ &\quad + \left[\tfrac12\beta'\text{Re}\gamma'(x^2+y^2)^2(x^2-y^2)^2 - 28\beta'(x^4+y^4) + \frac{b}{4}\right] + \left[\tfrac12\beta'\,\text{Re}\gamma'(x^2+y^2)^2(x^2-y^2)^2 - 8\alpha\beta'(x^2-y^2)^2 + \frac{b}{4}\right] \end{aligned}$$

$$+\left[\beta'\,\mathrm{Re}\gamma'(x^2+y^2)^2(x^2-y^2)^2-\tfrac{1}{2}\beta'(x^2-y^2)^2\frac{(x^2+y^2)^2}{x^2y^2}+\frac{b}{4}\right].$$

It remains to be proved that a suitable choice of b makes the last four terms positive. The first term is positive by taking $b/4$ not smaller than the lowest eigenvalue of the positive self-adjoint operator $|\mathrm{Im}\gamma'|(-\Delta+4x^2+4y^2)$. The second term is positive if $28\beta'(x^4+y^4)\leqslant\frac{1}{2}\beta'\,\mathrm{Re}\gamma'(x^2+y^2)^2(x^2-y^2)^2+(b/4)$, i.e., $(x^4+y^4)/(x^2-y^2)^2\leqslant(\mathrm{Re}\gamma'/56)(x^4+y^4)+(\mathrm{Re}\gamma'/28)x^2y^2+[(b)/4\beta'(x^2-y^2)^2]$ which is true for some b large enough, uniformly with respect to β' in any open interval $(0,\beta'_0)$. The third term is clearly positive for b large enough, which can be chosen independently of β' in any interval $(0,\beta'_0)$. The fourth term is positive when

$$\frac{x^4+y^4}{2x^2y^2}+1\leqslant\mathrm{Re}\gamma'(x^4+y^4)+2\,\mathrm{Re}\gamma'x^2y^2+\frac{b}{4\beta'(x^2-y^2)^2},$$

which is true for b large enough, again independent of β' for $0<\beta'<\beta'_0$. The Lemma is proved.

As a consequence we have the following, in analogy with Theorem II.9.2 of Ref. 15:

Theorem A.1: Le $\beta'>0$, $\mathrm{Re}\gamma'>0$. Then A_m is closed on $D(S_m)\cap D(V)$, and represents a holomorphic family of type A in γ' for $\mathrm{Re}\gamma'>0$.

Lemma A.3: The maximal multiplication operators by $(x^2-y^2)^2$, x^2-y^2, x^4-y^4 are relatively bounded with respect to A_m, with relative bound zero, uniformly on compacts in $\mathrm{Re}\gamma'>0$, $0<\beta'<\beta_0$.

Proof: The assertion is a consequence of Eq. (A2) because

$$(x^2-y^2)^2,(x^2-y^2),(x^4-y^4)=(x^2-y^2)(x^2+y^2)$$

are all relatively bounded with respect to $V=(x^2-y^2)^2(x^2+y^2)$ with relative bound zero.

We can now define

$$V_m=A_m+\eta'(x^2-y^2)+\delta'(x^2-y^2),\quad \eta'\in C,\ \delta'\in C \tag{A3}$$

as a closed operator on $D(A_m)=D(S_m)\cap D(V)$. By Lemma A.3 and problem IV.1.2 of Ref. 19, we can immediately conclude the following:

Lemma A.4: The assertion of Lemma A.3 remains true if we replace A_m by V_m.

A computation completely analogous to that performed in obtaining Eq. (A2) yields the following for $\beta'=0$:

Lemma A.5: Let $\mathrm{Re}\gamma'>0$. Then the maximal multiplication operator by (x^2+y^2) in $L^2(\mathbb{R}^2_{++})$ is relatively bounded with respect to S_m.

Lemma A.6 (Symanzik scaling property): Let $\lambda=\lambda(\alpha,\beta',\gamma',\delta',\eta')$ be an eigenvalue of $V_m(\alpha,\beta',\gamma',\delta',\eta')$. Then, if $\omega>0$, $\omega^{-1}\lambda(\alpha,\omega^4\beta',\omega^2\gamma',\omega^3\delta',\omega^2\eta')$ is an eigenvalue of $\omega^{-1}V_m(\alpha,\omega^4\beta',\omega^2\gamma',\omega^3\delta',\omega^2\eta')$.

Proof: Exactly as in Theorem II.2.1 of Ref. 15.

Remark: The above relation extends to all complex ω whenever the analytic continuation is possible.

Lemma A.7: Let $\delta'=\mu_1\sqrt{\beta'}$, $\eta'=\mu_2\sqrt{\beta'}$, $|\arg\mu_1|<\pi$, $|\arg\mu_2|<\pi$. Then V_m converges in the norm resolvent sense to S_m as $\beta'\to0$, uniformly on compacts for $\mathrm{Re}\gamma'>0$, $|\arg\mu_1|<\pi$, $|\arg\mu_2|<\pi$.

Proof: Let us proceed as in Ref. 15, Lemma II.9.3. First remark that the union $\cup$ of the numerical ranges of V_m for $\beta'\geqslant0$ in the above described regions is not the whole complex plane, so that $\|[V_m-\lambda]^{-1}\|$ is bounded uniformly in β' for some λ. For these values of λ we can write

$$\begin{aligned}(V_m-\lambda)^{-1}-(S_m-\lambda)^{-1}&=-(S_m-\lambda)^{-1}[\eta'(x^2-y^2)+\delta'(x^4-y^4)+\beta'(x^2-y^2)^2(x^2+y^2)](V_m-\lambda)^{-1}\\&=-\mu_1\sqrt{\beta'}(S_m-\lambda)^{-1}[(x^2-y^2)(V_m-\lambda)^{-1}]-\mu_2\sqrt{\beta'}\,[(S_m-\lambda)^{-1}(x^2+y^2)]\\&\quad\times[(x^2-y^2)(V_m-\lambda)^{-1}]-\beta'[(S_m-\lambda)^{-1}(x^2+y^2)][(x^2-y^2)^2(V_m-\lambda)^{-1}].\end{aligned}$$

By Lemmas A.4 and A.5 each operator within the square bracket is bounded, so that we can repeat word by word the remaining part of the argument of Lemma II.9.3 of Ref. 15 and the Lemma is proved.

We are now in a position to prove Theorem 4.2.

Proof of Theorem 4.2: By rescaling the phase of $-e^{4\theta}\equiv e^{i\pi}e^{4\theta}$, instead of $e^{\theta}H_m(\theta)$ we can equivalently look at

$$\begin{aligned}&H_m\left(\left|\frac{\gamma}{4}e^{4\theta}\right|,\mathrm{Re}\theta\right)\\&=e^{i[(\pi/4)+\mathrm{Im}\theta]}\cdot e^{-2\theta}\left\{\left[-\Delta+(m^2-\tfrac{1}{4})\left(\frac{1}{x^2}+\frac{1}{y^2}\right)-\frac{4\alpha}{x^2+y^2}\right]+ie^{2\,\mathrm{Re}\theta}\left(\frac{W^2}{c^2}+W\right)(x^2+y^2)\right.\\&\quad\left.-e^{-i(\pi/4)}e^{\mathrm{Re}\theta}\frac{4ZW}{c^2}-i\,e^{2\,\mathrm{Re}\theta}\eta(x^2-y^2)+e^{-i(3\pi/4)}e^{3\,\mathrm{Re}\theta}\frac{\delta}{2}(x^4-y^4)+\left|\frac{\gamma}{4}e^{4\theta}\right|(x^2-y^2)^2(x^2+y^2)\right\}.\end{aligned}$$

Then $H_m(|(\gamma/4)e^{4\theta}|,\mathrm{Re}\theta)$ coincides with V_m times a constant if we set

$$\gamma'=ie^{2\,\mathrm{Re}\theta}\left(\frac{W^2}{c^2}+W\right),\quad \beta'=\left|\frac{\gamma}{4}e^{4\theta}\right|,\quad \eta'=-i\,e^{2\,\mathrm{Re}\theta}\eta,\quad \delta'=e^{-i(3\pi/4)}e^{3\,\mathrm{Re}\theta}\delta/2.$$

Hence assertion (d) of Theorem 4.2 is Lemma A.7, and assertion (b) is Lemma A.6. Assertion (a) is a consequence of (d), given the fact that $H_m(\theta)$ is a holomorphic family of type A. Assertion (b) is a consequence of (d), and again of the holomorphic nature (of type A) of the operator family $H_m(\theta)$ (see Ref. 19, Theorem VII.1.10).

Finally, we omit the proof of part (e), since given the above assertions it is an almost word by word repetition of the arguments of Ref. 15, Sec. II.10.

Proof of Theorem 4.4: Let θ belong to the stated region. Again by the uniform sectoriality of $H_m(\theta)$ there are values of λ for which $\|[H_m(F,\theta,c)-\lambda]^{-1}\|$ is bounded uniformly with respect to c.

Then, proceeding as in Lemma A.7, write

$$\begin{aligned}
&[H_m(F,\theta,c)-\lambda]^{-1}-[K_m(F,\theta)-\lambda]^{-1}\\
&\quad=-\frac{Z^2}{c^2}e^{-2\theta}[K_m(F,\theta)-\lambda]^{-1}(x^2+y^2)^{-1}[H_m(F,\theta,c)-\lambda]^{-1}\\
&\quad-\frac{W^2}{c^2}e^{2\theta}[K_m(F,\theta)-\lambda]^{-1}(x^2+y^2)[H_m(F,\theta,c)-\lambda]^{-1}\\
&\quad-\frac{4ZW}{c^2}e^{\theta}[K_m(F,\theta)-\lambda]^{-1}[H_m(F,\theta,c)-\lambda]^{-1}\\
&\quad+\frac{2ZF}{c^2}e^{2\theta}[K_m(F,\theta)-\lambda]^{-1}(x^2-y^2)[H_m(F,\theta,c)-\lambda]^{-1}\\
&\quad+\frac{W}{c^2}e^{3\theta}[K_m(F,\theta)-\lambda]^{-1}(x^4-y^4)[H_m(F,\theta,c)-\lambda]^{-1}\\
&\quad-\frac{F^2}{4c^2}e^{4\theta}[K_m(F,\theta)-\lambda]^{-1}(x^2-y^2)^2(x^2+y^2)[H_m(F,\theta,c)-\lambda]^{-1}.
\end{aligned}$$

Now the inequality $(x^2+y^2)^{-1}\leqslant(1/x^2)+(1/y^2)$ clearly shows that first term is bounded; x^2+y^2 is relatively bounded with respect to $K_m(\theta)$ (see, for example, Ref. 4), so that $[K_m(F,\theta)-\lambda]^{-1}(x^2+y^2)$ is bounded, and the same is true for $[K_m(F,\theta)-\lambda]^{-1}(x^2+y^2)$. Hence, the first four terms vanish in the norm as $c\to\infty$. Now recall (see always Ref. 4) that also (x^4-y^4) is relatively bounded with respect to $K_m(F,\theta)$; since x^2+y^2 is by Lemma A.4 relatively bounded with respect to $H_m(F,\theta,c)$ uniformly on c, also the last two terms vanish in the norm as $c\to\infty$ and the theorem is proved.

[1] J.E. Avron and I.W. Herbst, Commun. Math. Phys. **52**, 239 (1977).
[2] L. Benassi, V. Grecchi, E. Harrell, and B. Simon, Phys. Rev. Lett. **42**, 704 (1979).
[3] L. Benassi and V. Grecchi, J. Phys. B (in press).
[4] S. Graffi and V. Grecchi, Commun. Math. Phys. **62**, 83 (1978).
[5] S. Graffi and V. Grecchi, J. Phys. B **12**, L265 (1979).
[6] S. Graffi, V. Grecchi, S. Levoni, and M. Maioli, J. Math. Phys. **20**, 685 (1979).
[7] S. Graffi, V. Grecchi, and B. Simon, J. Phys. A **12**, L193 (1979).
[8] I.W. Herbst, Commun. Math. Phys. **53**, 285 (1977).
[9] I.W. Herbst, Commun. Math. Phys. **64**, 179 (1977).
[10] I.W. Herbst and B. Simon, Phys. Rev. Lett. **42**, 81 (1978).
[11] I.W. Herbst and B. Simon (in preparation).
[12] H.A. Bethe, *Intermediate Quantum Mechanics* (Benjamin, New York, 1965).
[13] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1978), Vol. IV.
[14] R. Weder, J. Funct. Anal. **20**, 319 (1975).
[15] B. Simon, Ann. Phys. (N.Y.) **58**, 76 (1970).
[16] B. Simon, Ann. Math. **97**, 247 (1973).
[17] W. Faris and R. Lavine, Commun. Math. Phys. **35**, 39 (1974).
[18] E. Nelson, J. Funct. Anal. **11**, 211 (1972).
[19] T. Kato, *Perturbation Theory of Linear Operators* (Springer, New York 1966).
[20] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1975), Vol. II.
[21] E. Balslev and J.M. Combes, Commun. Math. Phys. **22**, 280 (1971).
[22] L.I. Schiff, *Quantum Mechanics* (Mc Graw-Hill, New York, 1949).

Continuum calculus. IV. The Laplace transform method in the evaluation of the Feynman path integrals with a Gaussian measure and applications in quantum mechanics

L. L. Lee
School C.E.M.S., University of Oklahoma, Norman, Oklahoma 73019

(Received 31 August 1978; accepted for publication 9 April 1979)

The continuum calculus, proposed earlier [J. Math. Phys. **17**, 1988 (1976)], is applied here to the development of a functional version of Laplace transform and a method for the evaluation of path integrals containing a Gaussian-like measure. Two methods in functional integration are proposed. First, the Gaussian integral for polynomial functionals and consequently functionals that can be expanded in Taylor series are examined. Formula (2.12) is derived. Next, we define the Laplace transform in the function space through the weak distribution formulation of Skorohod. Comparison of both approaches enables us to determine the expression of the functional integeral through a series of Laplace transformations. The second formula is given in Eq. (4.5). The latter formula is applicable to all Laplace transformable functionals. For illustration of the utility of the formulas derived, we evaluate the integral of a cosine functional by methods 1 and 2, and obtain consistent results. Further applications to quantum mechanics are also presented. We examine the cases of a free particle,the quantum harmonic oscillator, the forced oscillator, and charged particle in a magnetic field. In all cases, we obtain correct results in comparison with known expressions. A numerical procedure is employed in the calculation of infinite products. The usefulness of the p-integral method is stressed.

I. INTRODUCTION

The conventional Feynman[1,2] path integral in quantum mechanics involves a kinetic energy term and a potential energy term in the Lagrangian. The nonrelativistic kinetic energy is usually quadratic in momentum. This term acts like a Gaussian (or pseudo-Gaussian) measure on the rest of the integral. Therefore path integrals in the function space with Gaussian-like measures are important and are most widely studied.[3]

In a series of papers[4–6] (referred hereafter as papers I, II, and III, respectively) we proposed an operational calculus, called the continuum calculus, designed to treat the integral of functionals. In this approach we are able to integrate functionals of various types with or without a Gaussian measure. In addition, the integral can have finite limits of integration. Therefore the calculus proposed is of a more general nature. Tests with probability theory and known physical formulas gave valid results.[4] In this investigation, we propose to look specifically at Gaussian integrals and apply the continuum calculus method.

We define the class of functionals to be investigated. Consider a Banach space, B, over the complex number field C. Let C^B be the collection of all functions, $y(t)$, from B to C, i.e., $y:B\to C$. We construct a Banach algebra, A_B, from C^B by the usual process of completion and definitions of sums and products of elements of C^B. We further require that B and A_B be Hausdorff in the norm topologies. Two measure spaces (B,S_B,μ) and (A_B,S_A,m) are defined on B and A_B, respectively, with S_B and S_A the σ algebras and μ, m the measures on B and A_B, respectively. A functional, $\phi\,[y]$, is a *form* from A_B to C, $\phi:A_B\to C$. A functional integral is generally of the form,

$$I_f[\phi\,] = \int_F m(dy)\cdot\phi\,[y], \quad F\subseteq A_B, \tag{1.1}$$

for some measure m on S_A. When the measure is of the Gaussian form,

$$m(dy) = \bar{m}(dy)\exp\Big[-\tfrac{1}{2}\int dr\,ds\,y(r)A\,(r,s)y(s)\Big], \tag{1.2}$$

and the ϕ functional can be represented as a certain integral, we have

$$I_f[\phi\,] = \int_F \bar{m}(dy)\exp\Big[-\tfrac{1}{2}\int dr\,ds\,y(r)A\,(r,s)y(s)\Big] \times \int_E \mu(dt)f(y(t)), \tag{1.3}$$

where $E\subseteq B, f:C\to C$. (1.3) is a Gaussian functional integral. $A\,(r,s)$ is the covariance matrix. When A is positive definite, we have a strictly Gaussian process. Otherwise, we have a Gaussian-like process. For example, in Feynman integrals, A can be imaginary. In this paper, we shall investigate integrals of type (1.3).

We start out with an analysis on the polynomial functionals, which have been studied extensively.[7–9] With results from polynomial functionals, we can next treat functionals that can be expanded into functional Taylor series[10] (Sec. II). We obtain formula (2.12) which is an infinite series in terms of the functional derivatives of ϕ. In order to generalize to a wider class of functionals, we propose a functional Laplace transform method in Secs. III and IV. Later demonstrations show that this method is not only more general, but also more powerful than the functional Taylor expansion method. From Sec. V onwards, we present illustrations and applications of the methods proposed. A mathematical functional, the cosine functional, is integrated by both methods and results in the same form. The quantum mechanics of a free particle, and harmonic oscillator is examined in Secs. VII

 0022-2488/80/051055-15$01.00

and VIII. The properties of the kinetic energy matrix are important in the final results and are examined carefully in Sec. VI. Its eigenvalues are displayed in Table I and Fig. 1. As further applications we treat the forced harmonic oscillator in Sec. X, and a charged particle in a magnetic field in Sec. XI. In all cases, correct results are obtained in comparision with known formulas. Due to the absence of analytical expression for the spectrum of the kinetic energy matrix, we evaluated a number of p integrals numerically. They converged rapidly to known functions.

In the following, we review previous developments briefly. The continuum calculus consists of two operations: (i) the r differentiation, and (ii) the p integration. In approximate language, the r derivative (*rational* derivative), denoted by Rf/Rt, of a function $f(t)$, is the ratio of the function at two neighboring points $[f(t+\Delta t)/f(t),\ f(t)\neq 0]$; in contrast to the differential derivative which is a measure of the difference of the function at two neighboring points $[f(t+\Delta t)-f(t)]$. The two kinds of derivatives are related by a correspondence theorem (paper I[4])

$$\frac{Rf}{Rt}(t)=\exp\Big(\frac{d}{dt}\ln f(t)\Big). \tag{1.4}$$

The p integral (*potential* integral) arose from a search for the primitive of the r differentiation. Another correspondence theorem[4] relates it directly to the ordinary integral,

$$\mathscr{P}_E\, dt \gg g(t)=\exp\int_E dt\,\ln g(t). \tag{1.5}$$

The functional integral is obtained as a consequence of the interaction between the p integration and the ordinary integration. For an exponential functional

$$\psi[y]=\exp\int_E \mu(dt)\cdot f(y(t)) \tag{1.6}$$

we have developed the formula for the functional integral, I_f

$$I_f[\psi]=\exp\int_E \mu(dt)\ln\int_F m(dy)\exp[f(y)]. \tag{1.7}$$

For details and proofs, see paper I.[4]

II. GAUSSIAN FUNCTIONAL INTEGRAL

In this section, we restrict B to be a real Banach space over the real number field, R. For convenience in notation, we take $B=R$. A_B is the complex Banach algebra, and ϕ is the functional, $\phi: A_B\to C$. For a general functional, $\phi[y]$, it is a difficult problem to evaluate its functional integral. We therefore first consider here the special class of functionals, $\{\phi\}$, that can be represented in functional Taylor series,[10] i.e.,

$$\begin{aligned}\phi[y]=\phi[y=0]&+\int\mu(ds_1)\frac{\delta\phi}{\delta y(s_1)}\Big|_0 y(s_1)\\&+\frac{1}{2!}\int\mu(ds_1)\mu(ds_2)\\&\times\frac{\delta^2\phi}{\delta y(s_1)\delta y(s_2)}\Big|_0 y(s_1)y(s_2)+\cdots\\&+\frac{1}{n!}\int\mu(ds_1)\ldots\mu(ds_n)\frac{\delta^n\phi}{\delta y(s_1)\ldots\delta y(s_n)}\Big|_0\\&\times y(s_1)\ldots y(s_n)+\cdots,\end{aligned} \tag{2.1}$$

where $\delta\phi/\delta y(s)$, etc., are functional derivatives; the subscript 0 denotes that the derivatives are evaluated at $y(s)\equiv 0$. For convenience, we shall use Borel measure ds instead of $\mu(ds)$ in the future to simplify writing.

Each term on the right-hand side of (2.1) is recognized as a so-called polynomial functional. The Gaussian integral of polynomial functionals has been studied.[7-9] In general, it contains a k-ic defined as,

$$P_k[y]\equiv\int ds_1\ldots ds_k\,K(s_1,\ldots,s_k)y(s_1)\ldots y(s_k),\quad k=0,1,2,\cdots. \tag{2.2}$$

For kernels, $K(s_1,\ldots,s_k)$, that are invariantly traceable,[8] the Gaussian integral,

$$I_f[P_k]=\int m(dy)\exp\Big[-\tfrac{1}{2}\int(dt)A(t,t)y(t)^2\Big]\times\Big[\int ds_1\ldots ds_k\,K(s,\ldots,s_k)\cdot y(s_1)\ldots y(s_k)\Big] \tag{2.3}$$

can be evaluated. Due to the property of Gaussian integral, polynomial k-ics contribute only when k is even, $k=2n$,

$$I_f[P_{2n}]=\Big(\frac{2\pi}{\det A}\Big)^{1/2}\frac{(2n)!}{n!2^n}(\mathrm{Tr})^n(P_{2n}),\quad n=0,1,2,\cdots \tag{2.4}$$

and $I_f[P_k]=0$ for k odd. In (2.4), the symbol, $(\mathrm{Tr})^n$, indicates the nth order trace of the kernel.[8] It is defined recursively,

$$(\mathrm{Tr})^0\equiv 1, \tag{2.5}$$

$$(\mathrm{Tr})^n(P_m)\equiv(\mathrm{Tr})((\mathrm{Tr})^{n-1}(P_m)),\quad \text{for } m\geqslant 2n, \text{ and } n=1,2,3,\cdots. \tag{2.6}$$

When the covariance matrix $A(t,t)$ is the unit matrix, $\delta(s,t)$, Friedrichs[8] gave the formula,

$$\begin{aligned}\mathrm{Tr}[P_m]_{\text{Friedrichs}}&=\mathrm{Tr}\int ds_1\ldots ds_m\,K(s_1,\ldots,s_{m-2},s_{m-1},s_m)y(s)\ldots y(s_m)\\&=\int ds_1\ldots ds_{m-2}\,dt\,K(s_1,\ldots,s_{m-2},t,t)y(s_1)\cdots y(s_{m-2}).\end{aligned} \tag{2.7}$$

In our case the (diagonalized) covariance matrix is $A(t,t)$ with eigenvalues $a(t)$. (2.7) must be modified to give

$$\mathrm{Tr}\,[P_m]\equiv\int ds_1\ldots ds_{m-2}\,dt\,K(s_1,\ldots,s_{m-2},t,t)\times y(s)\ldots y(s_{m-2})/a(t). \tag{2.8}$$

Here we have assumed, without loss of generality, that $A(t,t)$ is diagonal. For a Gaussian integral in the strict sense, we require that the covariance matrix be positive definite. However, for Gaussian-like measures, (e.g., in Feynman path integrals where the matrix can be imaginary), we relax this condition. The eigenvalues of a continuum matrix (matrix with continuous indices) were defined in a previous work.[4] (2.8) is consistent with the known discrete formula of moments,

$$\int_{-\infty}^{+\infty}dx\,x^{2n}\exp(-\tfrac{1}{2}bx^2)=\frac{(2n)!}{n!2^nb^n}\Big(\frac{2\pi}{b}\Big)^{1/2}. \tag{2.9}$$

Now we are in a position to apply (2.4) to the Taylor series (2.1). We observe that (2.1) consists of a sum of 0-ics, 1-ics,..., and n-ics, with the functional derivatives as kernels. Due to the commutativity in the order of partial differentiations (under suitable continuity conditions) the kernels are totally symmetric, e.g., if

$$K(q,r,s,t) \equiv \frac{\delta^4 \phi}{\delta y(q) \delta y(r) \delta y(s) \delta y(t)}, \tag{2.10}$$

then

$$K(q,r,s,t) = K(r,q,s,t) = K(t,s,r,q) = K(\text{perm } qrst), \tag{2.11}$$

where perm $qrst$ means all permutations of the four arguments (q,r,s,t). The functional integral of $\phi[y]$ is then

$$\begin{aligned} I_f[\phi] &= \left(\frac{2\pi}{\det A}\right)^{1/2} \left(P_0 + 0 + \frac{1}{2!}\frac{2!}{1!2}\,\mathrm{Tr}(P_2) + 0 + \frac{1}{4!}\frac{4!}{2!2^2}\,(\mathrm{Tr})^2(P_4) + \cdots\right) \\ &= \left(\frac{2\pi}{\det A}\right)^{1/2} \left(P_0 + \frac{1}{1!2}\,\mathrm{Tr}(P_2) + \frac{1}{2!2^2}\,(\mathrm{Tr})^2(P_4) + \frac{1}{3!2^3}\,(\mathrm{Tr})^3(P_6) + \cdots\right) \\ &= \left(\frac{2\pi}{\det A}\right)^{1/2} \sum_{n=0}^{\infty} \frac{1}{n!2^n}\,(\mathrm{Tr})^n(P_{2n}), \end{aligned} \tag{2.12}$$

where

$$P_0 = \phi_0, \tag{2.13}$$

$$\mathrm{Tr}(P_2) = \int ds\, K(s,s) a(s)^{-1}. \tag{2.14}$$

Here we write $K(s,s)$ for the derivative $\delta^2\phi/\delta y(s)^2$. In general, we shall write $K(s_1 \ldots s_n)$ for the derivative $\delta^n \phi/\delta y(s_1) \ldots \delta y(s_n)$. The higher order traces require some care here. Due to the different treatments of the Gaussian integral to different moments in the random variable, and to maintain consistency with the discrete multivariate case, we rederive the expressions for higher order traces in (2.12), which represent modifications of the Friedrichs formulas.[8]

$$\begin{aligned} (\mathrm{Tr})^2(P_4) = \int ds\, K(ssss) a(s)^{-2} + \frac{2!2!2!}{4!1!1!} &\times \sum_{\text{perm } rr'} \int dr\, dr'\, K(rrr'r')[1 - \delta(rr')] \\ &\times [a(r)a(r')]^{-1}, \end{aligned} \tag{2.15}$$

where $\Sigma_{\text{perm } rr'}$ denotes the sum of distinct permutations of the arguments $rrr'r'$; e.g., $rr'rr'$, $rr'r'r$, etc. The factor $1 - \delta(rr')$, where $\delta(rr')$ is the Dirac delta, is to remove the diagonal elements $K(rrrr)$ which have already been accounted for in the first term on the right-hand side. The fact that this separation is necessary can be seen from the discrete case that

$$\int_{-\infty}^{+\infty} dy\, y^4 \exp\left(-\frac{a}{2}y^2\right) = \frac{4!}{2!2^2a^2}\left(\frac{2\pi}{a}\right)^{1/2}, \tag{2.16}$$

while

$$\int_{-\infty}^{+\infty} dx\, dy\, x^2 y^2 \exp\left(-\frac{a}{2}x^2 - \frac{b}{2}y^2\right) = \frac{2!2!}{1!2\cdot 1!2ab}\left(\frac{2\pi}{ab}\right)^{1/2}. \tag{2.17}$$

The coefficient of the fourth moment is different from that of the two second moments. A study of Gaussian integrals by Kollmann[9] also indicated this difference. Higher order traces can be similarly analyzed. In general,

$$\begin{aligned} (\mathrm{Tr})^n(P_{2n}) = &\int dr\, K(r \cdots r) a(r)^{-n} + \sum_{\text{perm } rr'} \frac{n!2!(2n-2)!}{(2n)!1!(n-1)!} \int dr\, dr'\, K(r \ldots rr'r')\,[1 - \delta(rr')]\, a(r)^{-n+1} a(r')^{-1} + \cdots \\ &+ \sum_{\substack{\text{perm} \\ r_1 \ldots r_n}} \frac{n!(2!)^n}{(2n!)(1!)^n} \int dr_1 \ldots dr_n\, K(r_1 r_1 \ldots r_n r_n) \left(1 - \sum_{i \neq j}\sum \delta(r_i r_j) - \sum_{i \neq j \neq k}\sum\sum \delta(r_i r_j)\delta(r_j r_k) - \ldots - \prod_{i \neq j} \delta(r_i r_j)\right) \prod_i a(r_i)^{-1}. \end{aligned} \tag{2.18}$$

Equation (2.12) is the desired result as the integral of the functional (2.1) under a Gaussian measure. We note that the traces are dependent upon the covariance matrix, A, of the Gaussian measure. This is natural since we expect the functional integral to be different for different Gaussian process.

We have succeeded in deriving a formula for the integral of functionals that can be expressed in Taylor series. However, we would like to extend it to more general cases. This is achieved by using the method of functional Laplace transform to be discussed in the next section.

III. FUNCTIONAL LAPLACE TRANSFORM

In this section we shall establish the Laplace transform in function space and exhibit some of its properties. The language of weak distributions of Skorohod[11] is again appropriate here and we shall follow his terminology. Let X be a real separable Hilbert space and S the σ algebra of measurable sets of X. Let L be a finite dimensional subspace of X under the action of the projection operator P_L. If A is a subset of L, the cylinder set, A_c, in X is defined as,

$$A_c \equiv \{x \in X:\ P_L x \in A\} \tag{3.1}$$

and A is called the base of A_c. The collection of all cylinder sets with bases in L form a σ algebra, $S_L \subset S$. Let $\{L_n\}$ be an increasing sequence of subspaces of X such that $L_{n-1} \subset L_n$, and $\cup_n L_n$ is dense in X. Then the σ closure of the corresponding σ algebras $\cup_n S_{L_n}$ gives S. Let m be some normalized measure on (X,S). This measure induces on L a projection measure m_L defined by

$$m_L(A) \equiv m(\{x \in X: P_L x \in A\}), \quad A \in S_L. \tag{3.2}$$

The family of projection measures m_L defined on all finite dimensional subspaces, L, of X and satisfying a compatibility condition (see Skorohod[11]) is called a weak distribution, $m_* = \{m_L\}$. Under suitable conditions, this weak distribution converges to m on (X,S) (see Lemma 1, p. 5, Skorohod[11]). In this paper we are not concerned with this approach. We shall follow instead the continuum calculus formulation developed so far.

To apply to our functional case, we identify X with A_B and set $B = R$. The inner product of $y, z \in A_B$ is defined as,

$$(y,z) \equiv \int_R \mu(dt) y(t) z(t). \tag{3.3}$$

To guarantee the existence of the inner product, we may restrict A_B further to be L_2. Now the Laplace transform of a functional, $\phi[y]: A_B \to C$, is given by

$$\mathscr{L}\phi[z] = \tilde{\phi}[z] \equiv \int_{X^+} m(dy) \exp[-(y,z)]\cdot\phi[y], \tag{3.4}$$

where X^+ is the collection of functions $y \in A_B$, $X^+ = \{y \in A_B : y(t) \geqslant 0, \ \forall t \in R\}$. In weak distribution language,

$$\tilde{\phi}[z] \doteq \int m_*(dy) \exp[-(y,z)_*]\, \phi_*[y] \tag{3.5}$$

where the subscript, $*$, denotes the weak distribution counterparts of the respective quantities. For example, the quadratic functional,

$$\gamma[y] \equiv \int dr\, K(r,r)\, y(r)^2 \tag{3.6}$$

under the action of the projection operator, P_n, where $P_n(y) = \{y_1, ..., y_n\}$ becomes the cylinder functional

$$\gamma_n[y] = \sum_j^n K_{jj} y_j^2. \tag{3.7}$$

The corresponding Laplace transform in discrete notation is

$$\int m_n(dy) \exp\Big(-\sum_j y_j z_j\Big)\cdot\gamma_n[y]. \tag{3.8}$$

Application of Fubini theorem[12] gives

$$\int_0^\infty dy_1\, e^{-y_1 z_1} \int_0^\infty dy_2\, e^{-y_2 z_2} \cdots \int_0^\infty dy_n\, e^{-y_n z_n}\Big(\sum_j K_{jj} y_j^2\Big)$$

$$= \Big(\prod_{j=1}^n z_j^{-1}\Big)\Big(\sum_{k=1}^n \frac{2! K_{kk}}{z_k^2}\Big), \tag{3.9}$$

where we have taken m_n to be the Borel measure. We obtained an n product and an n sum. As $n \to \infty$, the n sum becomes the Riemann integral,

$$\sum_{k=1}^n \frac{2! K_{kk}}{z_k^2} \xrightarrow{n\to\infty} 2! \int dr\, K(rr) z(r) \quad . \tag{3.10}$$

As for the n product, the p integral in continuum calculus[4] can be applied to give

$$\prod_{j=1}^n z_j^{-1} \xrightarrow{n\to\infty} \exp\Big[-\int dt\, \ln z(t)\Big]. \tag{3.11}$$

As $n \to \infty$, we pass, in the sense of continuum calculus, from m_* to m. Therefore, we define the Laplace transform of this quadratic functional to be

$$\mathscr{L}\gamma[z] \equiv 2! \exp\Big[-\int dt\, \ln z(t)\Big]\cdot\int dr\, K(rr) z(r)^{-2}. \tag{3.12}$$

This example is illuminating. We thus can generalize the same procedure to other polynomial functionals. For example, let

$$P_k'[y] \equiv \int dr\, K(r...r) y(r)^k. \tag{3.13}$$

The Laplace transform, after going through the necessary algebra, is

$$\mathscr{L}P_k'[z] = k!\, \exp\Big[-\int dt\, \ln z(t)\Big]\cdot\int dr\, K(r\cdots r) z(r)^{-k}. \tag{3.14}$$

This enables us to give the following definition for a general polynomial functional.

Definition 3.A: The functional Laplace transform of polynomial functionals. Given a k-ic of the form

$$P_k[y] \equiv \int ds_1 ... ds_k\, K(s_1, ..., s_k)\, y(s_1) ... y(s_k), \tag{3.15}$$

the Laplace transform is defined as

$$\mathscr{L}P_k[z] \equiv \Big[-\int dt\, \ln z(t)\Big]\cdot\Big\{k! \int dr\, K(r,...,r) z(r)^{-k} + (k-1)!1! \sum_{\text{perm.}} \int dr\, ds\, K(r...rs)[1-\delta(rs)]$$
$$\times z(r)^{k-1} z(s)^{-1} + \cdots + (1!)^k \sum_{\text{perm.}} \int ds_k ... ds_k\, K(s_1, ..., s_k)\Big[1 - \sum_{i\neq j}\sum \delta(s_i s_j) - \cdots - \prod_{i\neq j}\delta(s_i s_j)\Big]\prod_m z(s_m)^{-1}\Big\} \tag{3.16}$$

when the integrals exist. The symbol Σ_{perm} denotes the sum of distinct permutations of the arguments of $K(s_1 ... s_k)$. The factors, $1 - \delta(rs)$, etc., are designed to remove the diagonal elements of $K(rs)$, etc., respectively.

The capability of defining Laplace transforms for polynomial functionals enables us to define Laplace transforms for functionals expressible as Taylor series, e.g., (2.1). In fact we can do better. By application of the weak distribution formalism, we can treat more general functionals. However, in the following, we return to the study of functional integrals.

IV. ALTERNATIVE FORMULATION

In this section we shall develop an alternative way of evaluating the integral of type (1.3). The use of the Laplace transform in the function space accrues to the functional integral in two ways at least: (i) for those functionals whose Laplace transforms are known, this method avoids the evaluation of multiple integrals (k-ics) and the summation of infinite series; (ii) it generalizes the integral to functionals that can be Laplace transformed but may not possess a Taylor series expansion, thus enlarging the class of functionals that can be integrated with a Gaussian measure.

We have already developed a formula for polynomial functional, Definition 3.A. This can now be applied to the Taylor series (2.1). Since our eventual Gaussian integral retains only the terms of even powers of the random variable, we Laplace transform only the even polynomial functionals in the Taylor series. We shall call this subseries $\phi_E[y]$. We have then,

$$\tilde{\phi}_E[z] = \exp\Big(-\int dt \ln z(t)\Big)\Big\{\phi_0 + \frac{1}{2!}\Big[2!\int dr\, K(rr)\, z(r)^{-2}\Big] + \frac{1}{4!}\Big[4!\int dr\, K(rrrr)z(r)^{-4} + 2!2!\sum_{\text{perm.}}\int dr\, ds\, K(rrss)[1-\delta(rs)]\, z(r)^{-2}\Big] z(s)^{-2} + \ldots\Big\}. \tag{4.1}$$

Now we multiply both sides of (4.1) by $\exp[-\int dt \ln z(t)]$ and transform the variable from z to w by $w(t) = z(t)^2$,

$$\exp\Big(-\frac{1}{2}\int dt \ln w(t)\Big)\cdot\tilde{\phi}_E[\sqrt{w}] = \exp\Big[-\int dt \ln w(t)\Big]\cdot\Big\{\phi_0 + \int dr\, K(rr) w(r)^{-1} + \int dr\, K(rrrr)\, w(r)^{-2} + \frac{2!2!}{4!}\sum_{\text{perm.}}\int dr\, ds\, K(rrss)\,[1-\delta(rs)]\, w(r)^{-1}\, w(s)^{-1} + \ldots\Big\}. \tag{4.2}$$

we now define the inverse Laplace transform $\mathscr{L}^{-1}$ as the natural extension of the discrete case and the weak distribution expression:

$$\mathscr{L}^{-1}\tilde{\phi}[x] \equiv (2\pi i)^{-1}\int_{W^*} m(dw)\, e^{-(w,x)}\cdot\tilde{\phi}[w], \tag{4.3}$$

where $w(t)$ is complex and ranges over the half-plane, $W^* = \{\alpha(t)+\beta(t)i: -\infty \leqslant \beta(t) \leqslant +\infty\}$. Application of (4.3) to (4.2) gives

$$\mathscr{L}^{-1}\Big\{\exp\Big[-\frac{1}{2}\int dt \ln w(t)\Big]\cdot\tilde{\phi}_E[\sqrt{w}]\Big\}[x] = \phi_0 + \int dr\, K(rr)x(r) + \frac{1}{2!}\int dr\, K(rrrr)x(r)^2$$
$$+ \frac{1}{4!}\frac{2!2!}{1!1!}\sum_{\text{perm.}}\int dr\, ds\, K(rrss)x(r)x(s)\,[1-\delta(rs)] + \cdots + \frac{1}{n!}\int dr\, K(rrr\ldots r)x(r)^n + \frac{1}{(2n)!}\frac{(2n-2)!2!}{(n-1)!1!}$$
$$\times\sum_{\text{perm.}}\int dr\, ds\, K(r\ldots rss)\,[1-\delta(rs)]\, x(r)^{n-1}x(s) + \cdots + \frac{(2!)^n}{(2n!)(1!)^n}\sum_{\text{perm.}}\int dr_1\ldots dr_n\, K(r_1r_1\ldots r_nr_n)$$
$$\times\Big[1 - \sum_{i\neq j}^{n}\delta(r_ir_j) - \cdots - \prod_{i\neq j}^{n}\delta(r_ir_j)\Big]\,[x(r_1)\ldots x(r_n)]^{-1} + \cdots. \tag{4.4}$$

Comparison of (4.4) with (2.12) shows that if we replace $x(r)$ in (4.4) by $[2a(r)]^{-1}$, then the two expressions give the same result, i.e.,

$$I_f[\phi] = (2\pi/\det A)^{1/2}\cdot\mathscr{L}^{-1}\Big\{\exp\Big[-\frac{1}{2}\int dt \ln w(t)\Big]\cdot\tilde{\phi}_E[\sqrt{w}]\Big\}[x]\Big|_{x=(2a)^{-1}}. \tag{4.5}$$

We have obtained a methodology of finding the functional integral through the Laplace transform route. We shall show in the following sections that indeed both methods give identical results for know functional integrals. The advantage of the Laplace transform alternative has been mentioned, in that it can be applied to more general functionals.

We summarize the steps in using the Laplace method:

(i) Extract the symmetrical part $\phi_E[y]$ out of a given functional $\phi[y]$. $\phi_E[y]$ has the property that, for any t,

$$|\phi_E[y]|_{-y(t)} = \phi_E[y], \tag{4.6}$$

where the subscript means that at t, the original value of y at t is replaced by its negative value.

(ii) Take the Laplace transform $\tilde{\phi}_E$ of ϕ_E.

(iii) Multiply $\tilde{\phi}_E[z]$ by $\exp[-\int dt \ln z(t)]$.

(iv) Transform the variables from $z(t)$ to $w(t)$ by $w(t) = z(t)^2$.

(v) Take Laplace inverse transform with respect to the new variable $w(t)$,
$\mathscr{L}^{-1}\{\exp[-\frac{1}{2}\int dt \ln w(t)]\cdot\tilde{\phi}_E[(w)^{1/2}]\}\cdot[x]$.

(vi) The functional integral of the original $\phi[y]$ is obtained by substituting $[2a(r)]^{-1}$ for $x(r)$ above and by multiplying the resulting expression by the factor $(2\pi/\det A)^{1/2}$. (4.7)

In the following sections, we shall test both methods on some mathematical functionals and a few known problems in quantum mechanics. In the known cases, valid results are obtained.

V. THE COSINE FUNCTIONAL

In the following, we examine a functional of the type,

$$\phi[y] = \int_0^1 dr\, B(r)\cos[y(r)]. \tag{5.1}$$

This cosine functional is interesting because it is simple enough that it can be expanded into Taylor series and it also has a functional Laplace transform. Thus we can calculate the Gaussian integral using both methods developed so far.

The Gaussian integral is written as,

$$I_f[\phi] = \int \bar{m}(dy)\exp\Big[-\frac{1}{2}\int dt\, A(t,t)y(t)^2\Big]$$

$$\cdot \int dr\, B(r) \cos[y(r)]. \tag{5.2}$$

First, we apply the Taylor series method. Equation (5.1) is expanded into the following series:

$$\phi[y] = \phi_0 + \int dr \frac{\delta\phi}{\delta y(r)}\bigg|_0 y(r) + \frac{1}{2!}\int dr\, ds \frac{\delta^2\phi}{\delta y(r)\delta y(s)}\bigg|_0 y(r)y(s) + \frac{1}{3!}\cdots, \tag{5.3}$$

where the functional integrals are evaluated as,

$$\frac{\delta\phi}{\delta y(s)} = -\int dr\, B(r) \sin[y(r)]\, \delta(r,s) = -B(s)\sin[y(s)] \tag{5.4}$$

$$\frac{\delta^2\phi}{\delta y(s)\delta y(t)} = -B(t)\cos[y(t)]\delta(s,t) \tag{5.5}$$

$$\frac{\delta^3\phi}{\delta y(s)\delta y(t)\delta y(u)} = B(t)\sin[y(t)]\,\delta(s,t)\delta(t,u), \quad \text{etc.} \tag{5.6}$$

Setting $y(r) = 0$ causes all odd order derivatives to vanish. Equation (5.3) becomes

$$\phi[y] = \phi_0 - \frac{1}{2!}\int dr\, B(r)y(r)^2 + \frac{1}{4!}\int dr\, B(r)y(r)^4 + \cdots + \frac{(-1)^n}{(2n)!}\int dr\, B(r)y(r)^{2n} + \cdots. \tag{5.7}$$

Application of the formula for $I_f[\phi]$ of (2.12) gives

$$\begin{aligned} I_f[\phi] &= (2\pi/\det A)^{1/2}\Big\{\int dr\, B(r) - \int dr\, B(r)[2a(r)]^{-1} + \frac{1}{2!}\int dr\, B(r)[2a(r)]^{-2} - \frac{1}{3!}\int dr\, B(r) \\ &\quad \times [2a(r)]^{-3} + - \ldots\Big\} \\ &= (2\pi/\det A)^{1/2} \cdot \int dr\, B(r) \cdot \Big\{1 - [2a(r)]^{-1} + \frac{1}{2}[2a(r)]^{-2} - \frac{1}{3!}[2a(r)]^{-3} + \ldots\Big\} \\ &= (2\pi/\det A)^{1/2} \cdot \int_0^1 dr\, B(r)\exp\Big[-\frac{1}{2a(r)}\Big], \end{aligned} \tag{5.8}$$

where $a(r)$ is the spectrum of the covariance matrix $A(t,t)$. The result is surprisingly simple, because we can resum the infinite series into an exponential.

Now we try the alternative method of the Laplace transform. Since the given cosine functional is already an even functional, $\phi_E[y] = \phi[y]$. We follow the steps starting from (ii) in (4.7). The Laplace transform of (5.1) is

$$\tilde{\phi}[z] = \exp\Big[-\int dt \ln z(t)\Big] \cdot \int_0^1 dr\, B(r)\Big(\frac{z(r)^2}{z(r)^2 + 1}\Big). \tag{5.9}$$

Next, multiply by $\exp[-\int dt \ln z(t)]$ and change to variable $w(t) = z(t)^2$,

$$\exp\Big[-\frac{1}{2}\int dt \ln w(t)\Big]\tilde{\phi}[\sqrt{w}] = \exp\Big(-\int dt \ln w(t)\Big)\int_0^1 dr\, B(r)\Big[\frac{w(r)}{w(r)+1}\Big]. \tag{5.10}$$

The inverse Laplace transform of (5.10) is known from mathematical tables,[13,14]

$$\mathscr{L}^{-1}\Big\{\exp\Big[-\frac{1}{2}\int dt \ln w(t)\Big]\tilde{\phi}[\sqrt{w}]\Big\}[x] = \int_0^1 dr\, B(r)\exp[-x(r)]. \tag{5.11}$$

Substituting $[2a(r)]^{-1}$ for $x(r)$ and multiplying by $(2\pi/\det A)^{1/2}$ give

$$(2\pi/\det A)^{1/2}\mathscr{L}^{-1}\Big\{\exp\Big[-\frac{1}{2}\int dt \ln w(t)\Big]\tilde{\phi}[\sqrt{w}]\Big\}[x]\bigg|_{x=(2a)^{-1}} = \Big(\frac{2\pi}{\det A}\Big)^{1/2}\int dr\, B(r)\exp\Big[-\frac{1}{2a(r)}\Big]. \tag{5.12}$$

Equation (5.12) is precisely the integral (5.8). Therefore both the Taylor expansion method and the Laplace transform method gave the same result for (5.1).

The Laplace transform method holds certain advantages over the Taylor series. With know Laplace transform for the functional under study, the process of getting the integral is much simplified. There is no need of evaluating the functional derivatives, collecting the traces, and resumming the infinite series. The Laplace method also generalizes at once to cases where the Laplace transforms exist for the given functionals, which may not possess Taylor series expansions. We shall make full use of these properties in future studies.

VI. THE COVARIANCE MATRIX AND KINETIC ENERGY

The Lagrangian of a particle undergoing translational motion in a potential field, $V(x)$ (for illustration, we consider here a one-dimensional one-body system) is

$$L(x, \dot{x}, t) = \tfrac{1}{2}m\dot{x}^2 - V(x). \tag{6.1}$$

The propagator kernel is given[2] by a path integral,

$$K(x_b, t_b; x_a, t_a) = B\int Dx(t)\exp\Big\{i\hbar^{-1}\int_{t_a}^{t_b} dt\Big[\frac{1}{2}m\dot{x}(t)^2 - V(x(t))\Big]\Big\}, \tag{6.2}$$

where m is the mass of the particle, $x(t)$ is the path as a function of time, t, which starts at x_a at time t_a and ends at x_b at time t_b. B is the normalization constant. Comparing expression (6.2) with the standard form of a Gaussian integral, (1.3), shows that the covariance matrix in this case is given by

$$\int dr\, ds\, y(r)A(r,s)y(s) = m(i\hbar)^{-1}\int dt\, \dot{x}(t)^2, \tag{6.3}$$

i.e., it is the kinetic energy term in the Lagrangian. Our developments show that the eigenvalues of this covariance matrix play an important role in the functional integral [see, e.g., Eqs. (2.18), and (4.5)]. Since the kinetic energy term appears constantly in Lagrangian and Hamiltonian mechanics, we devote a few paragraphs to the discussion of its eigenvalues.

First we note that the Gaussian measure can be written in the Riemann sum,

$$\int_0^T dr\, ds\, y(r) A(r,s) y(s) = \lim_{\Delta t \to 0} \sum_i \sum_j \Delta t^2 y_i A_{ij} y_j, \quad (6.4)$$

where $T = N\Delta t$, $y_i = y(i\Delta t)$, $A_{ij} = A(i\Delta t, j\Delta t)$. The limit is taken as $N \to \infty$, and $\Delta t \to 0$, while keeping $N\Delta t = T$. Now if we consider the velocity, $\dot{x}(i\Delta t)$, as given by $\lim[x((i+1)\Delta t) - x(i\Delta t)]/\Delta t$, the rhs of (6.3) can be written as,

$$m(i\hbar)^{-1} \lim_{\Delta t \to 0} \frac{1}{\Delta t} \sum_i \sum_j \Delta t^2 \frac{x_i}{\Delta t} Q_{ij} \frac{x_j}{\Delta t}, \quad (6.5)$$

where the matrix Q is defined as (order $N-1 \times N-1$)

$$Q = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 \\ & & \cdots & & & & \\ 0 & \cdots & & 0 & -1 & 2 & -1 \\ 0 & \cdots & & 0 & 0 & -1 & 2 \end{pmatrix}. \quad (6.6)$$

(N.B. We have set $x_a = x_0 = 0$, and $x_b = x_N = 0$, in this case. Reasons will become apparent later.) The correspondence between (6.4) and (6.5) is then,

$$y_j = x_j/\Delta t, \quad A_{ij} = m(i\hbar\Delta t)^{-1} Q_{ij}. \quad (6.7)$$

Therefore in formulas (2.18) and (4.5), the eigenvalues $a(t_j)$ of A_{ij} correspond to $m(i\hbar\Delta t)^{-1} q(t_j)$, where $q(t_j)$ are eigenvalues of the matrix Q. For finite N, both $a(t_j)$ and $q(t_j)$ are dependent on the partition of $T = N\Delta t$, the time interval. But as $N \to \infty$, $(\Delta t \to 0)$, $a(t_j)$ should approach the eigenvalues $a(t)$, $0 \leqslant t \leqslant T$, of the continuum matrix, $A(r,s)$, as discussed in paper I. The same limit is expected of $q(t_j)$, even though Q is of the discrete form (6.6). The reason for this can be seen in the marker, the time interval Δt in the quadratic form, $(x_j/\Delta t)$; i.e., the path $x(t)$ is divided into N intervals at points, $x_1, x_2, ..., x_{N-1}$. And the entries of the matrix Q correspond to these x_j. As $N \to \infty$ and $\Delta t \to 0$, $x(t_j)$ becomes the continuous path $x(t)$, and the eigenvalues $q(t_j)$ approach the continuum eigenvalues $q(t)$. To show this numerically, we have evaluated the eigenvalues of the matrix Q of (6.6) for $N-1 = 10$, 20, 40, and 80. The results are presented in Fig. 1 and Table I. It is seen from the graph that the eigenvalues $q(t_j)$ lie between 0 and 4, and as N increases, $q^N(t_j)$ approach a limiting curve. The values of $q(t_j)$ of $N-1 = 40$ are already very close to $q(t_j)$ of $N-1 = 80$, and are hardly distinguishable from each other. We evaluated $q(t_j)$ for even higher values of N. The shape of the curve did not change much for $N > 81$.

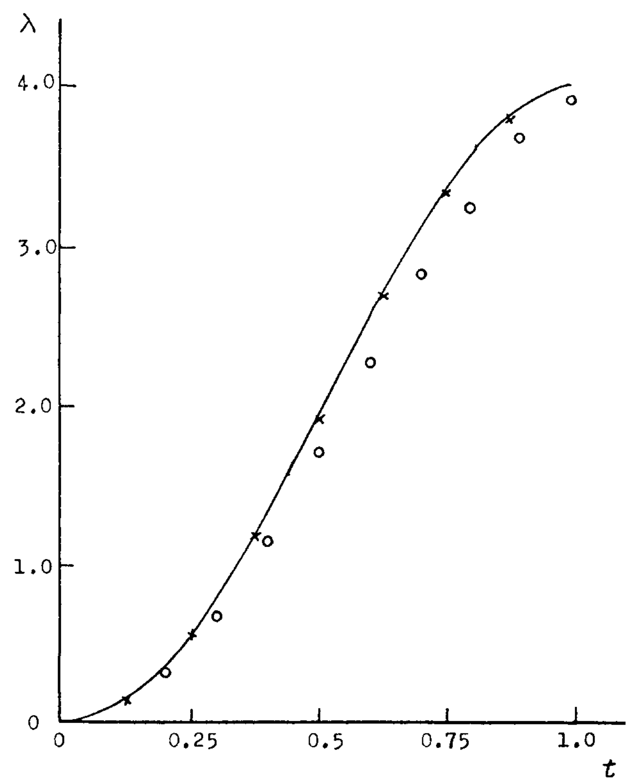

FIG. 1. The eigenvalues of the matrix Q. Circles, o, are for $\Delta t = 0.0909$; Crosses x, are for $\Delta t = 0.0244$. The continuous curve is for $\Delta t = 0.01234$. The unit of time is T.

TABLE I. Eigenvalues of the matrix Q (rank = 80).

0.1502E–02	0.5012E–02	0.1352E–01	0.2402E–01
0.3749E–01	0.5391E–01	0.7328E–01	0.9550E–01
0.1206	0.1486	0.1793	0.2127
0.2489	0.2877	0.3290	0.3729
0.4192	0.4679	0.5189	0.5721
0.6275	0.6850	0.7444	0.8057
0.8688	0.9336	1.000	1.068
1.137	1.208	1.280	1.352
1.426	1.501	1.577	1.653
1.729	1.806	1.884	1.961
2.039	2.116	2.194	2.271
2.347	2.423	2.499	2.574
2.647	2.720	2.792	2.863
2.932	3.000	3.066	3.131
3.194	3.255	3.315	3.372
3.428	3.481	3.532	3.581
3.627	3.671	3.712	3.751
3.787	3.821	3.851	3.879
3.904	3.927	3.946	3.982
3.976	3.986	3.994	3.998

VII. THE FREE PARTICLE IN QUANTUM MECHANICS

In this section and in the following, we shall evaluate the Feynman path integrals for certain known problems in quantum mechanics, e.g., the free particle and the harmonic oscillator. This will serve the purpose of testing the methodology developed so far. It will also shed new insights into some old problems and point out new directions of application of the present method.

The one-dimensional free particle which is at position x_a at time t_a and position x_b at time t_b $(t_a < t_b)$ is described by a propagator $K(b,a)$ given by Feynman,[2]

$$K(b,a) = \int_\Gamma Dx \exp[(i/\hbar) S(b,a)], \quad (7.1)$$

where $S(b,a)$ is the action,

$$S(b,a) = \int_{t_a}^{t_b} dt\, L(x, \dot{x}, t) \quad (7.2)$$

and $L(x,\dot{x},t)$ is the Lagrangian. For a free particle,

$$L(x,\dot{x},t) = \tfrac{1}{2}m\dot{x}(t)^2. \tag{7.3}$$

Note that Γ is the collection of all paths in the function space which pass through x_a at time t_a and x_b at time t_b, $(t_a < t_b)$. To take care of the boundary conditions properly, we follow Feynman[2] in first expanding the path, $x(t) = \bar{x}(t) + y(t)$, around the classical path, $\bar{x}(t)$. The difference function $y(t)$ defined above now has the boundary conditions:

$$y(t_a) = 0, \text{ and } y(t_b) = 0. \tag{7.4}$$

Substitution of $x = \bar{x} + y$ into (7.2) gives

$$S[x] = \int_{t_a}^{t_b} dt \left[L(x,\dot{x},t) + \frac{\partial L}{\partial x}\bigg|_{\bar{x}} y + \frac{\partial L}{\partial \dot{x}}\bigg|_{\bar{x}} \dot{y} + \frac{1}{2}\left(\frac{\partial^2 L}{\partial x^2}\bigg|_{\bar{x}} y^2 + \frac{\partial^2 L}{\partial \dot{x}^2}\bigg|_{\bar{x}} \dot{y}^2 \right) + \cdots \right]. \tag{7.5}$$

The first term on the right-hand side gives the classical action. The first-order terms in y vanish, since the action is stationary with respect to $x(t)$. For the Lagrangian of the free particle (7.3), we have

$$L(x,\dot{x},t) = \tfrac{1}{2}m\dot{\bar{x}}^2 + \tfrac{1}{2}m(2\dot{\bar{x}}\dot{y}) + \tfrac{1}{2}m\dot{y}^2. \tag{7.6}$$

We have then,

$$S[x] = S_{\rm cl} + \int_{t_a}^{t_b} dt\, \tfrac{1}{2}m\dot{y}(t)^2, \tag{7.7}$$

where $S_{\rm cl}$ is the classical action,

$$S_{\rm cl} = \tfrac{1}{2}m(x_b - x_a)^2/(t_b - t_a). \tag{7.8}$$

The propagator is then,

$$K(b,a) = \exp(i\hbar^{-1} S_{\rm cl}) \int_{\Gamma^*} Dy\cdot\exp\left(i\hbar^{-1} \int_{t_a}^{t_b} dt\, \tfrac{1}{2}m\dot{y}(t)^2 \right), \tag{7.9}$$

where Γ^* refers to all paths with boundary conditions $y(t_a) = 0, y(t_b) = 0$. It is the usual practice in evaluating this integral to discretize the kinetic energy term,

$$\int_{t_a}^{t_b} dt\, \tfrac{1}{2}m\dot{y}(t)^2 = \lim_{e\to 0} \frac{1}{2}\frac{m}{e}\sum_{i=1}^{N}(y_i - y_{i-1})^2. \tag{7.10}$$

where

$y_N = y(t_b) = 0$, $y_0 = y(t_a) = 0$, and $e = (t_b - t_a)/N$. This is a quadratic form. We can write,

$$\sum_{i=1}^{N}(y_i - y_{i-1})^2 = \mathbf{y}^T Q \mathbf{y}, \tag{7.11}$$

with

$$\mathbf{y} = (y_1, y_2, \ldots, y_{N-1}) \tag{7.12}$$

and the matrix, as defined before, is,

$$Q = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ & & & \cdots & & & & & \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix} (N-1)\times(N-1), \tag{7.13}$$

i.e., Q is the covariance matrix with value "2" on the diagonal and value "-1" on the off diagonals. We are to find the integral

$$I_* = B\int m_*(dy) \exp\left[-\frac{1}{2}\left(\frac{m}{i\hbar e}\right) \mathbf{y}^T Q \mathbf{y} \right], \tag{7.14}$$

where B is a normalization constant. It is determined by the formula,

$$K(b,a) = \int_{\Gamma_c} Dx_c\, K(b,c)K(c,a) \tag{7.15}$$

as given by Feynman.[2] In our case this is

$$B = m/(2\pi i\hbar e). \tag{7.16}$$

In the continuum case, we are to evaluate the equivalent integral,

$$I_f = B\int m(dy) \exp\left[-\tfrac{1}{2}(m/i\hbar e)\int dr\, ds\, y(r) Q(r,s) y(s) \right]. \tag{7.17}$$

Comparison or (7.17) with (1.3) shows that this is a Gaussian-like integral with covariance matrix $(m/i\hbar e)Q$ and the functional $\phi[y]$ is simply unity. The result is given immediately by formula (2.12) where now $\phi_0 = 1$, and all $K(r\ldots s) = 0$:

$$I_f = B\cdot(2\pi/\det Q)^{1/2}(i\hbar e/m)^{1/2} = (m/2\pi i\hbar e\det Q)^{1/2}. \tag{7.18}$$

To find the determinant of Q, $\det Q$, we note some recursion relations. For the matrix (7.13), if we take the lower right corner and carve off a square matrix, D_n, of order n, $0 \leqslant n \leqslant N-1$, we have

$$D_1 = (2), \quad D_2 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

$$D_3 = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \text{ etc.} \tag{7.19}$$

and the determinants of these matrices obey the recursion relation:

$$\det D_n = 2\det D_{n-1} - \det D_{n-2}, \quad n = 2,3,\ldots,N-1. \tag{7.20}$$

Rearrangement gives

$$(\det D_n - \det D_{n-1}) - (\det D_{n-1} - \det D_{n-2}) = 0. \tag{7.21}$$

In the continuum case,

$$\frac{\partial^2[\det D(t)]}{\partial t^2}=0. \tag{7.22}$$

The appropriate boundary conditions are $\det D(0)=1$, $\partial[\det D(0)]/\partial t=1$. We have the solution,

$$\det D(t)=t+1. \tag{7.23}$$

Therefore $\det D(N-1)=\det Q=(N-1)+1=N$, and

$$I_f=(m/2\pi i\hbar\epsilon N)^{1/2}=[m/(2\pi i\hbar(t_b-t_a))]^{1/2} \tag{7.24}$$

and the propagator $K(b,a)$ is given by

$$K(b,a)=\left(\frac{m}{2\pi i\hbar(t_b-t_a)}\right)^{1/2}\exp(i\hbar^{-1}S_{\rm cl}). \tag{7.25}$$

This is precisely the result given by Feynman.[2]

The above result has been obtained by the Taylor series method. The method of the Laplace transform can be equally well applied. However since in this case, the functional $\phi[y]$ to be integrated is identically unity, the calculations are trivial. To show consistency, we nonetheless carry out the evaluations.

The Laplace transform of $\phi[y]=1$ is

$$\tilde{\phi}[z]=\exp\left[-\int dt\ln z(t)\right]. \tag{7.26}$$

Multiply by $\exp(-\int dt\ln z(t))$ and change to $w(t)=z(t)^2$,

$$\exp\left[-\tfrac{1}{2}\int dt\ln w(t)\right]\tilde{\phi}[z]=\exp\left[-\int dt\ln w(t)\right]. \tag{7.27}$$

The inverse Laplace transform gives back

$$\mathscr{L}^{-1}\left\{\exp\left[-\frac{1}{2}\int dt\ln w(t)\right]\tilde{\phi}\left[\sqrt{w}\,\right]\right\}[x]=1. \tag{7.28}$$

Therefore the functional integral I_f is given by,

$$I_f=B(2\pi i\hbar\epsilon/m\det Q)^{1/2}\mathscr{L}^{-1}\left\{\exp\left[-\frac{1}{2}\int dt\ln w\right]\times\tilde{\phi}\left[\sqrt{w}\,\right]\right\}[x]\bigg|_{x=(2\alpha)}$$
$$=B(2\pi i\hbar\epsilon/m\det Q)^{1/2}=(m/2\pi i\hbar(t_b-t_a))^{1/2}. \tag{7.29}$$

Thus we obtain results identical to (7.18) from the Taylor series method, and the two methods are equivalent.

VIII. HARMONIC OSCILLATOR

That the continuum calculus is applicable to the quantum harmonic oscillator has already been demonstrated in paper I.[4] There we considered integrals of exponential functionals:

$$\psi[y]=\exp\int\mu(dt)\cdot f(y(t)). \tag{8.1}$$

The method presented there[4] is naturally suited to analyze cases such as harmonic oscillators, and is recommended for use over the materials to be given in this section for exponential functionals. The purpose of this section is to demonstrate further the validity of the present method for Gaussian-like integrals. Valid results are obtained. The mathematics, though involved, is quite interesting.

The Lagrangian of the harmonic oscillator is

$$L(x,\dot{x},t)=\tfrac{1}{2}m(\dot{x}^2-\omega^2x^2), \tag{8.2}$$

with the boundary conditions,
$x(t_a)=x_a,\ x(t_b)=x_b,\ t_a<t_b$. The propagator is then

$$K(b,a)=\exp\left\{i\hbar^{-1}\int_{t_a}^{t_b}dt\frac{1}{2}m\,[\dot{x}(t)^2-\omega^2x(t)^2]\right\}. \tag{8.3}$$

We follow Feynman in expanding $x(t)=\bar{x}(t)+y(t)$ around its classical path, $\bar{x}(t)$. Then we obtain, as before,

$$K(b,a)=\exp(i\hbar^{-1}S_{\rm cl})\int_{\Gamma^*}m(dy)\times\exp\left\{-\frac{1}{2}\left(\frac{m}{i\hbar}\right)\int_{t_a}^{t_b}dt\,[\dot{y}(t)^2-w^2y(t)^2]\right\}, \tag{8.4}$$

where Γ^* consists of all paths in A_B with end conditions, $y(t_a)=0$, and $y(t_b)=0$. $S_{\rm cl}$ is the classical action,

$$S_{\rm cl}=(m\omega/2\sin\omega T)\,[(x_a^2-x_b^2)\cos\omega T-2x_ax_b]. \tag{8.5}$$

To evaluate (8.4), we again discretize the integrand. The kinetic energy part is

$$-\frac{1}{2}\left(\frac{m}{i\hbar}\right)\int_{t_a}^{t_b}dt\,\dot{y}(t)^2=-\frac{1}{2}\left(\frac{m}{i\hbar\epsilon}\right)\sum_{i=1}^{N}(y_i-y_{i-1})^2$$
$$=-\frac{1}{2}(m/i\hbar\epsilon)\mathbf{y}^TQ\mathbf{y}, \tag{8.6}$$

where $\mathbf{y}$ and Q are defined as in (7.11) and (7.13). The potential energy part becomes

$$-\left(\frac{im\omega^2}{2\hbar}\right)\int_{t_a}^{t_b}dt\,y(t)^2=-\left(\frac{im\epsilon\omega^2}{2\hbar}\right)\sum_{j=1}^{N}y_j^2$$
$$\equiv-\alpha\mathbf{y}^T\mathbf{y}, \tag{8.7}$$

where

$$\alpha\equiv im\epsilon\omega^2/2\hbar. \tag{8.8}$$

We consider first the weak distribution integral

$$I_*=B\int m_*(dy)\cdot\exp[-\tfrac{1}{2}(m/i\hbar\epsilon)\,\mathbf{y}^TQ\mathbf{y}]\cdot\exp[-\alpha\mathbf{y}^T\mathbf{y}], \tag{8.9}$$

where B is a normalization constant and $B=m/(2\pi i\hbar\epsilon)$. The continuum analog is then obtained as

$$I_f=B\int m(dy)\exp\left[-\tfrac{1}{2}\int dr\,ds\,y(r)Q(r,s)y(s)\right]\times\exp\left[-\alpha\int dt\,y(t)^2\right]. \tag{8.10}$$

Comparison with (1.3) shows that the ϕ functional to be integrated is

$$\phi[y]=\exp\left[-\alpha\int dt\,y(t)^2\right]. \tag{8.11}$$

The covariance matrix is identified as $(m/i\hbar\epsilon)Q$. The kinetic energy part acts as Gaussian measure on a potential energy functional. This behavior is general for Feynman path integrals in physics. Because the nonrelativistic kinetic energy is quadratic, the integral is always the expectation of some quasi-Gaussian process.

To evaluate (8.10), we can use either the Taylor series method or the Laplace transform method. In the former case, we need to find the functional derivatives of all orders for the functional (8.11). This process degenerates very soon into a tedious task (however, it is similar to the cumulant method in statistical mechanics).[15] Instead, we look into the Laplace transform method.

We follow the steps outline in (4.7). First we note that the functional (8.11), is already even, $\phi_E[y] = \phi[y]$. Its Laplace transform can be found from standard transform tables[13,14] as

$$\tilde{\phi}[z] = (\pi/4\alpha)^{1/2} \exp\left[(4\alpha)^{-1}\int dt\, z(t)^2\right]\cdot \mathscr{P}\, dt \gg \operatorname{erfc}[(z(t)^2/4\alpha)^{1/2}], \tag{8.12}$$

where $\mathscr{P}\, dt \gg (.)$ is the p integral of the continuum calculus.[4] For definition and properties, see paper I. The erfc($\cdot$) is the complementary error function. Next multiply by $\exp[-\int dt\, \ln z(t)]$ and change to $w(t) = z(t)^2$,

$$\exp(-\tfrac{1}{2}\int dt\, \ln w)\tilde{\phi}[\sqrt{w}]$$
$$= (\pi/4\alpha)^{1/2}\exp(-\tfrac{1}{2}\int dt\, \ln w)$$
$$\times \exp\left[(4\alpha)^{-1}\int dt\, w(t)\right]\mathscr{P}\, dt \gg \operatorname{erfc}[(w/4\alpha)^{1/2}]. \tag{8.13}$$

From the inverse transform table,[14] we have the relationship,

$$\mathscr{L}^{-1} s^{-\nu}\exp(as)\operatorname{erfc}(\sqrt{as}) = \frac{t^{\nu-1/2}{}_2F_1[1,\tfrac{1}{2};\ -\tfrac{1}{2};\ -t/a]}{\sqrt{\pi a}\,\Gamma(\nu+\tfrac{1}{2})}, \tag{8.14}$$

where ${}_2F_1[a,b;c;x]$ is the hypergeometric function defined as,

$${}_2F_1[a,b;c;x] \equiv \sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{x^n}{n!} \tag{8.15}$$

and

$$(a)_n \equiv a(a+1)\cdots(a+n-1), \text{ etc.} \tag{8.16}$$

Application to our case yields

$$\mathscr{L}^{-1}\left\{\exp(-\tfrac{1}{2}\int dt\, \ln w)\,\tilde{\phi}[\sqrt{w}]\right\}[x] = \mathscr{P}\, dt \gg {}_2F_1[1,\tfrac{1}{2};1;\ -4\alpha x(t)]. \tag{8.17}$$

By using (8.15), we can evaluate ${}_2F_1$,

$$\begin{aligned}{}_2F_1[1,\tfrac{1}{2};1;\ -4\alpha x] &= \sum_{n=0}^{\infty}\frac{(1)_n(\tfrac{1}{2})_n}{(1)_n}\frac{(-4\alpha x)^n}{n!} \\ &= 1 + \left(\frac{1}{2}\right)\frac{(-4\alpha x)}{1!} + \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\frac{(-4\alpha x)^2}{2!} \\ &\quad + \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\left(\frac{5}{2}\right)\frac{(-4\alpha x)^3}{3!} + \cdots \\ &\quad + \frac{1\cdot3\cdot5\cdots(2n-1)}{2^n}\frac{(-4\alpha x)^n}{n!} + \cdots.\end{aligned} \tag{8.18}$$

This is precisely the binomial expansion of the function

$$(1+4\alpha x)^{-1/2} = {}_2F_1[1,\tfrac{1}{2};1;\ -4\alpha x]. \tag{8.19}$$

Therefore, we have

$$\mathscr{L}^{-1}\left\{\exp\left[-\frac{1}{2}\int dt\, \ln w\right]\tilde{\phi}[\sqrt{w}]\right\}[x] = \mathscr{P}\, dt \gg (1+4\alpha x(t))^{-1/2}. \tag{8.20}$$

Substitution into (4.5) gives

$$\begin{aligned}I_f &= B(2\pi i\hbar/m\, \det Q)^{1/2}\mathscr{L}^{-1}\left\{\exp\left(-\frac{1}{2}\int dt\, \ln\omega\right)\times\tilde{\phi}[\sqrt{w}]\right\}[x]\Big|_{x=(2\alpha m/i\hbar)} \\ &= (m/2\pi i\hbar\, \det Q)^{1/2}\,\mathscr{P}\, dt \gg [1 + 2\alpha i\hbar/ma(t)]^{-1/2},\end{aligned} \tag{8.21}$$

where $a(t)$ is the eigenspectrum of the matrix Q. From (8.8) we can further write,

$$I_f = (m/2\pi i\hbar\, \det Q)^{1/2}\,\mathscr{P}\, dt \gg [1 - e^2\omega^2/a(t)]^{-1/2} \tag{8.22}$$

we know already that $\det Q = N$. In Eq. (8.22) we also need the eigenvalues, $a(t)$, of Q. To evaluate the p integral, we have written a computer program. Since computer calculation deals with discrete numbers, the p integral (8.22) was discretized. The details are presented in Sec. IX. Here we give the results.

The p integral in (8.22) was found to be equal to,

$$\mathscr{P}\, dt \gg \left(1 - \frac{e^2\omega^2}{a(t)}\right)^{-1/2} = \left(\frac{\omega T}{\sin\omega T}\right)^{1/2}, \tag{8.23}$$

where $T = t_b - t_a$.

Substitution of (8.23) into (8.22) and then into (8.4) gives

$$K(b,a) = (m\omega/2\pi i\hbar\, \sin\omega T)^{1/2}\exp(i\hbar^{-1}S_{\rm cl}). \tag{8.24}$$

This is the well-known results for quantum harmonic oscillator (see, e.g., Feynman *et al.*[2]).

IX. NUMERICAL METHOD

In this section we report the numerical method used in evaluating the p integral (8.23). There is up to now no known method of obtaining an analytical expression for the eigenvalues, $a(t)$, of the matrix Q (7.13). Thus we cannot used directly the analytical formula of the p integral given in paper I.[4] The evaluation has to be done numerically. Numerical methods, when carried out properly, are equivalent to analytical methods. The results should be equally valid.

First we observe some characteristics. The matrix Q is given in a discrete form (7.13). In order to obtain an approximation to the continuum case, the rank, which is $N-1$, of Q will have to be very large, ideally $N\to\infty$. The limitations will come in from the capacity of the computer and the numerical precision of the working programs seeking the eigenvalues. The computer handles discrete numbers. Thus the p integral must be put in discrete form. A natural way of doing this is to discretize (8.23) consistent with the existing matrix Q.

Secondly, Feynman[2] integrated the path integral using the Fourier series method. The resulting expression corresponding to (8.23) was also an infinite product,

$$\left(\frac{\omega T}{\sin\omega T}\right)^{1/2} = \prod_{n=1}^{\infty}\left(1 - \frac{\omega^2T^2}{n^2\pi^2}\right)^{-1/2}. \tag{9.1}$$

Our expression is

$$\mathscr{P}\, dt \gg \left(1-\frac{e^2\omega^2}{a(t)}\right)^{-1/2} \rightarrow \prod_{j=1}^{N-1}\left(1-\frac{\omega^2T^2}{N^2a_j}\right)^{-1/2}. \tag{9.2}$$

where a_j, $1 \leqslant j \leqslant N-1$, are the eigenvalues of the $(N-1)\times(N-1)$ matrix Q, and we used $e = T/N$, as defined before. Our n product (9.2) is not the same as Feynman's n product even though both give the same limit. In other words, term by term, there is no correspondence between (9.1) and (9.2). They represent different ways that the function $(\omega T/\sin\omega T)^{1/2}$ can be decomposed.

We evaluated the product (9.2) numerically taking $N = 11, 21, 41, 81$, for the values of $T = 0.1, 0.2, 0.3,...,3.0$. Part of the results are presented in Table II. Together we have also calculated the Feynman product (9.1). The eigenvalues of Q at $N-1 = 80$ are given in Table I. For two values of $T = 0.6$ and 0.8, we plot the results in Fig. 2. We see that (9.2) converges to the exact values $(\omega T/\sin\omega T)^{1/2}$ as N becomes large. The contribution to the product (9.2) comes essentially from the small eigenvalues, since for large eigenvalues, $(\omega T/N)^2/a_j$ is essentially zero. For the Feynman product (9.1), the convergence is from below and is well behaved. The rate of convergence is less rapid than (9.2).

We conclude that numerically we have demonstrated that product (9.2) converges to

$$\lim_{N\to\infty}\prod_{j=1}^{N-1}\left[1-\frac{\omega^2T^2}{N^2a_j}\right]^{-1/2} = \left(\frac{\omega T}{\sin\omega T}\right)^{1/2}. \tag{9.3}$$

This result is quite fascinating since it arose from the eigenvalues of the matrix Q. The eigenvalues lie from close to zero to a maximum of just below 4 and form a slightly S-shaped curve. Even when N becomes very large, the maximum values of a_j never exceeds 4. We have thus shown that the present Laplace transform method gives valid results for the harmonic oscillator.

X. HARMONIC OSCILLATOR IN AN EXTERNAL FIELD

The forced harmonic oscillators are of interest in quantum electrodynamics because the electromagnetic field can be represented as a set of forced harmonic oscillators. If the

TABLE II. Convergence of the N-product formula (9.2) in the case of harmonic oscillator to the function $[\omega T/\sin(\omega T)]^{1/2}$.

ωT \ Δt[a]	0.0909	0.0476	0.0244	0.01234	$(\omega T/\sin\omega T)^{1/2}$ exact
0.1	1.0008 (1.00078)	1.0008 (1.0008)	1.0008 (1.0008)	1.0008 (1.0008)[b]	1.0008
0.2	1.0033 (1.0031)	1.0033 (1.0032)	1.0033 (1.0033)	1.0033 (1.0033)	1.0033
0.3	1.0075 (1.0070)	1.0075 (1.0073)	1.0075 (1.0074)	1.0076 (1.0075)	1.0076
0.4	1.0134 (1.0126)	1.0135 (1.0131)	1.0134 (1.0133)	1.0135 (1.0134)	1.0135
0.5	1.0211 (1.0199)	1.0212 (1.0206)	1.0211 (1.0209)	1.0213 (1.02105)	1.0212
0.6	1.0306 (1.0289)	1.0307 (1.0299)	1.0306 (1.0303)	1.0309 (1.0306)	1.0308
0.7	1.0421 (1.0397)	1.0423 (1.0411)	1.0421 (1.0417)	1.0424 (1.04204)	1.0424
0.8	1.0556 (1.0524)	1.0559 (1.0543)	1.0557 (1.0552)	1.0561 (1.0556)	1.0560
0.9	1.0713 (1.0673)	1.0717 (1.0696)	1.0714 (1.0708)	1.0720 (1.0713)	1.0719
1.0	1.0894 (1.0843)	1.0899 (1.0873)	1.0895 (1.0887)	1.0902 (1.0894)	1.0901

[a]Time interval $\Delta t = T/N$, where $T = t_b - t_a$ is taken to be unity.

[b]Values in the parentheses refer to the Feynman product (9.1).

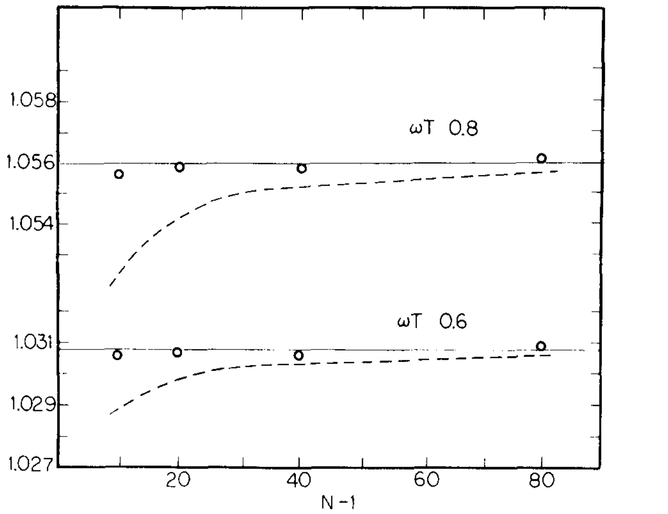

FIG. 2. Convergence of the n product (9.2) to the function $[\omega T/\sin(\omega T)]^{1/2}$. Circles: the n product (9.2). Horizontal lines: the exact values. The upper set of curves is for $\omega T = 0.8$. The lower set of curves is for $\omega T = 0.6$.

external force is given by $f(t)$, it is coupled to the oscillator by

$$L = \tfrac{1}{2}m\dot{x}(t)^2 - \tfrac{1}{2}m\omega^2 x(t)^2 + f(t)x(t). \qquad (10.1)$$

By using the results of the harmonic oscillator obtained previously, it will be a simple matter to obtain the path integral for the propagator,

$$K(x_b, t_b; x_a, t_a) = B\int Dx \exp i\hbar^{-1}\int_{t_a}^{t_b} dt\,(\tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}m\omega^2 x^2 + fx). \qquad (10.2)$$

To extract the classical path, we write the path, $x(t) = \bar{x}(t) + y(t)$, where $\bar{x}(t)$ is the classical path that makes the action $S = \int dt\, L$ stationary; i.e., the first order variation $\delta S = 0$. Substitution of $x(t)$ into the action gives

$$S = \int dt\,(\tfrac{1}{2}m\dot{\bar{x}}^2 - \tfrac{1}{2}m\omega^2\bar{x}^2 + f\bar{x}) + \int dt\,(m\dot{\bar{x}}\dot{y} - m\omega^2\bar{x}y + fy) + \int dt\,(\tfrac{1}{2}m\dot{y}^2 - \tfrac{1}{2}m\omega^2 y^2). \qquad (10.3)$$

The first term on the right-hand side gives the classical action,

$$S_{\rm cl} = m\omega(2\sin\omega T)^{-1}\Big[\cos\omega T(x_b^2 - x_a^2) - 2x_b x_a + 2x_b(m\omega)^{-1}\int_{t_a}^{t_b} ds\, f(s)\sin q + 2x_a(m\omega)^{-1} \times \int_{t_a}^{t_b} dt\, f(t)\sin p - \frac{2}{m^2\omega^2}\int ds\, dt\, f(s)f(t)\sin p\,\sin q\Big], \qquad (10.4)$$

where $p = \omega(t_b - t)$, $q = \omega(s - t_a)$, and $T = (t_b - t_a)$. The second term of (10.3) is zero because it is first order in the perturbation $y(t)$, i.e., $\delta S = 0$. Therefore we have

$$K = B\exp(i\hbar^{-1}S_{\rm cl}) \times \int Dx \exp i\hbar^{-1}\int dt\,(\tfrac{1}{2}m\dot{y}^2 - \tfrac{1}{2}m\omega^2 y^2). \qquad (10.5)$$

This integral is the same as (8.4), the unperturbed harmonic oscillator case. So the answer is given immediately,

$$K = \left(\frac{m\omega}{2\pi i\hbar\sin(\omega T)}\right)^{1/2}\exp(i\hbar^{-1}S_{\rm cl}). \qquad (10.6)$$

The difference of the forced harmonic oscillator from (8.24) lies in the classical action, $S_{\rm cl}$. The normalization factor, which was obtained from the path integral, is the same in both cases.

XI. CHARGED PARTICLE IN AN EXTERNAL MAGNETIC FIELD

The Lagrangian for a particle of charge e and mass m in a constant external magnetic field B, in the z direction, is

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + (eB/2c)(x\dot{y} - y\dot{x}). \qquad (11.1)$$

Let the classical path be $(\bar{x}(t), \bar{y}(t), \bar{z}(t))$. We can extract the classical action, $S_{\rm cl}$, by writing $x = \bar{x} + u$, $y = \bar{y} + v$, $z = \bar{z} + w$, where u, v, w are perturbations over the classical path. The action is then

$$S = \int dt\,(\tfrac{1}{2}m(\dot{\bar{x}}^2 + \dot{\bar{y}}^2 + \dot{\bar{z}}^2) + \tfrac{1}{2}m\omega(\bar{x}\dot{\bar{y}} - \bar{y}\dot{\bar{x}})) + \int dt\,(m(\dot{\bar{x}}\dot{u} + \dot{\bar{y}}\dot{v} + \dot{\bar{z}}\dot{w}) + \tfrac{1}{2}m\omega(u\dot{\bar{y}} + \bar{x}\dot{v} - v\dot{\bar{x}} - \bar{y}\dot{u})) + \int dt\,(\tfrac{1}{2}m(\dot{u}^2 + \dot{v}^2 + \dot{w}^2) + \tfrac{1}{2}m\omega(u\dot{v} - v\dot{u})). \qquad (11.2)$$

The first term on the right-hand side is the classical action,

$$S_{\rm cl} = \tfrac{1}{2}m\omega\,\{(z_b - z_a)^2/T + \tfrac{1}{2}\omega\cot(\tfrac{1}{2}\omega T) \times [(x_b - x_a)^2 + (y_b - y_a)^2] + \omega(x_a y_b - x_b y_a)\}, \qquad (11.3)$$

where $\omega = eB/(mc)$. The second term is zero, because first order variations over the classical path, $\delta S = 0$. Therefore we have as the propagator,

$$K = B'\exp(i\hbar^{-1}S_{\rm cl})\int DuDvDw \times \exp\left[i\hbar^{-1}\int_{t_a}^{t_b} dt\,(\tfrac{1}{2}m(\dot{u}^2 + \dot{v}^2 + \dot{w}^2) + \tfrac{1}{2}m\omega(u\dot{v} - v\dot{u}))\right]. \qquad (11.4)$$

We note that this integral is in three-dimensional space. Comparison with the Gaussian integral of (1.3) shows that the covariance matrices are

$$\lim \frac{m}{i\hbar\Delta t}(\mathbf{u}^T Q\mathbf{u} + \mathbf{v}^T Q\mathbf{v} + \mathbf{w}^T Q\mathbf{w}), \qquad (11.5)$$

where $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are the vectors $(u_1, u_2, \ldots, u_{N-1})$, $(v_1, \ldots, v_{N-1})$, and $(w_1, \ldots, w_{N-1})$, and Q is the same for isometric motion, [see (6.6)], in all three directions. The functional whose Gaussian expectation is to be evaluated is

$$\phi[u,v,w] = \exp\int dt\,\tfrac{1}{2}i\hbar^{-1}m\omega(u\dot{v} - v\dot{u}) = \lim\exp(\tfrac{1}{2}i\hbar^{-1}m\omega\,\mathbf{u}^T R\mathbf{v}), \qquad (11.6)$$

where

$$R = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ & \cdots & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 0 \end{pmatrix}, \tag{11.7}$$

i.e., a matrix with 0 on the diagonal, $+1$ on the upper off-diagonal, and -1 on the lower off-diagonal. This matrix is obtained when the velocities $\dot{u}$, and $\dot{v}$, are discretized as $[u((i+1)\Delta t) - u(i\Delta t)]/\Delta t$, and $[v((i+1)\Delta t) - v(i\Delta t)]/\Delta t$, respectively (see section VI). The limit in (11.6) was taken as $\Delta t \to 0$ and $N \to \infty$. The functional ϕ is independent of $w(t)$. The functional integration over Dw then gives a factor of $[m/(2\pi i\hbar T)]^{1/2}$. We have

$$K = B''[m/(2\pi i\hbar T)]^{1/2} \exp(i\hbar^{-1} S_{\rm cl}) \int m_*(dv) m_*(dw)$$
$$\times \exp[-\tfrac{1}{2}[m/(\Delta t i\hbar)](\mathbf{u}^T Q \mathbf{u} + \mathbf{v}^T Q \mathbf{v})]$$
$$\times \exp[-\tfrac{1}{2}(m\omega/i\hbar)\, \mathbf{u}^T R \mathbf{v}]. \tag{11.8}$$

We note that R is an antisymmetric matrix, which can be diagonalized by some regular matrix, C,; i.e., $CRC^{-1} = \Lambda$, where Λ is a diagonal matrix with the eigenvalue λ_i displayed on the diagonal. For (11.7), all the eigenvalues are imaginary. Table III gives the spectrum of R. Now if we define the new integration variables, $\mathbf{u} = C^T\mathbf{a}$, with Jacobian of transformation $J(\partial u/\partial a) = \det C$, and $\mathbf{v} = C^{-1}\mathbf{b}$, with Jacobian of transformation $J(\partial v/\partial b) = \det C^{-1} = 1/\det C$, (11.8) will be

$$K = B''\left(\frac{m}{2\pi i\hbar T}\right)^{1/2} \exp\left(\frac{i}{\hbar} S_{\rm cl}\right) \int m_*(da) m_*(db)$$
$$\times \exp\left\{-\frac{1}{2}\frac{m}{i\hbar\Delta t}[\mathbf{a}^T CQC^T \mathbf{a} + \mathbf{b}^T (C^{-1})^T QC^{-1}\mathbf{b}]\right\}$$
$$\times \exp\left(-\frac{1}{2}\frac{m\omega}{i\hbar}\mathbf{a}^T \Lambda \mathbf{b}\right). \tag{11.9}$$

There is no change in the differential volume due to the cancellation of the Jacobians. We diagonalized the matrix in the ϕ-functional part in order to facilitate the application of the Laplace transform,

$$\mathscr{L}\phi = \int m_*(da) m_*(db) \cdot \exp(-(a,r))$$
$$\times \exp[-(b,s)] \cdot \phi[a,b], \tag{11.10}$$

according to (3.4). The procedure of (4.7) is then activated. Namely, we first seek out the even functional, $\phi_E[a,b]$. This is relatively easy since Λ is already diagonalized. We then transform ϕ_E into $(r-s)$ space, and multiply by $\exp[-\int dt \ln r(t)]$ and $\exp[-\int dt \ln s(t)]$. Changing the variables, $r^2 = r'$, and $s^2 = s'$, and inverse-transforming back to the $(p-q)$ space give the path integral after replacing the variables, p_i, by the eigenvalues, in the form of $(2\mu_i)^{-1}$, of CQC^T, and q_i by $(2v_i)^{-1}$ of $(C^{-1})^T QC^{-1}$, and multiplication of the normalization factor $(2\pi/\det Q)$. The following shows this construction step by step.

The even functional, ϕ_E, of ϕ, in (11.9) is simply the hypergeometric functional,

$$\phi_E[a,b] = \mathscr{P}\, dt \gg {}_1F_2\, [1;1,\tfrac{1}{2};\tfrac{1}{4}\alpha^2 a^2 b^2], \tag{11.11}$$

where α is the constant $\frac{1}{2}[m\omega\lambda\,(i\hbar)^{-1}]$. Its Laplace transform can be obtained directly from the standard mathematical table,[14] and is given by

$$\mathscr{L}\phi_E = \left[\exp\left(-\int dt \ln r(t)\right)\exp\left(-\int dt \ln s(t)\right)\right]\mathscr{P}\, dt$$
$$\gg {}_4F_1\left[\tfrac{1}{2},1,1,1;1;\frac{4\alpha^2(t)}{(r(t)s(t))^2}\right]. \tag{11.12}$$

Multiplication by $\exp[-\int dt \ln(rs)]$, and transformation of variables, $r' = r^2$, and $s' = s^2$ gives

$$\mathscr{L}\phi_E[r',s'] = \left(\exp\left[-\int dt \ln(r's')\right]\right)\mathscr{P}\, dt$$
$$\gg {}_4F_1\left[\tfrac{1}{2},1,1,1;1;\frac{4\alpha^2(t)}{r'(t)s'(t)}\right]. \tag{11.13}$$

The inverse transform can be easily obtained again from the mathematical table:[14]

$$\mathscr{L}^{-1}(\mathscr{L}\phi_E[r',s'])[p,q] = \mathscr{P}\, dt$$
$$\gg {}_2F_1\,[\tfrac{1}{2},1;1;4\alpha^2 p(t)q(t)]. \tag{11.14}$$

TABLE III. Eigenvalues of the matrix R (rank = 80).

real	imaginary	real	imaginary
0.0	1.998	0.0	−1.998
0.0	1.993	0.0	−1.993
0.0	1.986	0.0	−1.986
0.0	1.975	0.0	−1.975
0.0	1.962	0.0	−1.962
0.0	1.945	0.0	−1.945
0.0	1.926	0.0	−1.926
0.0	1.904	0.0	−1.904
0.0	1.879	0.0	−1.879
0.0	1.851	0.0	−1.851
0.0	1.820	0.0	−1.820
0.0	1.787	0.0	−1.787
0.0	1.751	0.0	−1.751
0.0	1.712	0.0	−1.712
0.0	1.671	0.0	−1.671
0.0	1.627	0.0	−1.627
0.0	1.580	0.0	−1.580
0.0	1.532	0.0	−1.532
0.0	1.481	0.0	−1.481
0.0	1.427	0.0	−1.427
0.0	1.372	0.0	−1.372
0.0	1.315	0.0	−1.315
0.0	1.255	0.0	−1.255
0.0	1.194	0.0	−1.194
0.0	0.3877*E*–01	0.0	−0.3877*E*–01
0.0	0.1163	0.0	−0.1163
0.0	0.1936	0.0	−0.1936
0.0	0.2706	0.0	−0.2706
0.0	0.3472	0.0	−0.3472
0.0	0.4234	0.0	−0.4234
0.0	0.4988	0.0	−0.4988
0.0	0.5735	0.0	−0.5735
0.0	0.6474	0.0	−0.6474
0.0	1.31	0.0	−1.131
0.0	0.7203	0.0	−0.7203
0.0	0.7921	0.0	−0.7921
0.0	0.8627	0.0	−0.8627
0.0	1.066	0.0	−1.066
0.0	0.9320	0.0	−0.9320
0.0	1.000	0.0	−1.000

TABLE IV. Eigenvalues of the matrix $Q^{-1}RQ^{-1}R$ (rank = 72).

real	imaginary	real	imaginary
−513.0	0.0	−513.0	0.0
−132.0	0.0	−132.7	0.0
−59.00	0.0	−32.98	0.0
−59.00	0.0	−32.98	0.0
−20.89	0.0	−14.32	0.0
−20.89	0.0	−14.32	0.0
−10.35	0.0	−10.35	0.0
−7.771	0.0	−6.005	0.0
−7.772	0.0	−6.406	0.0
−4.743	0.0	−4.743	0.0
−3.809	0.0	−3.800	0.0
−3.100	0.0	−3.100	0.0
−2.549	0.0	−2.549	0.0
−2.113	0.0	−1.762	0.0
−2.113	0.0	−1.762	0.0
−1.976	0.0	−1.240	0.0
−1.976	0.0	−1.240	0.0
−1.043	0.0	−1.044	0.0
−0.8784	0.0	−0.8785	0.0
−0.7387	0.0	−0.5198	0.0
−0.7388	0.0	−0.5199	0.0
−0.5182	0.0	−0.4310	0.0
−0.5182	0.0	−0.4310	0.0
−0.3559	0.0	−0.2912	0.0
−0.3559	0.0	−0.2912	0.0
−0.2355	0.0	−0.1877	0.0
−0.2355	0.0	−0.1877	0.0
−0.1467	0.0	−0.1118	0.0
−0.1467	0.0	−0.1118	0.0
−0.8249*E*–01	0.0	−0.5817*E*–01	0.0
−0.8250*E*–01	0.0	−0.5817*E*–01	0.0
−0.3845*E*–01	0.0	−0.3845*E*–01	0.0
−0.2302*E*–01	0.0	−0.2303*E*–01	0.0
−0.4567*E*–03	0.0	−0.4627*E*–01	0.0
−0.4170*E*–02	0.0	−0.1166*E*–01	0.0
−0.1166–01	0.0	−0.4178*E*–02	0.0

This whole procedure is reminiscent of the case of the harmonic oscillator cited in Sec. VIII. The hypergeometric function ${}_2F_1\,[\frac{1}{2},1;1;\,4\alpha^2 pq]$ is a representation of the function $(1-4\alpha^2pq)^{-1/2}$. Substitution of the eigenvalues $\mu(t)$ and $\nu(t)$ into (11.14) gives

$$I_f=\left(\frac{mN/2\pi}{\det Q\, i\hbar T}\right)^{3/2}\mathscr{P}\,dt \gg\left(1-\left(\frac{\omega T}{2N}\right)^2\mu(t)^{-1}\lambda(t)\nu(t)^{-1}\lambda(t)\right)^{-1/2}\cdot(I_{\rm cl}). \tag{11.15}$$

The result is similar to Eq. (8.22), except (11.15) should be treated in matrix form. $\mu(t)^{-1}\lambda(t)\nu(t)^{-1}\lambda(t)$ is the eigenvalue of the matrix product,

$$(CQC^T)^{-1}\Lambda(C^{-T}QC^{-1})^{-1}\Lambda = C^{-T}Q^{-1}RQ^{-1}R^TC^T, \tag{11.16}$$

where we have used the diagonalized matrix relation, $CRC^{-1}=\Lambda$. The eigenvalues of (11.16) are the same as $Q^{-1}RQ^{-1}R^T$, from a well-known spectral theorem of matrices. Since R is antisymmetric [see (11.7)], $R^T=-R$. Therefore we need only to find the eigenvalues of the matrix product, $-Q^{-1}RQ^{-1}R$, and substitute for the eigenvalue term of the p-integral (11.15). [We note the use of the sym-

TABLE V. Convergence of the N product (11.17) in the case of charged particle to the function $\frac{1}{2}\omega T/\sin(\frac{1}{2}\omega T)$.

ωT \ Δt[a]	0.11111	0.05882	0.1538	0.0137	$\frac{1}{2}\omega T/\sin(\frac{1}{2}\omega T)$ exact
0.2	1.0012	1.0014	1.00155	1.0015	1.0016
0.4	1.0046	1.0056	1.0062	1.0062	1.0066
0.6	1.010	1.013	1.0141	1.014	1.0152
0.8	1.019	1.023	1.025	1.026	1.0271
1.0	1.03	1.036	1.04	1.041	1.043
1.2	1.043	1.052	1.058	1.06	1.063
1.6	1.08	1.1	1.11	1.11	1.115
2.0	1.13	1.155	1.17	1.173	1.188

[a]The time interval, $\Delta t = T/N$, and $T=t_b-t_a$ is set to unity.

bols: C^{-T} is the transpose of the inverse of the matrix C, and I_{cl} is the classical part of the path integral in (11.4)].

It remains to be shown that the p integral in (11.15) conforms to the answer, $\frac{1}{2}\omega T/\sin(\frac{1}{2}\omega T)$, i.e.,

$$\mathscr{P}\ dt \succ \left(1-\left(\frac{\omega T}{2N}\right)^2 \gamma(t)\right)^{-1/2} = \frac{\frac{1}{2}\omega T}{\sin(\frac{1}{2}\omega T)}, \qquad (11.17)$$

where $\gamma(t)$ is the eigenvalue of the matrix, $-Q^{-1}RQ^{-1}R$. Table IV lists the eigenvalues γ for a discrete matrix of rank 72. (This corresponds to a time interval of 0.0137 T, where $T = T_b - T_a$.) It is interesting to note that the eigenvalues appear in pairs. This fact, upon closer scrutinization, is responsible for the result being the square of $(\frac{1}{2}\omega T/\sin(\frac{1}{2}\omega T))^{1/2}$ which was obtained in the one-dimensional case of the harmonic oscillator [see Eq. (8.23)]. The antisymmetrical nature (the exterior product) of the movement in a magnetic field is effectively two dimensional, and thus we have $\frac{1}{2}\omega T/\sin(\frac{1}{2}\omega T)$. Extension to a suitable three-dimensional case could have given a $\frac{3}{2}$ power.

To calculate the p integral, we again used numerical quadrature on a computer. The results are presented in Table V. We see as the time interval, $\Delta t = T/N$, is reduced, the p integral (11.17) approaches the function $\frac{1}{2}\omega T/\sin(\frac{1}{2}\omega T)$ asymptotically for various values of the frequency ωT. The p-integral as calculated in the numerical program was sensitive to the round-off errors. Therefore double precision arithmetic or high precision was required. The 4 bytes precision in the search of the eigenvalues could lead to imprecision in the small eigenvalues. Thus we urge the use of a high-accuracy eigenvalue search subroutine.

The final result for the propagator kernel is

$$K(b,a) = \exp(i\hbar^{-1}S_{cl})\left(\frac{m}{2\pi i\hbar T}\right)^{3/2}\left(\frac{\frac{1}{2}\omega T}{\sin(\frac{1}{2}\omega T)}\right), \qquad (11.18)$$

which checks with the known correct answer.[2]

XII. CONCLUDING REMARKS

In this paper we derived two new methods for the calculation of Gaussian path integrals. The first method was based on the consideration of polynomial functionals, and the resulting formula, (2.12), is applicable to functionals expressible in Taylor series, i.e., the class C^∞ in function space, A_B. The second method was derived by using functional Laplace transform, and the procedure of integration (4.7) is applicable to all Laplace transformable functionals in A_B. The latter class is certainly wider than the class C^∞ (since all polynomial functionals with well-behaved kernel are Laplace transformable, while not vice versa: e.g., the Dirac delta functional has no Taylor series representation.)

The proofs of these methods are given in Secs. II and IV. They constitute the mathematical basis for the two methods. We underscore here that the present study is a first step in the development of the Laplace transform method. Further applications and comparison with other integration methods shall come in the future. However, for the sake of illustration we looked at four physical applications, the quantum mechanics of a free particle, the harmonic oscilator, the oscillator in an external field, and the charged particle in a magnetic field. The mathematics of the harmonic oscillator and the charged particle is of particular interest, not only in the correct answers obtained but also in the unexpected *tour de force* of the mathematics: the identification of the hypergeometric function with a simple binomial, the pairing of the eigenvalues of the matrix product, $-Q^{-1}RQ^{-1}R$, to cite just a few points. Since the Laplace transform is a general method, the second method shall find applications in other interesting physical problems that can be formulated in terms of a Gaussian path integral. Under study are the generating functional of the Bogoliubov equation.[16,17] where the interaction potential acts as a Gaussian measure, and the kernel of the characteristic functionals in Hopf's turbulence theory, where the quadratic form is provided by the two-point correlation tensor.[18] We shall report the results in due course.

[1] R.P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).
[2] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[3] J.T. Devreese, Ed., *Proceedings of the NATO Advanced Study Institute on Path Integrals and Their Applications* (Plenum, New York, 1977); W. Martin and I. Segal, Eds., *Analysis in Function Space* (MIT Press, Cambridge MA, 1964); A.M. Arthurs, Ed., *Functional Integral and Its Applications* (Clarendon, Oxford, 1975).
[4] L.L. Lee, J. Math. Phys. **17**, 1988 (1976).
[5] L.L. Lee, J. Math. Phys. **18**, 1752 (1977).
[6] L.L. Lee, J. Math. Phys. **19**, 2055 (1978).
[7] K.O. Friedrichs, *Mathematical Aspects of the Quantum Theory of Fluids* (Interscience, New York, 1953).
[8] K.O. Friedrichs *et al.*, *Integration of Functionals*, Seminar (New York University, New York, 1957).
[9] W. Kollmann, J. Math. Phys. **16**, 2197 (1975).
[10] V. Volterra, *Theory of Functionals and of Integral and Integrodifferential Equations* (Dover, New York, 1959).
[11] A.V. Skorohod, *Integration in Hilbert Space* (Springer, New York, 1974).
[12] A.C. Zaanen, *Integration* (North-Holland, Amsterdam, 1974).
[13] F. Oberhettinger and L. Badii, *Tables of Laplace Transforms* (Springer, New York, 1972).
[14] G.E. Roberts, and H. Kaufman, *Table of Laplace Transforms* (Saunders, Philadelphia, 1966).
[15] S.A. Rice and P. Gray, *The Statistical Mechanics of Simple Liquids* (Wiley, New York, 1965).
[16] C. Blomberg, Acta Polytech. Scand. Phys. 49 (1967).
[17] N.N. Bogoliubov, Zh. Fiz. (USSR) **10**, 256 (1946).
[18] G. Rosen, J. Math. Phys. **12**, 812 (1971).

Systematic improvement of Hall–Post–Stenschke lower bounds to eigenvalues in the few-body problem

Robert Nyden Hill
Physics Department, University of Delaware, Newark, Delaware 19711

(Received 6 August 1979; accepted for publication 14 December 1979)

A method for systematically improving Hall–Post–Stenschke lower bounds to the bound state eigenvalues of three-body Schrödinger equations is given. The improved bounds are obtained by solving coupled one variable integral equations; the bounds get better as the number of coupled equations is increased. The method generates explicit wave functions which can be used to obtain complementary upper bounds via the Rayleigh–Ritz variational method. Either identical or nonidentical particles can be handled. The method is illustrated by calculations for three identical particles bound by Hooke's law forces. A brief discussion of extensions to more than three particles is given.

I. INTRODUCTION AND SUMMARY

The quantum mechanical N-body problem can, in general, not be solved exactly for $N > 2$. Approximate methods must then be used. The present paper is concerned with the approximate computation of bound state energy eigenvalues in the three-body problem. More specifically, it describes and illustrates a method for computing improvable lower bounds to such eigenvalues. The method is expected to be most useful in nuclear physics and in molecular physics.

The most desirable approxmation methods are those for which error estimates can be obtained; one would like these error estimates to be both rigorous and realistic. The most popular methods for the computation of bound state eigenvalues in the three-body problem have been (1) the numerical solution to Faddeev's equations[1] and (2) the Rayleigh–Ritz variational method.[2] Both statistical error estimates and rigorous error bounds are in principle available for the numerical methods used on Faddeev's equations;[3] unfortunately, the statistical estimates are not rigorous while the rigorous bounds are unrealistically large. In practice, neither the statistical error estimates nor the rigorous error bounds are usually calculated. The Rayleigh–Ritz variational method gives rigorous upper bounds (if the errors made in its numerical implementation are rigorously bounded), but gives no information about the difference between these upper bounds and the true eigenvalues. Complementary lower bounds are needed to complete the picture. The present paper will present a new method for obtaining such complementary bounds.

There is an enormous literature on bounds to eigenvalues. The author has found the books of S.H. Gould,[4] of A. Weinstein and W. Stenger,[5] and of H.F. Weinberger[6] particularly helpful. In atomic physics, impressive results have been obtained by Bazley:[7] rigorous lower bounds to the ground state energy of helium which agree to five significant figures with Rayleigh–Ritz upper bounds. Bazley's method and its extensions and generalization[8] depend on the fact that, in atomic physics, exactly solvable "base problems" can be obtained by deleting the Coulomb repulsion between electrons. Such base problems are not available in nuclear and molecular physics, where typical potentials have the form shown in Fig. 1. Thus, an alternative to Bazley's approach is needed if good lower bounds are to be obtained for nuclear and molecular problems.

Hall and Post[9,10] and, independently, Stenschke[11] have given an unimprovable lower bound on the lowest energy eigenvalue of the identical particle N-body Schrödinger equation. Better bounds were devised for the N-fermion problem by Hall[12] and by Carr and Post.[13] Hall[14] obtained a better bound for the mixed symmetry representations which occur in the three-body problem of nuclear physics, and showed[15] how lower bounds to excited states could be obtained. Generalizations have been given by Calogero and Marchioro[16] and by Savchenko.[17] Post's original bound[9] has been tested[18] for N spinless bosons bound by attractive Coulomb forces; in combination with a simple Rayleigh–Ritz

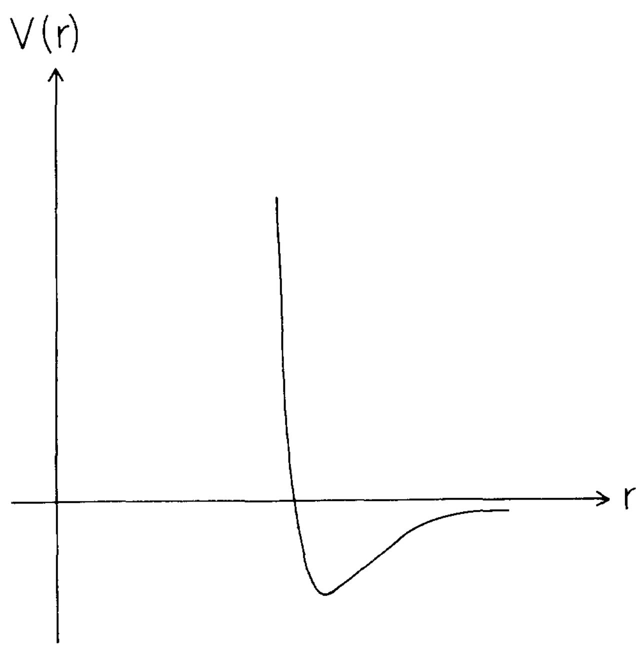

FIG. 1. A typical potential in nuclear or molecular physics.

 0022-2488/80/051070-13$01.00

upper bound, results accurate to $\pm$ 8% were obtained. Tests of the Hall–Post–Stenschke (HPS) bounds for N particles bound via either square wells, exponential wells, Gauss wells, or Hulthen wells have been carried out by Hall and Post[10] and by Hall.[19] A comparison of the HPS lower bound with Rayleigh–Ritz upper bounds and with bound state energies obtained via numerical solution of the Faddeev equations has been conducted by Humberston, Hall, and Osborne.[20] The usefulness of HPS lower bounds for the trinucleon has been explored by Brady, Harms, Laroze, and Levinger,[21] who concluded that "the Hall–Post bounds work moderately well for central potentials, which may include soft cores or weak tensor forces; but the bound is far from the true energy when hard cores or strong tensor forces are included." The HPS bound has been used in molecular problems by Stenschke,[11] by Bruch and McGee,[2] and by Lim and Zuniga,[2] who also found that HPS lower bounds are poor for interactions with hard cores.

None of the papers discussed in the preceding paragraph, which derive and/or apply lower bounds of HPS type, contains an algorithm for pushing the lower bound closer and closer to the true eigenvalue. The present paper provides such an algorithm for the three-body problem, and illustrates it with model calculations for particles bound by Hooke's law forces. Implementation of this algorithm requires the solution of coupled one-variable integral equations; the lower bound can be improved by increasing the number of coupled equations. The lower bounds obtained are of interest not only in their own right, but also as a step in the implementation of a strategy such as Bazley[8] employed on helium: The algorithm can be used to construct the lower bound to the first excited state which is needed for the Temple[22] lower bound to the ground state energy. Several authors[23] have applied the Temple formula to the three-body problem with the lower bound to the first excited state provided by the unproven assumption that there is only one one bound state. Lower bounding methods are needed to either prove such an assumption or replace it with a rigorous lower bound to the first excited state. Section II establishes notation and constructs the internal Hamiltonian in a convenient form. Section III shows how to obtain the one-variable equations whose eigenvalues give lower bounds to the eigenvalues of the original internal Hamiltonian. Section IV illustrates the method by calculating lower bounds to the eigenvalues of some exactly solvable problems of three particles bound by Hooke's law forces. Section V discusses a difficulty which can arise for some potentials, and discusses the extension to more than three particles.

II. NOTATION. THE THREE-BODY HAMILTONIAN

The three-body Schrödinger Hamiltonian is assumed to have the form

$$H = \sum_{i=1}^{3} p_i^2/(2m_i) + V_1(\mathbf{r}_2 - \mathbf{r}_3) + V_2(\mathbf{r}_3 - \mathbf{r}_1) + V_3(\mathbf{r}_1 - \mathbf{r}_2)\,, \tag{2.1}$$

where $\mathbf{r}_i,\mathbf{p}_i$ are individual particle coordinates and momenta and the m_i are particle masses. It is convenient to introduce the normalized Jacobi coordinates

$$\mathbf{R} = (m_1\mathbf{r}_1 + m_2\mathbf{r}_2 + m_3\mathbf{r}_3)/(m_1 + m_2 + m_3), \tag{2.2}$$

$$\mathbf{r}_{i,j} = \mathbf{r}_i - \mathbf{r}_j\,, \tag{2.3}$$

$$\boldsymbol{\rho}_k = (2/\sqrt{3})[\mathbf{r}_k - (m_i\mathbf{r}_i + m_j\mathbf{r}_j)/(m_i + m_j)]\,, \tag{2.4}$$

where $i \neq j \neq k \neq i$. The normalizing factor $2/\sqrt{3}$ in $\boldsymbol{\rho}_k$ is included so that permutations of particle labels will be equivalent to rotations in the space of the internal coordinates $\mathbf{r}_{i,j},\boldsymbol{\rho}_k$ when the masses are identical. The corresponding conjugate momenta are

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3, \tag{2.5}$$

$$\mathbf{p}_{i,j} = (m_j\mathbf{p}_i - m_i\mathbf{p}_j)/(m_i + m_j)\,, \tag{2.6}$$

$$\boldsymbol{\Pi}_k = (\sqrt{3}/2)[(m_i + m_j)\mathbf{p}_k - m_k(\mathbf{p}_i + \mathbf{p}_j)]/(m_1 + m_2 + m_3)\,. \tag{2.7}$$

The center of mass can be separated off to obtain

$$H = P^2/[2(m_1 + m_2 + m_3)] + H_{\text{int}}\,,$$

where the internal Hamiltonian H_{int} can be written in the form

$$\begin{aligned} H_{\text{int}} = {} & (m_1 + m_2)p_{1,2}^2/(2m_1m_2) + 2(m_1 + m_2 + m_3) \\ & \times \Pi_3^2/[3(m_1 + m_2)m_3] \\ & + V_1\{-(\sqrt{3}\boldsymbol{\rho}_3/2) - [m_1\mathbf{r}_{1,2}/(m_1 + m_2)]\} \\ & + V_2\{(\sqrt{3}\boldsymbol{\rho}_3/2) - [m_2\mathbf{r}_{1,2}/(m_1 + m_2)]\} \\ & + V_3(\mathbf{r}_{1,2})\,. \end{aligned} \tag{2.8}$$

Choosing $\mathbf{r}_{1,2}$ and $\boldsymbol{\rho}_3$ as internal coordinates results in the form (2.8) for H_{int}; the form appropriate to other choices may be obtained from Eq. (2.8) by cyclic permutation of the indices (1,2,3). An alternative form which will be needed later is

$$H_{\text{int}} = H_1(\mathbf{r}_{2,3},\mathbf{p}_{2,3}) + H_2(\mathbf{r}_{3,1},\mathbf{p}_{3,1}) + H_3(\mathbf{r}_{1,2},\mathbf{p}_{1,2})\,, \tag{2.9}$$

where

$$H_3(\mathbf{r}_{1,2},\mathbf{p}_{1,2}) = (m_1 + m_2)^2p_{1,2}^2/[2m_1m_2(m_1 + m_2 + m_3)] + V_3(\mathbf{r}_{1,2})\,. \tag{2.10}$$

H_1 and H_2 can be obtained from H_1 by cyclic permutation of the indices (1,2,3).

III. LOWER BOUNDS

A. The basic theorem

The basic tool to be used is a comparison theorem[24] well-known among mathematicians who work on eigenvalue problems:

Theorem 1: Let $H^{(1)}$ and $H^{(2)}$ be two essentially self-adjoint (Hermitian) Hamiltonians whose discrete eigenvalues below the bottom of the essential spectrum (i.e., the continuous spectrum plus limit points of the discrete spectrum and discrete eigenvalues of infinite multiplicity) can be characterized by the familiar variational principle $E = \min\langle\psi|H|\psi\rangle|\langle\psi|\psi\rangle$, with the minimization for excited states carried out subject to the constraint that $|\psi\rangle$ be orthogonal to preceding eigenvectors. Denote the ordered eigenvalues of $H^{(i)}$ by $E_1^{(i)} \leqslant E_2^{(i)} \leqslant \cdots \leqslant E_n^{(i)} \leqslant \cdots \leqslant E_{\text{ess}}^{(i)}$, where $E_{\text{ess}}^{(i)}$ is the energy at which the essential spectrum (if any) begins. Assume $\langle\psi|H^{(1)}|\psi\rangle$ is defined for all vectors $|\psi\rangle$ for which $\langle\psi|H^{(2)}|\psi\rangle$ is defined. Then if $\langle\psi|H^{(1)}|\psi\rangle\leqslant\langle\psi|H^{(2)}|\psi\rangle$ holds for all admissible state vectors $|\psi\rangle$, $E_n^{(1)}\leqslant E_n^{(2)}$ holds for all n, and $E_{\text{ess}}^{(1)}\leqslant E_{\text{ess}}^{(2)}$.

The result $E_1^{(1)}\leqslant E_1^{(2)}$ for the ground state energy follows immediately from the "familiar variational principle." Proofs of the result for the excited states are usually based on one of the minimax characterizations[25] of eigenvalues. In practical applications of theorem 1 to the computation of lower bounds, $H^{(2)}$ is the original Hamiltonian, while $H^{(1)}$ is something more tractable. The results of the present paper will be obtained by letting $H^{(2)}$ be the internal Hamiltonian H_{int} while $H^{(1)}$ is something for which the Schrödinger equation is reducible to one-particle Schrödinger equations.

Model potentials used in nuclear physics and in molecular physics typically have a short range attraction [the Yukawa $r^{-1}\exp(-\mu r)$ tail in nuclear physics, or the r^{-6} van der Waals tail in molecular physics], become repulsive at sufficiently short distances (a hard core or an r^{-12} repulsion as in the Lennard-Jones potential are possibilities), and may, in nuclear physics, also include a long range Coulomb repulsion. Most rigorous treatments of the N-body problem (for $N\geqslant 3$) known to the author[26] place conditions on the potentials that exclude repulsions which are too singular. However, a recent paper of McKean[27] proves the essential self-adjointness of the Hamiltonian for singular potentials which are nonnegative. The conditions placed on the potential by McKean permit extending his result to a potential which is bounded below by shifting the zero of energy; this extension does not include hard cores, but is sufficient to include most other potentials of interest in nuclear and molecular physics. Since essential self-adjointness implies, via the spectral theorem, that bound states can be characterized by the familiar variational principle in the hypothesis of Theorem 1, the application of Theorem 1 to the model Hamiltonians of nuclear and molecular physics is justified if the potentials do not contain hard cores. No attempt will be made here to supply the justification for the hard core case.

B. Lower bounds via truncation

In what follows it will be assumed that the lower part of the spectrum of the two-particle Hamiltonians H_1, H_2, and H_3 is discrete. Denote the ordered eigenvalues of H_i by $\epsilon_1^{(i)}\leqslant\epsilon_2^{(i)}\leqslant\cdots\leqslant\epsilon_n^{(i)}\leqslant\cdots\leqslant\epsilon_{\text{ess}}^{(i)}$, where $\epsilon_{\text{ess}}^{(i)}$ is the energy at which the essential spectrum (if any) begins. The corresponding eigenfunctions will be called $\phi_n^{(i)}$. The $\phi_n^{(i)}$ are assumed normalized:

$$\int|\phi_n^{(i)}(\mathbf{r})|^2 d^3\mathbf{r}=1. \tag{3.1}$$

It will now be convenient to write operators in continuous matrix notation. Define $H_i^{(n_i)}$, which is a two-particle operator in the three-particle internal space, by

$$H_i^{(n_i)}(\mathbf{r}_{j,k},\boldsymbol{\rho}_i;\mathbf{r}'_{j,k},\boldsymbol{\rho}'_i)=\left[\sum_{n=1}^{n_i-1}(\epsilon_n^{(i)}-\epsilon_{n_i}^{(i)})\phi_n^{(i)}(\mathbf{r}_{j,k})\overline{\phi_n^{(i)}(\mathbf{r}'_{j,k})}+\epsilon_{n_i}^{(i)}\delta(\mathbf{r}_{j,k}-\mathbf{r}'_{j,k})\right]\delta(\boldsymbol{\rho}_i-\boldsymbol{\rho}'_i)\,. \tag{3.2}$$

As a consequence of the familiar variational principle referred to in Theorem 1, the expectation value of $H_i^{(n_i)}$ can never exceed the expectation value of H_i: Expectation values of H_i and $H_i^{(n_i)}$ are the same for functions of the form $\phi_n^{(i)}(\mathbf{r}_{jk})\psi(\boldsymbol{\rho}_i)$, $n=1,2,\ldots,n_i$ (where ψ is arbitrary); for functions orthogonal to these $H_i^{(n_i)}$ has the expectation value $\epsilon_{n_i}^{(i)}$ while the expectation value of H_i cannot be less than $\epsilon_{n_i}^{(i)}$. It follows from Theorem 1 that the eigenvalues of

$$H^{(n_1,n_2,n_3)}\equiv H_1^{(n_1)}+H_2^{(n_2)}+H_3^{(n_3)} \tag{3.3}$$

are lower bounds to the eigenvalues of H_{int}. Clearly, the eigenvalues E of $H^{(n_1,n_2,n_3)}$ satisfy

$$\epsilon_1^{(1)}+\epsilon_1^{(2)}+\epsilon_1^{(3)}\leqslant E\leqslant\epsilon_{n_1}^{(1)}+\epsilon_{n_2}^{(2)}+\epsilon_{n_3}^{(3)}\,. \tag{3.4}$$

$H^{(1,1,1)}$ has only the infiniely degenerate eigenvalue $\epsilon_1^{(1)}+\epsilon_1^{(2)}+\epsilon_1^{(3)}$, which becomes the HPS lower bound when the particles are identical.

The method used above to obtain $H_i^{(n_i)}$ from H_i is known as truncation. It was introduced by H.F. Weinberger[28] and developed for quantum mechanical problems by N.W. Bazley and D.W. Fox.[29]

C. One-particle equations

The Hamiltonian $H^{(n_1,n_2,n_3)}$ is a sum of separable potentials with no kinetic energy. It has been known for a long time that the three-body Schrödinger equation for separable potentials is reducible to one-particle equations even with kinetic energy present.[30] Thus, a reduction to one-particle equations must be possible for $H^{(n_1,n_2,n_3)}$. This reduction can be achieved by making the definition

$$f_l^{(i)}(\boldsymbol{\rho}_i)=(\epsilon_{n_i}^{(i)}-\epsilon_l^{(i)})^{1/2}\int\overline{\phi_l^{(i)}(\mathbf{r}'_{j,k})}\times\delta(\boldsymbol{\rho}_i-\boldsymbol{\rho}'_i)\psi(\mathbf{r}'_{j,k},\boldsymbol{\rho}'_i)d^\delta\mathbf{r}'_{j,k}d^\delta\boldsymbol{\rho}'_i\,. \tag{3.5}$$

Here ψ is an eigenfunction of

$$H^{(n_1,n_2,n_3)}|\psi\rangle=E\,|\psi\rangle \tag{3.6}$$

and $\delta=1,2$, or 3 is the dimension of the physical space in which the vectors $\boldsymbol{\rho}_i$ and $\mathbf{r}_{j,k}$ live. Writing out the Schrödinger equation (3.6) explicitly using Eqs. (3.2), (3.3), and (3.5) yields an equation which can be solved for ψ to obtain

$$\psi=(\epsilon_{n_1}^{(1)}+\epsilon_{n_2}^{(2)}+\epsilon_{n_3}^{(3)}-E)^{-1}\sum_{\substack{i=1\\ i\neq j\neq k\neq i}}^{3}\left\{\sum_{l=1}^{n_i-1}(\epsilon_{n_i}^{(i)}-\epsilon_l^{(i)})^{1/2}\phi_l^{(i)}(\mathbf{r}_{j,k})f_l^{(i)}(\boldsymbol{\rho}_i)\right\}. \tag{3.7}$$

The expression (3.7) for ψ can be inserted in the definition (3.5) of $f_l^{(i)}$ to obtain coupled equations for the $f_l^{(i)}$. The relations

$$d^\delta\mathbf{r}'_{1,2}d^\delta\boldsymbol{\rho}'_3=d^\delta\mathbf{r}'_{2,3}d^\delta\boldsymbol{\rho}'_1=d^\delta\mathbf{r}'_{3,1}d^\delta\boldsymbol{\rho}'_2\,, \tag{3.8}$$

$$r_{j,k}=\frac{(\surd 3)(m_j+m_k)}{2m_k(m_1+m_2+m_3)}\,[m_i\,\boldsymbol{\rho}_i+(m_i+m_k)\boldsymbol{\rho}_j\,]\,, \tag{3.9}$$

and

$$\delta(\rho_i - \rho_i') = \left[\frac{(\sqrt{3})(m_i + m_k)(m_j + m_k)}{2m_k(m_1 + m_2 + m_3)}\right]^{\delta} \delta\left\{\mathbf{r}'_{k,i} + \frac{(\sqrt{3})(m_i + m_k)}{2m_k(m_1 + m_2 + m_3)}\left[(m_j + m_k)\rho_i + m_j\,\rho_j'\right]\right\}, \tag{3.10}$$

where $i \neq j \neq k \neq i$, can be used to bring these coupled equations to the form

$$(E - \epsilon_{n_1}^{(1)} - \epsilon_{n_2}^{(2)} - \epsilon_{n_3}^{(3)}) f_l^{(i)}(\rho_i) = L_l^{(i)} f_l^{(i)}(\rho_i) + \sum_{\substack{j=1 \\ j\neq i\neq k\neq j}}^{3} \sum_{m=1}^{n_j - 1} \int M_{l,m}^{(i,j)}(\rho_i, \rho_j) f_m^{(j)}(\rho_j) d^{\delta}\rho_j\,. \tag{3.11}$$

Here

$$L_l^{(i)} = -(\epsilon_{n_i}^{(i)} - \epsilon_l^{(i)}) \tag{3.12}$$

and

$$M_{l,m}^{(i,j)} = -(\epsilon_{n_i}^{(i)} - \epsilon_l^{(i)})^{1/2}(\epsilon_{n_j}^{(j)} - \epsilon_m^{(j)})^{1/2}\left[\frac{(\sqrt{3})(m_i + m_k)(m_j + m_k)}{2m_k(m_1 + m_2 + m_3)}\right]^{\delta} \overline{\phi_l^{(i)}\left(\frac{\epsilon_{i,j,k}(\sqrt{3})(m_j + m_k)[m_i\,\rho_i + (m_i + m_k)\rho_j]}{2m_k(m_1 + m_2 + m_3)}\right)}$$
$$\times \phi_m^{(j)}\left(\frac{-\epsilon_{i,j,k}(\sqrt{3})(m_i + m_k)[(m_j + m_k)\rho_i + m_j\rho_j]}{2m_k(m_1 + m_2 + m_3)}\right), \tag{3.13}$$

where $\epsilon_{i,j,k}$ is the Levi-Civita symbol

$$\epsilon_{i,j,k} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3). \end{cases} \tag{3.14}$$

When $E = \epsilon_{n_1}^{(1)} + \epsilon_{n_2}^{(2)} + \epsilon_{n_3}^{(3)}$, ψ can not be constructed from Eq. (3.7) and the above reduction must be reexamined. This will not be done here, however, because $\epsilon_{n_1}^{(1)} + \epsilon_{n_2}^{(2)} + \epsilon_{n_3}^{(3)}$ is the upper bound to the spectrum of $H^{(n_1,n_2,n_3)}$ and is therefore of no interest for the construction of lower bounds.

D. Analysis of the one-particle equations

If the functions $f_l^{(i)}(\rho_i)$, $l = 1,2,...,n_i - 1$, $i = 1,2,3$ are thought of as components of a vector, the eigenvalue problem (3.11) can be written in the abstract form

$$(L + M)|f\rangle = \mu|f\rangle, \tag{3.15}$$

where L and M are given by Eqs. (3.12) and (3.13), respectively, and

$$\mu = E - \epsilon_{n_1}^{(1)} - \epsilon_{n_2}^{(2)} - \epsilon_{n_3}^{(3)}\,. \tag{3.16}$$

Because the eigenfunctions $\phi_l^{(i)}$ which appear in M are normalized, M itself is square integrable and therefore completely continuous. L, although not completely continuous, is already in diagonal form so that its spectrum can be immediately written down. Now it is known that the essential spectrum of an operator is not changed when a completely continuous operator is added to it.[31] Thus, the essential spectrum of $L + M$ is determined by L alone. Since L has only discrete eigenvalues of infinite multiplicity (which are the accumulation points of the spectrum of $L + M$), *no lower bounds to discrete eigenvalues of the original problem can be obtained which lie above the bottom of the spectrum of L.* Thus, only the eigenvalues of $K = L + M$ which lie below the bottom of the spectrum of L are of interest.

In practical applications of the present method, it will usually be necessary to find the eigenvalues of $L + M$ by approximate methods. The most obvious approximation is the use of a numerical integration rule to convert the integration which appears in $M|f\rangle$ into a finite sum; this carries the integral equation (3.15) into a finite dimensional matrix eigenvalue problem. Now it might be objected that such a procedure cannot work, because the full kernel $L + M$ is not completely continuous and can therefore not be approximated by a finite dimensional kernel. Such an objection is, however, not valid if the only objective is the computation of the discrete eigenvalues of $L + M$ which lie below the bottom of the spectrum of L. For μ below the bottom of the spectrum of L, $(\mu I - L)^{-1}$ exists (I is the identity), and the integral equation (3.15) can be rewritten as

$$|f\rangle = (\mu I - L)^{-1} M|f\rangle. \tag{3.17}$$

The kernel of Eq. (3.17) is completely continuous since $(\mu I - L)^{-1}$ is bounded (the product of a bounded operator with a completely continuous operator is completely continuous). The use of a numerical integration rule converts Eq. (3.17) into a matrix eigenvalue problem equivalent to the matrix eigenvalue problem obtained by using the same numerical integration rule on the original integral equation (3.15). Since such an approximation is valid for (3.17), it must also be valid for the computation of the discrete eigenvalues of (3.15) which lie below the bottom of the spectrum of L. This point has been discussed in more detail by Hetherington.[32] Approximate computation of the eigenvalues of (3.15) will lead to rigorous lower bounds to the eigenvalues of the original problem only *if* error bounds can be found for the difference between the exact eigenvalues of (3.15) and their approximations. Error bounds on the eigenvalues of integral equations which have been solved approximately by numerical integration have been discussed by Mysovskih.[3]

The existence of discrete eigenvalues below the bottom of the spectrum of L can be discussed with the aid of a trick. The result is as follows:

Theorem 2: Consider the eigenvalue problem (3.15), where both L and M are bounded Hermitian operators on a Hilbert space $\mathscr{H}$ and M is completely continuous. Let l_0 be the greatest lower bound to the spectrum of L:

$$\langle\psi|L|\psi\rangle \geqslant l_0\langle\psi|\psi\rangle \tag{3.18}$$

for all $|\psi\rangle$ in $\mathscr{H}$. Let μ_n be the largest number for which the inequality

$$\langle\psi|(L + M)|\psi\rangle \geqslant \mu_n\langle\psi|\psi\rangle \tag{3.19}$$

holds for all $|\psi\rangle$ in $\mathscr{H}_n$, where $\mathscr{H}_n$ is the subspace of vectors $|\psi\rangle$ in $\mathscr{H}$ which satisfy the constraints $\langle\psi|\psi_i\rangle = 0$,

$j = 0,1,2,\ldots,n-1$. Here the vectors $|\psi_j\rangle$ are eigenvectors of the problem (3.15) belonging to eigenvalues $\mu_j < \mu_n$. Then, if $\mu_n < l_0$, μ_n is a discrete eigenvalue with corresponding eigenvector $|\psi_n\rangle$.

Proof of Theorem 2: Let

$$|\phi\rangle = (L - \mu_n I)^{1/2}|\psi\rangle\,, \tag{3.20}$$

where $(L-\mu_n I)^{1/2}$ is the unique positive definite square root of $(L-\mu_n I)$ (a unique positive definite square root exists because $l_0 > \mu_n$ and L is Hermitian). Then the inequality (3.14) is equivalent to

$$\langle\phi|N|\phi\rangle \leqslant \nu\langle\phi|\phi\rangle\,, \tag{3.21}$$

where $\nu = 1$ and

$$N = -(L-\mu_n I)^{-1/2} M (L-\mu_n I)^{-1/2}\,, \tag{3.22}$$

with $(L-\mu_n I)^{-1/2}$ the inverse of $(L-\mu_n I)^{1/2}$. [This inverse exists in $\mathscr{H}_n$ because $\langle\psi|(L-\mu_n I)^{1/2}|\psi\rangle \geqslant (l_0-\mu_n)^{1/2} \times \langle\psi|\psi\rangle$ for all $|\psi\rangle$ in $\mathscr{H}_n$.] Because $(L-\mu_n I)^{-1/2}$ is bounded, N is completely continuous. Hence, the Euler equation which follows from the variational principle (3.21) is the Hilbert–Schmidt integral equation

$$N|\phi\rangle = \nu|\phi\rangle. \tag{3.23}$$

Because μ_n is the largest number for which the inequality (3.19) holds, equality in (3.19) can be approached arbitrarily closely (or perhaps even achieved) by an appropriate choice of $|\psi\rangle$. Hence, equality can be approached aribitrarily closely in (3.21) with $\nu = 1$. Now Hilbert–Schmidt theory implies that the integral equation (3.23) with $|\phi\rangle$ related to a $|\psi\rangle$ in $\mathscr{H}_n$ by (3.20) has at least one discrete eigenvalue; the fact that equality can be approached arbitrarily closely in (3.21) with $\nu = 1$ implies that $\nu = 1$ is an eigenvalue of (3.23). The corresponding eigenvector is an eigenvector of (3.23) with eigenvalue $\nu = 1$ *and* an eigenvector of (3.15) with eigenvalue μ_n.

E. Identical particles

If $m_1 = m_2 = m_3 = m$ and $V_1 = V_2 = V_3 = V$ so that the particles are identical, the equations of the preceding section simplify somewhat. For three identical particles, the wave function ψ must transform (under permutation of particle indices) as a partner in a basis for one of the irreducible representations of the permutation group on three objects. This permutation group has three irreducible representations: the symmetric (identity) representation, the antisymmetric (alternating) representation, and the (two dimensional faithful) mixed representation. All three can occur for real physical systems; to save space only the results for the symmetric and antisymmetric cases, needed for the examples of the next section, will be recorded here.

For identical particles, the two-particles eigenfunctions $\phi_j^{(i)}$ and eigenvalues $\epsilon_j^{(i)}$ no longer depend on i, so that a simpler notation can be used. If the particles are identical, then the potential V must possess the reflection symmetry $V(\mathbf{r}) = V(-\mathbf{r})$, which implies that the two-particle eigenfunctions can be chosen to be symmetric or antisymmetric under reflection. The superscript i will be dropped; superscripts S and A will be used to distinguish the two kinds of behavior under reflection:

$$\phi_j^{(S)}(\mathbf{r}) = \phi_j^{(S)}(-\mathbf{r})\,, \tag{3.24}$$

$$\phi_j^{(A)}(\mathbf{r}) = -\phi_j^{(A)}(-\mathbf{r})\,. \tag{3.25}$$

The eigenvalue $\epsilon_j^{(S)}$ will belong to $\phi_j^{(S)}$; $\epsilon_j^{(A)}$ will belong to $\phi_j^{(A)}$.

In order to keep the lower bounding Hamiltonian $H^{(n_1,n_2,n_3)}$ symmetric under permutation of particle indices, it is necessary that $n_1 = n_2 = n_3 = n$. It is then easy to show that, for a ψ which transforms according to the symmetric (identity) representation of the permutation group (3.5), (3.7), (3.11), (3.12), and (3.13) are replaced, respectively, by

$$f_i(\boldsymbol{\rho}_3) = (\epsilon_n^{(S)} - \epsilon_i^{(S)})^{1/2} \int \overline{\phi_i^{(S)}(\mathbf{r}'_{1,2})}\, \delta(\boldsymbol{\rho}_3 - \boldsymbol{\rho}'_3) \times \psi(\mathbf{r}'_{1,2}, \boldsymbol{\rho}'_3)\, d^\delta\mathbf{r}'_{1,2}\, d^\delta\boldsymbol{\rho}'_3\,, \tag{3.26}$$

$$\psi(\mathbf{r}_{1,2}, \boldsymbol{\rho}_3) = (3\epsilon_n^{(S)} - E)^{-1} \sum_{j=1}^{n-1} (\epsilon_n^{(S)} - \epsilon_j^{(S)})^{1/2} [\phi_j^{(S)}(\mathbf{r}_{1,2}) \times f_j(\boldsymbol{\rho}_3) + \phi_j^{(S)}(\mathbf{r}_{2,3}) f_j(\boldsymbol{\rho}_1) + \phi_j^{(S)}(\mathbf{r}_{3,1}) f_j(\boldsymbol{\rho}_2)]\,, \tag{3.27}$$

$$(E - 3\epsilon_n^{(S)}) f_i(\boldsymbol{\rho}_3) = \sum_{j=1}^{n-1} \int [L_{i,j}(\boldsymbol{\rho},\boldsymbol{\rho}') + M_{i,j}(\boldsymbol{\rho},\boldsymbol{\rho}')] f_j(\boldsymbol{\rho}'_3)\, d^\delta\boldsymbol{\rho}'_3\,, \tag{3.28}$$

$$L_{i,j}(\boldsymbol{\rho},\boldsymbol{\rho}') = -(\epsilon_n^{(S)} - \epsilon_i^{(S)}) \delta(\boldsymbol{\rho} - \boldsymbol{\rho}') \delta_{i,j}\,, \tag{3.29}$$

and

$$M_{i,j}(\boldsymbol{\rho},\boldsymbol{\rho}') = -2(2/\sqrt{3})^\delta (\epsilon_n^{(S)} - \epsilon_i^{(S)})^{1/2} (\epsilon_n^{(S)} - \epsilon_j^{(S)})^{1/2} \times \overline{\phi_i^{(i)}[(\boldsymbol{\rho} + 2\boldsymbol{\rho}')/\sqrt{3}]}\, \phi_j^{(S)}[-(2\boldsymbol{\rho} + \boldsymbol{\rho}')/\sqrt{3}] \tag{3.30}$$

It should be noted that only the functions $\phi_n^{(S)}$ which are symmetric under reflection appear. If the internal wave function ψ is to transform according to the antisymmetric (alternating) representation of the permutation group, only the functions $\phi_n^{(A)}$ appear; the relevant equations can be obtained by replacing S by A everywhere in Eqs.(3.26)–(3.30). The mixed representation (which can occur for half-integral spin particles) is somewhat more complex, but is still simpler than the case of nonidentical particles.

F. Angular decomposition for identical particles

For three dimensional problems with spherically symmetric pair potentials, the three dimensional integral equations (3.28) can be reduced to one-variable equations by carrying out an angular decomposition. The eigenfunctions of the two-body Hamiltonian introduced in (2.10) have the form

$$\phi_{l_1,m_1,n_1}(\mathbf{r}) = Y_{l_1,m_1}(\theta,\phi) R_{n_1,l_1}(r), \tag{3.31}$$

where Y_{l_1,m_1} is a spherical harmonic. The eigenvalues ϵ_{l_1,n_1} of the two-body Hamiltonian depend only on l_1 and n_1; the approximate Hamiltonian introduced in (3.3), whose eigenvalues are lower bounds, now takes the form

$$H^{(\epsilon)} = \sum_{\substack{i=1\\ i\neq j\neq k\neq i}}^{3} \Big[{\sum_{n_1,l_1}}' \sum_{m_1=-l_1}^{l_1} (\epsilon_{n_1,l_1} - \epsilon) \phi_{l_1,m_1,n_1}(\mathbf{r}_{j,k})$$

$$\times \overline{\phi_{l_1,m_1,n_1}(\mathbf{r}_{j,k})} + \epsilon\delta(\mathbf{r}_{j,k} - \mathbf{r}'_{j,k})\big]\delta(\boldsymbol{\rho}_i - \boldsymbol{\rho}'_i) . \tag{3.32}$$

Here the prime on the sum over n_1 and l_1 means that the sum is restricted to values of n_1 and l_1 for which $\epsilon_{n_1,l_1} < \epsilon$. For ψ belonging to the symmetric (identity) representation, only terms with even l_1 will contribute; for antisymmetric ψ only odd l_1 contributes. The reduction of the Schrödinger equation for the approximate Hamiltonian to an integral equation [Eqs. (3.26)–(3.30)] is still valid with some trivial changes in notation: The single index i is replaced by the triple l_1,m_1,n_1; j is replaced by l'_1,m'_1,n'_1; $\epsilon_n^{(S)}$ is replaced by ϵ.

Because $H^{(\epsilon)}$ in Eq. (3.32) is invariant under rotations, its eigenfunctions ψ can be taken to have definite values for the total angular momentum quantum number L and the z component of angular momentum quantum number M. Equation (3.26) and the addition theory for angular momentum then implies that

$$f_{l_1,m_1,n_1}(\boldsymbol{\rho}) = \sum_{l_2=|L-l_1|}^{L+l_1} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & M-m_1 & -M \end{pmatrix} \times Y_{l_2,M-m_1}(\theta,\phi) g^{(L)}_{l_1,l_2,n_1}(\rho)/\rho , \tag{3.33}$$

where ρ, θ, ϕ are the spherical coordinates of $\boldsymbol{\rho}$ and the coefficient ( ) is the Wigner 3-j symbol.[33] It can be readily shown with the aid of the orthogonality properties of the Wigner 3-j symbol and the spherical harmonics that the functions $g^{(L)}_{l_1,l_2,n}(\rho)$ satisfy the coupled one dimensional integral equations

$$(E-3\epsilon)g^{(L)}_{l_1,l_2,n_1}(\rho) = \sum_{l'_1,n'_1}{}' \sum_{l'_2=|L-l'_1|}^{L+l'_1} \int \big[L^{(L)}_{l_1,l_2,n_1}(\rho,\rho') + M_{l_1,l_2,n_1;l'_1,l'_2,n'_1}(\rho,\rho')\big] g^{(L)}_{l'_1,l'_2,n'_1}(\rho')\,d\rho' , \tag{3.34}$$

where

$$L^{(L)}_{l_1,l_2,n_1;l'_1,l'_2,n'_1}(\rho,\rho') = -(\epsilon - \epsilon_{l_1,n_1})\delta(\rho-\rho')\delta_{l_1,l'_1}\delta_{l_2,l'_2}\delta_{n_1,n'_1} \tag{3.35}$$

and

$$M^{(L)}_{l_1,l_2,n_1;l'_1,l'_2,n'_1}(\rho,\rho') = -\frac{16}{3\sqrt{3}}(2L+1)(\epsilon-\epsilon_{l_1,n_1})^{1/2}(\epsilon-\epsilon_{l'_1,n'_1})^{1/2} \sum_{m_1,m'_1} \begin{pmatrix} l_1 & l_2 & L \\ m_1 & (M-m_1) & -M \end{pmatrix} \begin{pmatrix} l'_1 & l'_2 & L \\ m_1 & (M-m_1) & -M \end{pmatrix}$$
$$\times \rho\rho' \int d\Omega\, d\Omega'\, \overline{Y_{l_2,M-m_1}(\theta,\phi)}\, Y_{l'_2,M-m'_1}(\theta',\phi')\, \overline{\phi_{l_1,m_1,n_1}((\boldsymbol{\rho}+2\boldsymbol{\rho}')/\sqrt{3})}\, \phi_{l'_1,m'_1,n'_1}(-(2\boldsymbol{\rho}+\boldsymbol{\rho}')/\sqrt{3}) . \tag{3.36}$$

Because l_1 and l'_1 are both even for the symmetric case and both odd for the antisymmetric case, the behavior of the integrand of Eq. (3.36) under inversion (parity) implies that $L^{(L)}$ vanishes unless l_2 and l'_2 are either both even or both odd.

IV. EXAMPLES: THREE IDENTICAL PARTICLES COUPLED BY SPRINGS

Particles coupled by springs are the traditional exactly solvable model on which bounds of HPS type have been tested. This section applies the methods of the present paper to three identical particles coupled by springs in one dimension and in three dimensions. No attempt will be made to present examples with nonidentical particles, because these examples appear to require that the integral equations (3.11) be solved numerically. The Schrödinger equation for the internal motion of three particles coupled by springs can be written in the dimensionless form

$$\big[-(p_1^2+p_2^2+p_3^2)/(2\hbar^2) + P^2/(6\hbar^2) - \tfrac{2}{3}(r_{1,2}^2 + r_{2,3}^2 + r_{3,1}^2) + E\big]\psi = 0 . \tag{4.1}$$

Construction of the exact eigenfunctions and eigenvalues of (4.1) is outlined in Appendix A. The exact eigenvalues are compared with the lower bounds in Tables I–III. Certain expansion theorems needed for the analytic solution of the integral equations (3.28) and (3.34) are developed in Appendix B.

A. One-dimensional examples

The two-body Hamiltonian introduced in (2.10) is now

$$H_3(r_{1,2},p_{1,2}) = \frac{2p_{1,2}^2}{3\hbar^2} + \frac{2}{3}r_{1,2}^2 . \tag{4.2}$$

The normalized eigenfunctions and eigenvalues of Eq. (4.2) in the one dimensional case are

$$\chi_l(x) = \pi^{-1/4}(2^l l!)^{-1/2} e^{-x^2/2} H_l(x) \tag{4.3}$$

and

$$\epsilon_l = \tfrac{2}{3}(2l+1) , \tag{4.4}$$

where $H_l(x)$ is the lth Hermite polynomial defined by the Rodrigues' formula

$$H_l(x) = (-1)^l e^{x^2}(d^l/dx^l)e^{-x^2} . \tag{4.5}$$

and $x = r_{1,2}$.

The integral equation (3.28) has been written out and its solutions obtained for $n = 2,3$, and 4 in both the symmetric and antisymmetric cases. The results are summarized in Tables I and II; $n=1$ is the HPS bound. Details are supplied below for two representative cases; the reader should have no trouble filling in the details for the other cases listed in Tables I and II. The variables $u = (2\rho+\rho')/\sqrt{3}$, $v = (\rho+2\rho')/\sqrt{3}$ have been introduced to save space. For $n=2$ in the symmetric case, $\epsilon_1^{(S)} = 2/3, \epsilon_n^{(S)} = \epsilon_2^{(S)} = 10/3, \phi_1^{(S)} = \chi_1$, and (3.28) becomes the integral equation

$$\left(E-\frac{22}{3}\right)f_1(\rho) = -\frac{32}{3\sqrt{3}}\int \chi_1(v)\chi_1(u)f_1(\rho')\,d\rho'$$
$$= -\frac{16}{3}\sum_{k=0}^{\infty}(-2)^{-k}\chi_k(\rho)$$
$$\times\int \chi_k(\rho')f_1(\rho')\,d\rho', \qquad (4.6)$$

where the second equality follows from one of the expansions given in Appendix B. The eigenvalues of (3.41) are

$$E_k = (22/3) - (16/3)(-2)^{-k}, \quad k = 0,1,2,\cdots, \qquad (4.7)$$

with corresponding eigenfunctions $f_1 = \chi_k$.

For $n = 4$ in the antisymmetric case, $\epsilon_1^{(A)} = 2$, $\epsilon_2^{(A)} = (14/3)$, $\epsilon_3^{(A)} = (22/3)$, $\epsilon_4^{(A)} = 10$, $\phi_1^{(A)} = \chi_1$, $\phi_2^{(A)} = \chi_3$, $\phi_3^{(A)} = \chi_5$, and Eq. (3.31) becomes the coupled integral equations

$$\begin{aligned}(E-22)f_1(\rho) &= [32/(3\sqrt{3})]\int \chi_1(v)[3\chi_1(u)f_1(\rho') + 6^{1/2}\chi_3(u)f_2(\rho') + \sqrt{3}\chi_5(u)f_3(\rho')]d\rho'\\
&= \int [8\chi_0(\rho)\chi_0(\rho') + 8\chi_1(\rho)\chi_1(\rho') - 10\chi_2(\rho)\chi_2(\rho') + 8\chi_3(\rho)\chi_3(\rho')]f_1(\rho')d\rho'\\
&\quad + 6\sqrt{2}\chi_2(\rho)\int \chi_0(\rho')f_2(\rho')\,d\rho' - \sum_{k=0}^{\infty}(-2)^{-k-2}\chi_{k+4}(\rho)\int \{2(3k+11)\chi_{k+4}(\rho')f_1(\rho')\\
&\quad + 6(k+1)[(k+3)(k+4)]^{1/2}\chi_{k+2}(\rho')f_2(\rho')\\
&\quad + 3(3k-5)[(k+1)(k+2)(k+3)(k+4)/10]^{1/2}\chi_k(\rho')f_3(\rho')\}\,d\rho',\end{aligned} \qquad (4.8)$$

$$\begin{aligned}(E-\tfrac{74}{3})f_2(\rho) &= [32/(3\sqrt{3})]\int \chi_3(v)[6^{1/2}\chi_1(u)f_1(\rho') + 2\chi_3(u)f_2(\rho') + \sqrt{2}\chi_5(u)f_3(\rho')]\,d\rho'\\
&= 6\sqrt{2}\chi_0(\rho)\int \chi_2(\rho')f_1(\rho')d\rho' + \int [\tfrac{4}{3}\chi_0(\rho)\chi_0(\rho') + \tfrac{16}{3}\chi_1(\rho)\chi_1(\rho')]f_2(\rho')d\rho'\\
&\quad - \sum_{k=0}^{\infty}(-2)^{-k-2}\chi_{k+2}(\rho)\int \{6(k+1)[(k+3)(k+4)]^{1/2}\chi_{k+4}(\rho')f_1(\rho')\\
&\quad + \tfrac{2}{3}(9k^3-45k-20)\chi_{k+2}(\rho')f_2(\rho') + (9k^3-72k^2+123k-20)\\
&\quad \times[(k+1)(k+2)/10]^{1/2}\chi_k(\rho')f_3(\rho')\}\,d\rho',\end{aligned} \qquad (4.9)$$

$$\begin{aligned}(E-\tfrac{82}{3})f_3(\rho) &= [32/(3\sqrt{3})]\int \chi_5(v)[\sqrt{3}\chi_1(u)f_1(\rho') + \sqrt{2}\chi_3(u)f_2(\rho') + \chi_5(u)f_3(\rho')]d\rho'\\
&= -\sum_{k=0}^{\infty}(-2)^{-k-2}\chi_k(\rho)\int \{3(3k-5)[(k+1)(k+2)(k+3)(k+4)/10]^{1/2}\chi_{k+4}(\rho')f_1(\rho')\\
&\quad + (9k^3-72k^2+123k-20)[(k+1)(k+2)/10]^{1/2}\chi_{k+2}(\rho')f_2(\rho')\\
&\quad + [(81k^5-1485k^4+8685k^3-18\,675k^2+11\,994k-40)/60]\\
&\quad \times\chi_k(\rho')f_3(\rho')\}d\rho'.\end{aligned} \qquad (4.10)$$

Among the solutions are $f_1 = 3\chi_2, f_2 = -\sqrt{2}\chi_0, f_3 = 0$ for $E = 8$ and the following solutions for $E = 30$: $f_1 = \chi_0, f_2 = f_3 = 0$; $f_1 = \chi_1, f_2 = f_3 = 0$; $f_1 = \chi_3, f_2 = f_3 = 0$; $f_1 = 0, f_2 = \chi_1, f_3 = 0$; $f_1 = \sqrt{2}\chi_2, f_2 = 3\chi_0, f_3 = 0$. The remaining solutions have the form $f_1 = a_k\chi_{k+4}$, $f_2 = b_k\chi_{k+2}$, $f_3 = c_k\chi_k$, where a_k, b_k, c_k, and E are determined by solving the set of equations

$$\begin{aligned}&[2(3k+11) + (-2)^{k+2}(E-22)]a_k + 6(k+1)[(k+3)(k+4)]^{1/2}b_k + 3(3k-5)\\
&\quad\times[(k+1)(k+2)(k+3)(k+4)/10]^{1/2}c_k = 0,\end{aligned} \qquad (4.11)$$

$$\begin{aligned}&6(k+1)[(k+3)(k+4)]^{1/2}a_k + [6k^3-30k-(40/3)+(-2)^{k+2}(E-\tfrac{74}{3})]b_k\\
&\quad + (9k^3-72k^2+123k-20)[(k+1)(k+2)/10]^{1/2}c_k = 0,\end{aligned} \qquad (4.12)$$

$$\begin{aligned}&3(3k-5)[(k+1)(k+2)(k+3)(k+4)/10]^{1/2}a_k + (9k^3-72k^2+123k-20)[(k+1)(k+2)/10]^{1/2}b_k\\
&\quad + \{[(81k^5-1485k^4+8685k^3-18\,675k^2+11\,994k-40)/60] + (-2)^{k+2}(E-\tfrac{82}{3})\}c_k = 0.\end{aligned} \qquad (4.13)$$

B. Harmonic oscillators in three dimensions

This subsection illustrates the angular decomposition of Sec. III F by using Eqs. (3.34)–(3.36) to calculate lower bounds for the symmetric and antisymmetric states of the Schrödinger equation (4.1) in three dimensions. Only the lowest states (of the appropriate symmetry) of the two-body Hamiltonian (4.2) are kept in each of the two examples. Table III summarizes the results.

The normalized eigenfunctions and eigenvalues of the two-body Hamiltonian (4.2) in the three dimensional case are

$$\chi_{l,m,n}(\mathbf{r}) = Y_{l,m}(\theta,\phi)R_{n,l}(r) \qquad (4.14)$$

and

$$\epsilon_{l,n} = 2(4n+2l+3)/3,$$

where

$$R_{n,l}(r) = [2n!/\Gamma((2n+2l+3)/2)]^{1/2}r^l \times\exp(-r^2/2)L_n^{(l+\frac{1}{2})}(r^2), \qquad (4.16)$$

with $L_n^{(\alpha)}$ the generalized Laguerre polynomial defined by the Rodrigues formula

$$L_n^{(\alpha)}(x) = x^{-\alpha}(n!)^{-1}(d^n/dx^n)(e^{-x}x^{n+\alpha}). \qquad (4.17)$$

If only the lowest state, for which $l_1 = m_1 = n_1 = 0$, is kept in the symmetric case, then $\epsilon_{0,0} = 2$ and $\epsilon = \epsilon_{0,1} = \epsilon_{2,0} = 14/3$. The kernel (3.36) can be evaluated with the aid of the expansion given in Eq. (B9) of Appendix B; the integral equation (3.34) for the L th partial wave then takes the form

$$\left(E - \tfrac{34}{3}\right) g_{0,L,0}^{(L)}(\rho) = -\tfrac{16}{3} \sum_{n=0}^{\infty} (-2)^{-2n-L} \rho R_{n,L}(\rho) \times \int_0^{\infty} \rho' R_{n,L}(\rho') g_{0,L,0}^{(L)}(\rho')\, d\rho' . \quad (4.18)$$

The eigenvalues of (4.18) are

$$E = \tfrac{34}{3} - \tfrac{16}{3}(-2)^{-2n-L}, \quad (4.19)$$

with corresponding eigenfunctions $g_{0,L,0}^{(L)} = \rho R_{n,L}(\rho)$.

If only the lowest states, for which $l_1 = 1, m_1 = 0, \pm 1$, and $n_1 = 0$, are kept in the antisymmetric case, then $\epsilon_{1,0} = 10/3$, and $\epsilon = \epsilon_{1,1} = \epsilon_{3,0} = 6$. The kernel (3.36) can be evaluated with the expansion given in Eq. (B23) of Appendix B; the coupled integral equations (3.34) for the L th partial wave then take the form

$$\left(E - \tfrac{46}{3}\right) g_{1,L+1,0}^{(L)}(\rho) = \tfrac{32}{3} \sum_{n=1}^{\infty} (-2)^{-2n-L} \rho R_{n-1,L+1}(\rho) \times \Big\{ [-2 + 3n(L+1)(2L+1)^{-1}] \int_0^{\infty} R_{n-1,L+1}(\rho') g_{1,L+1,0}^{(L)}(\rho') \rho'\, d\rho' + 3[nL(L+1)(n+L+\tfrac{1}{2})]^{1/2}(2L+1)^{-1} \int_0^{\infty} R_{n,L-1}(\rho') g_{1,L-1,0}^{(L)}(\rho')\rho'\, d\rho' \Big\}, \quad (4.20)$$

$$\left(E - \tfrac{46}{3}\right) g_{1,L,0}^{(L)}(\rho) = \tfrac{32}{3} \sum_{n=0}^{\infty} (-2)^{-2n-L} \rho R_{n,L}(\rho) \int_0^{\infty} R_{n,L}(\rho') g_{1,L,0}(\rho')\rho'\, d\rho', \quad L \neq 0, \quad (4.21)$$

and

$$\left(E - \tfrac{46}{3}\right) g_{1,L-1,0}^{(L)}(\rho) = \tfrac{32}{3} \sum_{n=0}^{\infty} (-2)^{-2n-L} \rho R_{n,L-1}(\rho) \Big\{ 3[nL(L+1)(n+L+\tfrac{1}{2})]^{1/2}(2L+1)^{-1} \times \int_0^{\infty} R_{n-1,L+1}(\rho') g_{1,L+1,0}^{(L)}(\rho')\rho'\, d\rho' + [-2 + 3L(n+L+\tfrac{1}{2})(2L+1)^{-1}] \int_0^{\infty} R_{n,L-1}(\rho') g_{1,L-1,0}^{(L)}(\rho')\rho'\, d\rho' \Big\}, \quad L \neq 0. \quad (4.22)$$

It should be noted that even l_2 and odd l_2 are uncoupled from one another, as they must be, in Eqs. (4.20)–(4.22). The restriction $L \neq 0$ in (4.21) and (4.22) arises from the summation limit $L + 1 \leqslant l_2' \leqslant |L-1|$ in Eq. (3.34) (with $l_1' = 1$) and the fact that l_2 must be nonnegative. The eigenvalues of (4.21) are

$$E = \tfrac{46}{3} + \tfrac{32}{3}(-2)^{-2n-L}, \quad L \neq 0, \quad (4.23)$$

with corresponding eigenfunctions $g_{1,L,0}^{(L)}(\rho) = \rho R_{n,L}(\rho)$. The eigenvalues of the coupled pair (4.20) and (4.22) are

$$E = \tfrac{46}{3} - \tfrac{64}{3}(-2)^{-2n-L}, \quad n \neq 0 \text{ and } L \neq 0, \quad (4.24)$$

and

$$E = \tfrac{46}{3} + \tfrac{16}{3}(-2)^{-2n-L}(6n + 3L - 4), \quad n \neq 0 \text{ or } L \neq 0, \quad (4.25)$$

with corresponding eigenfunctions $g_{1,L+1,0}^{(L)}(\rho) = a\rho R_{n-1,L+1}(\rho)$, $g_{1,L-1,0}^{(L)}(\rho) = b\rho R_{n,L-1}(\rho)$, where the constants a and b can be calculated from an easily obtained secular equation.

V. DISCUSSION

Upon reflection, two obvious questions arise with respect to the method of the preceding sections. Can the lower bounds obtained be pushed arbitrarily close to the exact ei-

TABLE I. Lower bounds to energy eigenvalues for the symmetric states of a one dimensional three-body problem with harmonic forces, $n = 1$ is the HPS bound.

Exact	$n=1$	$n=2$	$n=3$	$n=4$
2.0000	2.0000	2.0000	2.0000	2.0000
6.0000	2.0000	6.0000	6.0000	6.0000
8.0000	2.0000	7.0000	8.0000	8.0000
10.0000	2.0000	7.2500	10.0000	10.0000
12.0000	2.0000	7.3125	11.2500	12.0000
14.0000	2.0000	7.3281	12.0000	14.0000
14.0000	2.0000	7.3320	12.3750	14.0000
16.0000	2.0000	7.3330	12.6250	15.3125
18.0000	2.0000	7.3333	12.6427	16.0000
18.0000	2.0000	7.3333	12.6632	16.6427
⋮	⋮	⋮	⋮	⋮
∞	2.0000	7.3333	12.6667	18.0000

TABLE II. Lower bounds to energy eigenvalues for the antisymmetric states of a one dimensional three body problem with harmonic forces. $n = 1$ is the HPS bound.

Exact	$n=1$	$n=2$	$n=3$	$n=4$
8.0000	6.0000	8.0000	8.0000	8.0000
12.0000	6.0000	9.5000	12.0000	12.0000
14.0000	6.0000	10.6250	13.3750	14.0000
16.0000	6.0000	11.0938	14.0000	16.0000
18.0000	6.0000	11.2578	14.6708	17.3438
20.0000	6.0000	11.3105	15.7500	18.0000
20.0000	6.0000	11.3267	15.8897	18.6699
22.0000	6.0000	11.3314	16.5350	19.4688
24.0000	6.0000	11.3328	16.5625	19.9916
24.0000	6.0000	11.3332	16.6454	20.0000
⋮	⋮	⋮	⋮	⋮
∞	6.0000	11.3333	16.6667	22.0000

TABLE III. Lower bounds to energy eigenvalues for a three-body problem with harmonic forces. The numbers in parentheses specify the number of times a listed eigenvalue occurs.

Exact	HPS lower bounds	Lower bounds from Eq. (4.19)
Symmetric States		
6.0000(1)	6.0000(1)	6.0000(1)
10.0000(6)	6.0000(6)	10.0000(6)
12.0000(10)	6.0000(10)	11.0000(10)
14.0000(21)	6.0000(21)	11.0000(5), 11.2500(16)
16.0000(45)	6.0000(45)	11.2500(12), 11.3125(33)
18.0000(83)	6.0000(83)	11.3125(12), 11.3281(66), 11.3320(5)
⋮	⋮	⋮
∞	6.0000	11.3333

Exact	HPS lower bounds	Lower bounds from Eq. (4.23)–(4.25)
Antisymmetric states		
10.0000(3)	10.0000(3)	10.0000(3)
12.0000(10)	10.0000(10)	12.0000(10)
14.0000(15)	10.0000(15)	13.5000(15)
16.0000(45)	10.0000(45)	13.5000(6), 14.0000(15), 14.6250(24)
18.0000(73)	10.0000(73)	14.6250(12), 15.0000(35), 15.0938(26)
20.0000(126)	10.0000(126)	15.0938(29), 15.2500(63), 15.2578(34)
⋮	⋮	⋮
∞	10.0000	15.3333

genvalues? Can the method be extended to more than three particles? These questions are discussed in order below.

If, as is typical in nuclear and molecular physics, the interparticle potentials V_i are such that the two-body Hamiltonians H_i introduced in Eq. (2.10) have a finite number of discrete levels below a continuum, the best that can be done is to choose $n_i - 1$ equal to the number of discrete levels of H_i and $\epsilon^{(i)}_{n_i}$ equal to the energy at which the continuum begins, so that the sum in $H_i^{(n_i)}$ defined in Eq. (3.2) includes all of the discrete levels of H_i. If the resulting lower bounds are not good enough, an alternative strategy is required. One possible way around the difficulty is to split the potential V into attractive and repulsive parts, put a coupling constant λ in front of the attractive part, and let λ instead of E be the eigenvalue. Such a strategy, which has the advantage of eliminating the continuum, was employed successfully by Bazley and Fox[34] on a one dimensional (radial) Schrödiner equation.

Extensions to more than three particles can be made in a variety of ways. The N-particle Hamiltonian

$$H = \sum_{i=1}^{N} p_i^2/(2m_i) + \sum_{i<j} V_{i,j}(\mathbf{r}_{i,j}) \tag{5.1}$$

can be brought to the form

$$H = \Big(\sum_{i=1}^{N} \mathbf{p}_i\Big)^2/(2M) + H_{\text{int}} , \tag{5.2}$$

where

$$M = \sum_{i=1}^{N} m_i \tag{5.3}$$

and

$$H_{\text{int}} = \sum_{i<j} H_{i,j} , \tag{5.4}$$

with

$$H_{i,j} \equiv (m_i + m_j)^2 p_{i,j}^2/(2m_i m_j M) + V_{i,j}(\mathbf{r}_{i,j}) . \tag{5.5}$$

In the above equations, $\mathbf{r}_{i,j}$ and $\mathbf{p}_{i,j}$ are the canonically conjugate internal coordinates and momenta defined by Eqs. (2.3) and (2.6). One obvious approach is to truncate the $H_{i,j}$ in (5.4); this will in general lead to equations in $N-2$ vector variables. However, it is possible to do better. For example, for four particles H_{int} can be expressed as

$$H_{\text{int}} = H_a + H_b + H_c , \tag{5.6}$$

where

$$\begin{aligned} H_a &= H_{12} + H_{34}, \quad H_b = H_{13} + H_{24}, \\ H_c &= H_{14} + H_{23} . \end{aligned} \tag{5.7}$$

The eigenvalue problems for H_a,H_b, and H_c can be reduced to eigenvalue problems for $H_{i,j}$ via separation of variables; if H_a,H_b, and H_c are truncated, equations in one vector variable (instead of two vector variables) are obtained.

VI. ACKNOWLEDGMENTS

The author is grateful to the University of Wisconsin Physics Department, and especially to Professor Lou Bruch, for hospitality during the summer of 1975 during which a major portion of the present work was completed. Support during that summer from a General Electric Company grant to the Wisconsin physics department was greatly appreciated.

APPENDIX A: EXACT SOLUTIONS FOR HARMONIC OSCILLATORS

This appendix derives the exact eigenfunctions and eigenvalues for the harmonic oscillator examples.

Introduce new coordinates r_i,θ_i via

$$r^{(i)}_{1,2} = r_i \cos\theta_i , \tag{A1}$$

$$\rho^{(i)}_3 = r_i \sin\theta_i , \tag{A2}$$

where $r^{(i)}_{1,2}$ and $\rho^{(i)}_3$ are the ith Cartesian coordinates of $\mathbf{r}_{1,2}$ and $\mathbf{\rho}_3$. These changes of variables bring the Schrödinger equation (4.1) to the form

$$\left\{\sum_{i=1}^{\delta}\left[\left(\frac{\partial^2}{\partial r_i^2} + \frac{1}{r_i}\frac{\partial}{\partial r_i} + \frac{1}{r_i^2}\frac{\partial^2}{\partial\theta_i^2} - r_i^2\right)\right] + E\right\}\psi = 0 , \tag{A3}$$

where, as before, $\delta = 1,2$, or 3 is the dimension of the space. Equations (A1) and (A2) are obtained by regarding $r^{(i)}_{1,2}$ and $\rho^{(i)}_3$ as a pair of Cartesian coordinates; r_i and θ_i are the corresponding polar coordinates. This transformation is motivated by the fact that particle permutations are equivalent to rotations in the $r^{(i)}_{1,2} - \rho^{(i)}_3$ planes for $0 \leqslant i \leqslant \delta$, as is spelled out explicitly below. The eigenvalues and (unnormalized) eigenfunctions of this are[35]

$$E = 2\sum_{i=1}^{\delta} (|m_i| + 2n_i + 1) \tag{A4}$$

and

$$\psi = \prod_{j=1}^{\delta} \exp(-r_j^2/2) r_j^{|m_j|} L_{n_j}^{(|m_j|)}(r_j^2) \exp(im_j\theta_j), \quad \text{(A5)}$$

where $L_n^{(\alpha)}$ is a generalized Laguerre polynomial as defined in Eq. (4.17). The m_j are integers, and the n_j are nonnegative integers.

The behavior required of ψ under permutation of particle coordinates imposes restrictions on the quantum numbers m_j Under interchange of $\mathbf{r}_1$ and $\mathbf{r}_2$, θ_j goes into $\pi - \theta_j$ all $j, 0 \leqslant j \leqslant \delta$. Under the cyclic permutation $\mathbf{r}_1 \to \mathbf{r}_2, \mathbf{r}_2 \to \mathbf{r}_3, \mathbf{r}_3 \to \mathbf{r}_1$, θ_j goes into $\theta_j + 2\pi/3$, all j, $0 \leqslant j \leqslant \delta$. Since one interchange and one cyclic permutation generate the permutation group on three objects, it is sufficient to require proper behavior of ψ under these two permutations. For a symmetric ψ, symmetry under the $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ interchange implies that the dependence of ψ on the θ_j is given by

$$\psi \propto \cos\left[\sum_{j=1}^{\delta} m_j(\theta_j - \pi/2)\right]. \quad \text{(A6)}$$

For an antisymmetric ψ, antisymmetry under the $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ interchange implies that the dependence of ψ on the θ_j is given by

$$\psi \propto \sin\left[\sum_{j=1}^{\delta} m_j(\theta_j - \pi/2)\right]. \quad \text{(A7)}$$

Both the symmetric and the antisymmetric ψ must be invariant under the cyclic permutation $\mathbf{r}_1 \to \mathbf{r}_2, \mathbf{r}_2 \to \mathbf{r}_3, \mathbf{r}_3 \to \mathbf{r}_1$. This fact used in Eqs. (A6) and (A7) implies that

$$\sum_{j=1}^{\delta} m_j = 0, \pm 3, \pm 6, \pm 9, \cdots. \quad \text{(A8)}$$

The expression (A4) for the eigenvalue E combined with the restriction (A8) and the fact that not all m_j can be zero in the antisymmetric case yields the exact eigenvalues which are compared with lower bounds in Tables I–III.

APPENDIX B: EXPANSION THEOREMS

This Appendix derives the expansion theorems used to solve the integral equations in the harmonic oscillator examples.

The expansion theorems used in the one dimensional three-particle examples of Sec. IV A are all obtained from the basic expansion theorem

$$\chi_0(u)\chi_0(v) = \tfrac{1}{2}\sqrt{3} \sum_{k=0}^{\infty} (-2)^{-k} \chi_k(\rho)\chi_k(\rho'), \quad \text{(B1)}$$

with the aid of raising and lowering operators. The functions χ_k in Eq. (B1) are the normalized one dimensional harmonic oscillator eigenfunctions defined in Eq. (4.3); u and v are related to ρ and ρ' by $u = (2\rho + \rho')/\sqrt{3}$ and $v = (\rho + 2\rho')/\sqrt{3}$. The basic theorem (B1) can be obtained by letting $x = \rho, y = \rho', z = -\frac{1}{2}$ in Mehler's formula[36]

$$\exp\left[\frac{2xyz - (x^2 + y^2)z^2}{1 - z^2}\right] = (1 - z^2)^{1/2} \sum_{k=0}^{\infty} \frac{(z/2)^k}{k!} H_k(z) H_k(y). \quad \text{(B2)}$$

The raising and lowering operators $x \pm \partial/\partial x$ have the properties

$$\left(x + \frac{\partial}{\partial x}\right)\chi_k(x) = (2k)^{1/2}\chi_{k-1}(x) \quad \text{(B3)}$$

and

$$\left(x - \frac{\partial}{\partial x}\right)\chi_k(x) = [2(k+1)]^{1/2}\chi_{k+1}(x). \quad \text{(B4)}$$

Define $S_l(u,v)$ by

$$S_l(u,v) = \sum_{k=0}^{\infty} a_k^{(l)}(\rho)\chi_k(\rho)\chi_{k+l}(\rho'), \quad \text{(B5)}$$

where the coefficients $a_k^{(l)}$ are for the moment left unspecified. Then it can be shown that

$$\left(u - \frac{\partial}{\partial u}\right)S_l(u,v) = \sum_{k=0}^{\infty} \{2(2k/3)^{1/2} a_{k-1}^{(l)} + [2(k+l)/3]^{1/2} a_k^{(l)}\} \times \chi_k(\rho)\chi_{k+l-1}(\rho') \quad \text{(B6)}$$

and

$$\left(v - \frac{\partial}{\partial v}\right)S_l(u,v) = \sum_{k=0}^{\infty} \{[2(k+1)/3]^{1/2} a_{k+1}^{(l)} + 2[2(k+l+1)/3]^{1/2} a_k^{(l)}\} \times \chi_k(\rho)\chi_{k+l+1}(\rho'). \quad \text{(B7)}$$

Repeated application of the raising operator $u - \partial/\partial u$ and $v - \partial/\partial v$ to the basic theorem (B1), with the results evaluated by using Eqs. (B3)–(B7), produces expansions of the form

$$\chi_m(u)\chi_{m+l}(v) = \sum_{k=0}^{\infty} (-2)^{-k} c_k^{(m,l)} \chi_k(\rho)\chi_{k+l}(\rho'). \quad \text{(B8)}$$

The coefficients required for the examples of Sec. IV A are
$c_k^{(0,0)} = \sqrt{3}/2$, $c_k^{(1,0)} = \sqrt{3}(-3k+1)/4$,
$c_k^{(2,0)} = \sqrt{3}(9k^2 - 21k + 2)/16$,
$c_k^{(3,0)} = \sqrt{3}(-9k^3 + 54k^2 - 63k + 2)/32$,
$c_k^{(4,0)} = \sqrt{3}(27k^4 - 306k^3 + 945k^2 - 762k + 8)/256$,
$c_k^{(5,0)} = \sqrt{3}(-81k^5 + 1485k^4 - 8685k^3 + 18\,675k^2 - 11\,994k + 40)$, $c_k^{(0,1)} = 3(k+1)^{1/2}/4$,
$c_k^{(1,1)} = 3(-3k+2)[2(k+1)]^{1/2}/16$,
$c_k^{(2,1)} = 3(3k^2 - 9k + 2)[3(k+1)]^{1/2}/32$,
$c_k^{(3,1)} = 3(-9k^3 + 63k^2 - 90k + 8)(k+1)^{1/2}/128$,
$c_k^{(4,1)} = 3(27k^4 - 342k^3 + 1197k^2 - 1122k + 40)[(k+1)/5]^{1/2}/512$,
$c_k^{(0,2)} = 3[3(k+1)(k+2)/2]^{1/2}/8$,
$c_k^{(1,2)} = 9(-k+1)[2(k+1)(k+2)]^{1/2}/32$,
$c_k^{(2,2)} = 9(3k^2 - 11k + 4)[(k+1)(k+2)]^{1/2}/128$,
$c_k^{(3,2)} = 3(-9k^3 + 72k^2 - 123k + 20) \times [3(k+1)(k+2)/5]^{1/2}/256$,
$c_k^{(0,3)} = 3[3(k+1)(k+2)(k+3)/2]^{1/2}/16$,
$c_k^{(1,3)} = 3(-3k+4)[3(k+1)(k+2)(k+3)/2]^{1/2}/64$,
$c_k^{(2,3)} = 3(9k^2 - 39k + 20) \times [3(k+1)(k+2)(k+3)/5]^{1/2}/256$,
$c_k^{(0,4)} = 9[(k+1)(k+2)(k+3)(k+4)/2]^{1/2}/64$,
$c_k^{(1,4)} = 9(-3k+5) \times [(k+1)(k+2)(k+3)(k+4)/10]^{1/2}/128$,
and
$c_k^{(0,5)} = 9[3(k+1)(k+2)(k+3)(k+4)(k+5)/10]^{1/2}/128$.

The expansion theorems used in the three dimensional three-particle examples of Sec. IV B are obtained from the basic expansion theorem

$$\chi_{0,0,0}(\mathbf{u})\overline{\chi_{0,0,0}(\mathbf{v})}$$

$$= (3/4)^{3/2} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{n=0}^{\infty} (-2)^{2n-l} \chi_{l,m,n}(\boldsymbol{\rho}) \overline{\chi_{l,m,n}(\boldsymbol{\rho}')} . \tag{B9}$$

The functions $\chi_{l,m,n}$ in (B9) are the normalized three dimensional oscillator eigenfunctions defined in Eqs. (4.14) and (4.16); $\mathbf{u}$ and $\mathbf{v}$ are related to $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$ by $\mathbf{u} = (2\boldsymbol{\rho} + \boldsymbol{\rho}')/\sqrt{3}$ and $\mathbf{v} = (\boldsymbol{\rho} + 2\boldsymbol{\rho}')/\sqrt{3}$. The basic theorem (B9) can be obtained from the well-known expansion[37]

$$\exp(z \cos\alpha) = (\pi/2)^{1/2} \sum_{l=0}^{\infty} (2l+1) P_l(\cos\alpha) z^{-1/2} I_{l+\frac{1}{2}}(z) , \tag{B10}$$

the addition theorem for the spherical harmonics

$$P_l(\cos\beta) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{l,m}(\theta,\phi) \overline{Y_{l,m}(\theta',\phi')}, \tag{B11}$$

and the Hille–Hardy formula[38]

$$[(1-t)t^{\alpha/2}]^{-1} \exp[-\tfrac{1}{2}(x+y)(1+t)/(1-t)] I_\alpha [2(xyt)^{1/2}/(1-t)]$$
$$= \sum_{n=0}^{\infty} [n!/\Gamma(n+\alpha+1)] t^n \exp[-\tfrac{1}{2}(x+y)](xy)^{\alpha/2} L_n^{(\alpha)}(x) L_n^{(\alpha)}(y) . \tag{B12}$$

P_l in (B10) and (B11) is a Legendre polynomial; $I_{l+\frac{1}{2}}$ in (B10) and I_α in (B12) are modified Bessel functions of the first kind. In Eq. (B11), β is the angle between the directions specified by the spherical coordinates (θ,ϕ) and (θ',ϕ'). The other functions appearing in (B10)–(B12) are defined in the text. By using (4.14) and (4.16) which define the $\chi_{l,m,n}$, formula (B10) with $z = 4\rho\rho'/3$, $\alpha = \pi - \beta$, and the fact that $P_l[\cos(\pi - \beta)] = (-1)^l P_l(\cos\beta)$, it is easy to show that

$$\chi_{0,0,0}(\mathbf{u})\chi_{0,0,0}(v) = (\pi/\sqrt{2})^{-1} \exp[-5(\rho^2+\rho'^2)/6] \sum_{l=0}^{\infty} (-1)^l (2l+1) P_l(\cos\beta)(4\rho\rho'/3)^{-1/2} I_{l+\frac{1}{2}}(4\rho\rho'/3), \tag{B13}$$

where β is the angle between $\boldsymbol{\rho}$ and $\boldsymbol{\rho}'$. The use of (B11) and (B12) with $x = \rho^2, y = \rho'^2, t = 1/4, \alpha = l + \frac{1}{2}$ and Eqs. (4.14) and (4.16) which define the $\chi_{l,m,n}$ in (B13) produces Eq. (B9).

It is convenient to define the raising and lowering operators $A_\pm$ by

$$A_\pm(1;\mathbf{r}) = -(1/\sqrt{2})[x + iy \mp (\partial/\partial x + i\partial/\partial y)] , \tag{B14}$$

$$A_\pm(0;\mathbf{r}) = z \mp \partial/\partial z, \tag{B15}$$

$$A_\pm(-1;\mathbf{r}) = (1/\sqrt{2})[x - iy \mp (\partial/\partial x - i\partial/\partial y)] . \tag{B16}$$

The $A_\pm$ are the spherical components of a vector operator. It can be shown (most easily with the aid of the Wigner–Eckart theorem) that

$$A_+(m_1;\mathbf{r})\chi_{l,m,n}(\mathbf{r}) = 2(-1)^{l+m+m_1+1}\left\{[(l+1)(2n+2l+3)/2]^{1/2}\begin{pmatrix} 1 & (l+1) & l \\ -m_1 & (m+m_1) & -m \end{pmatrix}\chi_{l+1,m+m_1,n}(\mathbf{r})\right.$$
$$\left. + [l(n+1)]^{1/2}\begin{pmatrix} l & (l-1) & l \\ -m_1 & (m+m_1) & -m \end{pmatrix}\chi_{l-1,m+m_1,n+1}(\mathbf{r})\right\} \tag{B17}$$

and

$$A_-(m_1;\mathbf{r})\chi_{l,m,n}(\mathbf{r}) = 2(-1)^{l+m+m_1}\left\{[(l+1)n]^{1/2}\begin{pmatrix} 1 & (l+1) & l \\ -m_1 & (m+m_1) & -m \end{pmatrix}\chi_{l+1,m+m_1,n-1}(\mathbf{r})\right.$$
$$\left. + [l(2n+2l+1)/2]^{1/2}\begin{pmatrix} 1 & (l-1) & l \\ -m_1 & (m+m_1) & -m \end{pmatrix}\chi_{l-1,m+m_1,n}(\mathbf{r})\right\}. \tag{B18}$$

It can also be shown that

$$A_+(m_1,\mathbf{u}) = (1/\sqrt{3})\left[2A_+(m_1,\boldsymbol{\rho}) + (-1)^{m_1}\overline{A_-(-m_1,\boldsymbol{\rho}')}\right] \tag{B19}$$

and

$$\overline{A_+(m_1,\mathbf{v})} = (1/\sqrt{3})\left[(-1)^{m_1}A_1(-m_1,\boldsymbol{\rho}) + 2\overline{A_+(m_1,\boldsymbol{\rho}')}\right] . \tag{B20}$$

Applying $\overline{A_+(m_1,\mathbf{v})}$ to Eq. (B9) and evaluating the result with the aid of Eqs. (B17)–(B20) yields the expansion theorem

$$\chi_{0,0,0}(\mathbf{u})\overline{\chi_{1,m_1,0}(\mathbf{v})} = (3/4)^{3/2}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sum_{n=0}^{\infty}(-1)^{m+m_1+1}2^{-2n-l}$$
$$\times\chi_{l,m,n}(\boldsymbol{\rho})\left\{[3(l+1)(2n+2l+3)/4]^{1/2}\begin{pmatrix} 1 & (l+1) & l \\ -m_1 & (m+m_1) & -m \end{pmatrix}\overline{\chi_{l+1,m+m_1,n}(\boldsymbol{\rho}')}\right.$$

$$+\,[3l(n+1)/2]^{1/2}\begin{pmatrix}1 & (l-1) & l\\ -m_1 & (m+m_1) & -m\end{pmatrix}\overline{\chi_{l-1,m+m_1,n+1}(\boldsymbol{\rho}')}\Bigg\}. \tag{B21}$$

Applying $\overline{(m_1',u)}$ to Eq. (B21) and rearranging the result with the aid of Eqs. (B17)–(B20) and the formula

$$\begin{pmatrix}1 & (l+\mu) & l\\ -m_1' & (m-m_1) & (-m+m_1+m_1')\end{pmatrix}\begin{pmatrix}1 & (l+\nu) & l\\ -m_1 & (m-m_1') & (-m+m_1+m_1')\end{pmatrix}$$
$$=(-1)^{m_1+m_1'}\sum_j(2j+1)\begin{Bmatrix}1 & (l+\mu) & l\\ 1 & (l+\nu) & j\end{Bmatrix}\begin{pmatrix}1 & (l+\mu) & j\\ m_1 & (m-m_1) & -m\end{pmatrix}\begin{pmatrix}1 & (l+\nu) & j\\ m_1' & (m-m_1') & -m\end{pmatrix} \tag{B22}$$

yields the expansion theorem

$$\chi_{1,m_1',0}(\mathbf{u})\overline{\chi_{1,m_1,0}(\mathbf{v})}$$
$$=(3/4)^{3/2}\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\sum_{n=0}^{\infty}(-2)^{2n-l+1}\Bigg\{-(2l+1)\begin{pmatrix}1 & l & l\\ m_1 & (m-m_1) & -m\end{pmatrix}\begin{pmatrix}1 & l & l\\ m_1' & (m-m_1') & -m\end{pmatrix}$$
$$\times\chi_{l,m-m_1,n}(\boldsymbol{\rho})\chi_{l,m-m_1',n}(\boldsymbol{\rho}')+2(2l+1)\begin{pmatrix}1 & (l+1) & l\\ m_1 & (m-m_1) & -m\end{pmatrix}\begin{pmatrix}1 & (l+1) & l\\ m_1 & (m-m_1') & -m\end{pmatrix}$$
$$\times\chi_{l+1,m-m_1,n-1}(\boldsymbol{\rho})\chi_{l+1,m-m_1',n-1}(\boldsymbol{\rho}')$$
$$+\begin{pmatrix}1 & (l-1) & l\\ m_1 & (m-m_1) & -m\end{pmatrix}\begin{pmatrix}1 & (l-1) & l\\ m_1' & (m-m_1') & -m\end{pmatrix}\chi_{l-1,m-m_1,m}(\boldsymbol{\rho})\chi_{l-1,m-m_1',n}(\boldsymbol{\rho}')$$
$$-3\Bigg[(n(l+1))^{1/2}\begin{pmatrix}1 & (l+1) & l\\ m_1 & (m-m_1) & -m\end{pmatrix}\chi_{l+1,m-m_1,n-1}(\boldsymbol{\rho})$$
$$+(l(n+l+\tfrac{1}{2}))^{1/2}\begin{pmatrix}1 & (l-1) & l\\ m_1 & (m-m_1) & -m\end{pmatrix}\chi_{l-1,m-m_1,n}(\boldsymbol{\rho})\Bigg]\Bigg[(n(l+1))^{1/2}\begin{pmatrix}1 & (l+1) & l\\ m_1' & (m-m_1') & -m\end{pmatrix}$$
$$\times\overline{\chi_{l+1,m-m_1',n-1}(\boldsymbol{\rho}')}+(l(n+l+\tfrac{1}{2}))^{1/2}\begin{pmatrix}1 & (l-1) & l\\ m_1' & (m-m_1') & -m\end{pmatrix}\overline{\chi_{l-1,m-m_1',n}(\boldsymbol{\rho}')}\Bigg]\Bigg\}. \tag{B23}$$

The formula (B22) can be obtained by rewriting formula (6.2.6) on p. 95 of Edmonds[33] in terms of the Wigner 3-j symbol. The needed values of the Wigner 6-j symbol which appears in Eq. (B22) can be obtained from Table V in Edmonds[33] with the aid of the symmetry properties of the 6-j symbol.

[1]Triton binding energy calculations in nuclear physics via the Faddeev equations have been discussed recently by R.A. Brandenburg, Y.E. Kim, and A. Tubis, Phys. Rev. C **12**, 1368 (1975); by E. Harper, Phys. Rev. Lett. **34**, 677 (1975); by O. Zohni, Phys. Rev. C **8**, 1164 (1973); by Michael I. Haftel, Phys. Rev. C 7, 80 (1973); and by C. Gignoux and A. Laverne, Phys. Rev. Lett. **29**, 436 (1972). A closely related approach has been used by Dean W. Halderson and Paul Goldhammer, Phys. Rev. C **11**, 237 (1975). Three-body binding energies in molecular physics have been calculated via the Faddeev equations by H. Krüger, Phys. Rev. Lett. **28**, 71 (1972). Krüger's results are, however, suspect because they violate the Hall–Post–Stenschke lower bound for some values of the parameters in the Morse potential which Krüger used, as was pointed out by T.K. Lim and Michael A. Zuniga, Ref. 2, this paper.

[2]Triton binding energy calculations in nuclear physics via the Rayleigh–Ritz variational method have been reported recently by M.R Strayer and P.U. Sauer, Nucl. Phys. A **231**, 1 (1974); by V.F. Demin, Y.E. Pokrovsky, and V.D. Efros, Phys. Lett. B **44**, 227 (1973); and by J. Bruinsma and R. van Wageningen, Phys. Lett. B **44**, 221 (1973). Variational calculations of three-body binding energies in molecular physics have been reported by T.K. Lim and Michael A. Zuniga, J. Chem. Phys. **63**, 2245 (1975); by L.W. Bruch and I.J. McGee, J. Chem. Phys. **59**, 409 (1973); and by S. Bosanac and J.N. Murrell, Mol. Phys. **26**, 349 (1973). The article of Lim and Zuniga contains a short critical review of earlier work.

[3]A useful general reference is L.M. Delves and J. Walsh, *Numerical Solution of Integral Equations* (Oxford University, New York, 1974). Error bounds to eigenvalues of integral equations are discussed by I.P. Mysovskih, AMS Trans. **35**, 237 (1964).

[4]S.H. Gould, *Variational Methods for Eigenvalue Problems: An Introduction to the Weinstein Method of Intermediate Problems, Second Edition*, (University of Toronto, Toronto, 1966).

[5]A. Weinstein and W. Stenger, *Methods of Intermediate Problems for Eigenvalues* (Academic, New York, 1972).

[6]H. F. Weinberger, *Variational Methods for Eigenvalue Approximation*, SIAM Regional Conference Series in Applied Mathematics (1974).

[7]N.W. Bazley, Phys. Rev. **120**, 144 (1960). See also Refs. 4–6 above.

[8]N.W. Bazley and D.W. Fox, Phys. Rev. **124**, 483 (1961). The difficulties of principle in extending the Bazley–Fox method to lithium are resolved in David W. Fox, SIAM (Soc. Ind. Appl. Math.) J. Math. Anal. **3**, 617

(1972); D.W. Fox and V.G. Sigillito, Chem. Phys. Lett. **13**, 85 (1972); **14**, 583 (1972); and David W. Fox and Vincent G. Sigillito, J. Appl. Math. Phys. **23**, 392 (1972). See also C.E. Reid, Int. J. Quantum Chem. **6**, 793 (1972). A proof that the H^- ion has only one bound state has been given via a generalization of Bazley's method by R.N. Hill, Phys. Rev. Lett. **38**, 643 (1977); J. Math. Phys. **18**, 2316 (1977).
[9]H.R. Post, Proc. Phys. Soc. London. Sect. A **69**, 936 (1956).
[10]R.L. Hall and H.R. Post, Proc. Phys. Soc. London **90**, 381 (1967).
[11]H. Stenschke, Diplomarbeit, Karlsruhe, Germany, 1966 (unpublished); J. Chem. Phys. **53**, 466 (1970).
[12]R.L. Hall, Proc. Phys. Soc. London **91**, 16 (1967).
[13]R.J.M. Carr and H.R. Post, J. Phys. A **1**, 596 (1968); **4**, 665 (1971).
[14]R.L. Hall, J. Phys. A **1**, 468 (1968).
[15]R.L. Hall, Phys. Lett. B **30**, 320 (1969).
[16]F. Calogero and C. Marchioro, J. Math. Phys. **10**, 562 (1969). See also J. Math. Phys. **15**, 139 (1974).
[17]O. Ya. Savchenko, Sov. J. Nucl. Phys. **8**, 728 (1969).
[18]H.R. Post, Proc. Phys. Soc. London **79**, 819 (1962).
[19]R.L. Hall, Proc. Phys. Soc. London **91**, 787 (1967).
[20]J.W. Humberston, R.L. Hall, and T.A. Osborne, Phys. Lett. B **27**, 195 (1968).
[21]T. Brady, E. Harms, L. Laroze, and J.S. Levinger, Nucl. Phys. A **168**, 509 (1971).
[22]G. Temple, Proc. R. Soc. (London) A **119**, 276 (1928). See also Weinstein and Stenger (Ref. 5), pp. 103–104 and Weinberger (Ref. 6), pp. 113–115.
[23]L.M. Delves, J.M. Blatt, C. Pask, and B. Davies, Phys. Lett. B **28**, 472 (1969); B. Davies, Nucl. Phys. A **103**, 165 (1967); Y.C. Tang, R.C. Herndon, and E.W. Schmid, Phys. Rev. **134**, B743 (1964).
[24]Essentially, this theorem appears as Theorem 1 on p. 21 of Weinstein and Stenger (Ref. 5), as Theorem 9 on p. 62 of Weinberger (Ref. 6), and is implicit in the discussion of intermediate problems of the second type in Gould (Ref. 4).
[25]There are actually two distinct minimax principles, one going back to H. Weyl and R. Courant, the other due to Poincaré. See Weinstein and Stenger (Ref. 5), Chap. 3 for a careful discussion of the distinction and of the relation between the two principles. See also Weinberger, Ref. 6, pp. 55–57.
[26]A.G. Sigalov, Usp. Mat. Nauk **22**, 3 (1967) [Russ. Math. Surveys **22**, 1 (1967)]; T. Kato, Suppl. Prog. Theor. Phys. **40**, 3 (1967); B. Simon, *Quantum Mechanics for Hamiltonians Defined as Quadratic Forms* (Princeton University, Princeton, 1971); B. Simon, in *Studies in Mathematical Physics, Essays in Honor of Valentine Bargmann* (Princeton University, Princeton, 1976), pp. 305–326; K. Jörgens and J. Weidmann, *Special Properties of Hamiltonian Operators*, Lecture Notes in Mathematics (Springer, Berlin, 1973), Vol. 313; W. Hunziker, Acta Phys. Austr. Suppl. **17**, 43 (1977).
[27]H.P. McKean, J. Math. Phys. **18**, 1277 (1977).
[28]H.F. Weinberger, A. Theory of Lower Bounds for Eigenvalues, Technical Note BN-183, IFDAM, University of Maryland, College Park, Md. 1959.
[29]N.W. Bazley and D.W. Fox, Ref. 8. See also Refs. 4–6.
[30]K.M. Watson, J. Nuttall, and J.S.R. Chisholm, *Topics in Several Particle Dynamics* (Holden-Day, San Francisco, 1967), Chap. 5; C. Lovelace, Phys. Rev. Sect. B **135**, 1225 (1964); A.N. Mitra, Nucl. Phys. **32**, 529 (1962); A.N. Mitra and V.S. Bhasin, Phys. Rev. **131**, 1265 (1963).
[31]With L bounded, as is the case here, this result is known as Weyl's theorem. See F. Riesz and B. Sz-Nagy, *Functional Analysis* (Ungar, New York, 1955), p. 367.
[32]J.H. Hetherington, J. Math. Phys. **11**, 1829 (1970).
[33]A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University, Princeton, 1968), p. 45.
[34]N.W. Bazley and D.W. Fox, Ref. 8, Sec. VII.C.
[35]Wilhelm Magnus, Fritz Oberhettinger, and Raj Pal Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics, Third Edition* (Springer, New York, 1966), p. 243.
[36]Magnus *et al.*, Ref. 35, p. 252; A. Erdelyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi, *Higher Transcendental Functions Vol. II* (McGraw-Hill, New York, 1953), p. 194.
[37]Magnus *et al.*, Ref. 35, p. 108.
[38]Magnus *et al.*, Ref. 35, p. 242; Erdelyi *et al.*, Ref. 36, p. 189.

An exactly solvable one dimensional three-body problem with hard cores

Robert Nyden Hill
Physics Department, University of Delaware, Newark, Delaware 19711

(Received 6 August 1979; accepted for publication 14 December 1979)

Three identical particles in one dimension interact via a potential which is infinite whenever one or more of the interparticle separations is less than a or greater than b, and zero when all interparticle separations lie between a and b. Their Schrödinger equation is solved by reducing it to the exactly solvable problem of the two dimensional Helmholtz equation inside an equilateral triangle.

I. INTRODUCTION AND SUMMARY OF RESULTS

Exactly solvable model problems provide useful testing grounds for approximation methods. This paper presents an exactly solvable one dimensional quantum mechanical three-body problem with hard cores. Stated explicitly, the interparticle potential is

$$V(x)=\begin{cases}\infty, & |x|<a \text{ or } |x|>b,\\ 0, & a<|x|<b,\end{cases} \tag{1}$$

where x is an interparticle coordinate. This model was invented as a testing ground for a method of improving Hall–Post–Stenschke (HPS) lower bounds to eigenvalues.[1] HPS lower bounds[2] are known to be poor for potentials with hard cores[3]; it was desired to see how well the improvements to HPS work for potentials with hard cores.

It is found that the Schrödinger Hamiltonian

$$H\equiv-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2}\right)+V(x_1-x_2)+V(x_2-x_3)+V(x_3-x_1), \tag{2}$$

has the spectrum

$$E=E_{\text{c.m.}}+E_{\text{int}}, \tag{3}$$

where $E_{\text{c.m.}}$ is the continuously variable center-of-mass energy and E_{int} is the internal energy, given by

$$E_{\text{int}}=4\pi^2\hbar^2(k^2+kl+l^2)/[3m(b-2a)^2], \tag{4}$$

with k and l integers subject to the restriction that neither k, l, or $k+l$ can be zero. The ground state energy is

$$E_0=4\pi^2\hbar^2/[m(b-2a)^2]. \tag{5}$$

The HPS lower bound[2] to the ground state energy is easily calculated; it is

$$E_{\text{HPS}}=(2\pi^2\hbar^2)/[m(b-a)^2]. \tag{6}$$

The lower bound (6) exhibits the poor quality of the HPS lower bound for potentials with hard cores: it differs from the exact energy (5) by a factor of 2 in the most favorable case, and gets worse as the ratio of a to b increases.

II. DERIVATION OF RESULTS

The center of mass can be separated off by introducing the Jacobi coordinates

$$X\equiv(x_1+x_2+x_3)/3, \tag{7}$$

$$x_{1,2}\equiv x_1-x_2, \tag{8}$$

$$\xi_3\equiv(2x_3-x_1-x_2)/\sqrt{3}. \tag{9}$$

The coordinates

$$x_{2,3}\equiv x_2-x_3=-(x_{1,2}+\sqrt{3}\xi_3)/2, \tag{10}$$

$$x_{3,1}\equiv x_3-x_1=-(x_{1,2}-\sqrt{3}\xi_3)/2 \tag{11}$$

will also be useful. The Hamiltonian (2) can be brought to the form

$$H=-[\hbar^2/(6m)]\partial^2/\partial X^2+H_{\text{int}}, \tag{12}$$

where the internal Hamiltonian H_{int} is

$$H_{\text{int}}=-\frac{\hbar^2}{m}\left(\frac{\partial^2}{\partial x_{1,2}^2}+\frac{\partial^2}{\partial \xi_3^2}\right)+U(x_{1,2},\xi_3), \tag{13}$$

with

$$U(x_{1,2},\xi_3)\equiv V(x_{1,2})+V(x_{2,3}(x_{1,2},\xi_3))+V(x_{3,1}(x_{1,2},\xi_3)). \tag{14}$$

The regions in which U is zero are equilateral triangles in the $x_{1,2}-\xi_3$ plane, as shown in Fig. 1. Thus, the eigenvalues of $mH_{\text{int}}/\hbar^2$ are the same as the eigenvalues of the interior problem for the two dimensional Helmholtz equation for the equilateral triangle. The exact solvability of the interior problem for the equilateral triangle seems to have been first noted by S.A. Schelkunoff,[4] who states the results without derivation. Because a derivation appears to be unavailable in the literature, one is sketched below.

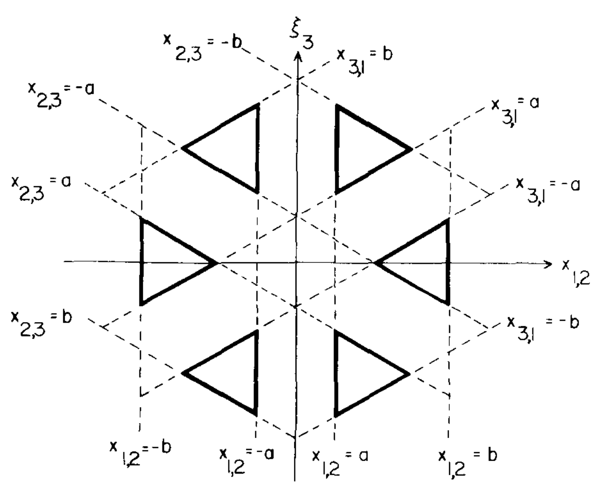

FIG. 1. Regions in the $x_{1,2}$–ξ_3 plane in which U is zero.

0022-2488/80/051083-03$01.00

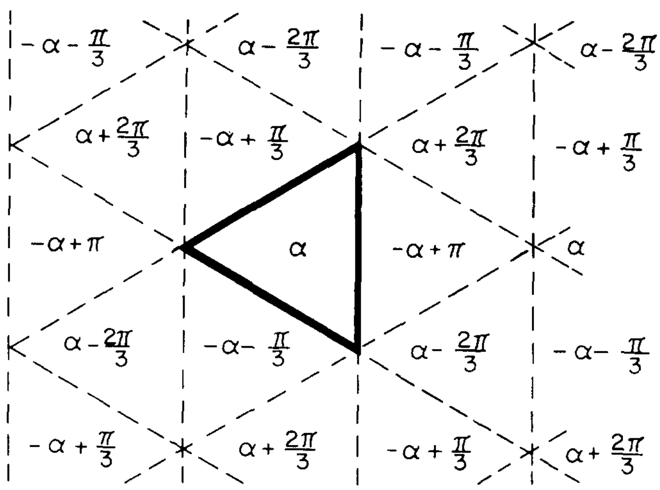

FIG. 2. Reflecting the box instead of the wave.

The derivation is based on de Broglie's old idea that an eigenfunction is a wave which interferes constructively with itself. One starts with a plane wave inside the triangular box and demands that, after several reflections off the sides of the box, a (multiply) reflected wave be obtained which interferes constructively with the original wave. Start with a plane wave whose direction of travel makes an angle α with the horizontal. Repeated reflections off the walls of the box then produce additional waves whose directions of travel make angles $-\alpha+\pi$, $-\alpha\pm\pi/3$, and $\alpha\pm 2\pi/3$ with the horizontal, but no others. These waves can be easily analyzed by the trick of reflecting the box instead of the wave as shown in Fig. 2, where each triangle is labeled by the direction of travel of the corresponding wave. The condition that the waves interfere constructively with themselves is equivalent to the condition that the waves have the periodicity of the hexagonal lattice in Fig. 2. A plane wave

$$\phi(x_{1,2},\xi_3)=\exp[i(K_{1,2}x_{1,2}+K_3\xi_3)], \qquad (15)$$

whose direction of travel makes an angle $\alpha=\tan^{-1}(K_3/K_{1,2})$ with the horizontal, will have this periodicity if and only if

$$K_{1,2}=\pi(k+l)/(b-2a), \qquad (16)$$

and

$$K_3=\pi(k-l)/[\sqrt{3}(b-2a)], \qquad (17)$$

where k and l are integers. Given the conditions (16) and (17) on $K_{1,2}$ and K_3, a superposition of waves traveling in each of the six directions and having the same value of $K=(K_{1,2}^2+K_3^2)^{1/2}$ is

$$\psi(x_{1,2},\xi_3)=\sum_{i=1}^{6}c_i\exp\{i\pi[\alpha_i x_{1,2}+(\beta_i\xi_3/\sqrt{3})]/(b-2a)\}, \qquad (18)$$

where $\alpha_1=-\alpha_4=k+l$, $\alpha_2=-\alpha_3=l$, $\alpha_5=-\alpha_6=-k$, $\beta_1=\beta_4=k-l$, $\beta_2=\beta_3=-2k-l$, and $\beta_5=\beta_6=k+2l$; the c_i are constants. A wavefunction ψ which satisfies the boundary condition $\psi=0$ on the sides of the triangle bounded by $x_{1,2}=b$, $x_{2,3}=-(x_{1,2}+\xi_3\sqrt{3})/2=-a$, and $x_{3,1}=-(x_{1,2}-\xi_3\sqrt{3})/2=-a$ can be obtained by imposing conditions on the c_i. The condition $\psi=0$ when $x_{1,2}=b$ is satisfied if

$$c_4/c_1=-\exp[2\pi i(k+l)b/(b-2a)], \qquad (19a)$$

$$c_3/c_2=-\exp[2\pi ilb/(b-2a)], \qquad (19b)$$

$$c_5/c_6=-\exp[2\pi ikb/(b-2a)]. \qquad (19c)$$

The condition $\psi=0$ when $x_{2,3}=-(x_{1,2}+\sqrt{3}\xi_3)/2=-a$ is satisfied if

$$c_2/c_1=-\exp[2\pi ika/(b-2a)], \qquad (20a)$$

$$c_3/c_6=-\exp[2\pi i(k+l)a/(b-2a)], \qquad (20b)$$

$$c_4/c_5=-\exp[2\pi ila/(b-2a)]. \qquad (20c)$$

The condition $\psi=0$ when $x_{3,1}=-(x_{1,2}-\sqrt{3}\xi_3)/2=-a$ is satisfied if

$$c_6/c_1=-\exp[2\pi ila/(b-2a)], \qquad (21a)$$

$$c_5/c_2=-\exp[2\pi i(k+l)a/(b-2a)], \qquad (21b)$$

$$c_4/c_3=-\exp[2\pi ika/(b-2a)]. \qquad (21c)$$

Equations (19) – (21) are just the conditions on the relative phases of incident and reflected waves. Their solution is

$$c_2=-c_1\exp[2\pi ika/(b-2a)], \qquad (22a)$$

$$c_3=c_1\exp[2\pi i(k+2l)a/(b-2a)], \qquad (22b)$$

$$c_4=-c_1\exp[4\pi i(k+l)a/(b-2a)], \qquad (22c)$$

$$c_5=c_1\exp[2\pi i(2k+l)a/(b-2a)], \qquad (22d)$$

$$c_6=-c_1\exp[2\pi ila/(b-2a)]. \qquad (22e)$$

The magnitude $|c_1|$ can be fixed by the demand that ψ be normalized; the phase of c_1 is arbitrary.

The result (4) for the eigenvalues of the internal Hamiltonian follows from Eqs. (16) and (17) with the aid of $E_{\rm int}=\hbar^2K^2/m=\hbar^2(K_{1,2}^2+K_3^2)/m$. It should be noted that the pairs (k,l), $(-k,k+l)$, $(-k-l,k)$, $(-l,-k)$, $(l,-k,-l)$, and $(k+l,-l)$ all correspond to the same energy and the same ψ (within an overall phase factor). If k, or l, or $k+l$ is zero, ψ vanishes. The form of ψ within each of other five triangles in which U is zero can be obtained by a sequence of rotations through $\pi/3$ in the $x_{1,2}-\xi_3$ plane. Because permutation of the particles just interchanges these disjoint pieces of the configuration space, wave functions which transform according to the symmetric, antisymmetric, and mixed representations of the permutation group on three objects have the same eigenvalues $E_{\rm int}$.

ACKNOWLEDGMENTS

The author would like to thank Ralph Kleinman of the University of Delaware mathematics department for the remark that the interior problem for the two dimensional Helmholtz equation can be solved exactly for the equilateral triangle. The author is grateful to the University of Wisconsin Physics Department, and especially to Lou Bruch, for hospitality during the summer of 1975 when this exactly solvable model was discovered. Support during that summer from a General Electric Company grant to the Wisconsin Physics Department was greatly appreciated.

[1]R.N. Hill, J. Math. Phys. **21**, 1070 (1980).
[2]H.R. Post, Proc. Phys. Soc. London Sect. A **69**, 936 (1956); R.L. Hall and

H.R. Post, Proc. Phys. Soc. London **90**, 381 (1967); H. Stenschke, Diplomarbeit, Karlsruhe, Germany, 1966 (unpublished); J. Chem. Phys. **53**, 466 (1970). References to additional work on lower bounds of this type are given in Ref. 1 above.
[3]T. Brady, E. Harms, L. Laroze, and J.S. Levinger, Nucl. Phys. A **168**, 509 (1971); T.K. Lim and Michael A. Zuniga, J. Chem. Phys. **63**, 2245 (1975).
[4]S.A. Schelkunoff, *Electromagnetic Waves* (Van Nostrand, New York, 1943), pp. 393–4.

Test of a method for finding lower bounds to eigenvalues of the three-body problem

James E. Bernard,[a] Robert Nyden Hill, and Maurice V. Barnhill III
Physics Department, University of Delaware, Newark, Delaware 19711

(Received 6 August 1979; accepted for publication 14 December 1979)

The simplest version of a method for systematically improving Hall–Post–Stenschke (HPS) lower bounds to eigenvalues is tested on an exactly soluble one dimensional three-body problem with hard cores. Significant improvement over the HPS bound is obtained, but considerable room for additional improvement remains.

I. INTRODUCTION

The preceding two papers[1,2] have presented a method for the systematic improvement of Hall–Post–Stenschke (HPS) lower bounds to eigenvalues and an exactly soluble three-body problem for testing the performance of the method with hard core potentials. The present paper tests the method on this model problem. The simplest version of the method, which is the only version tested here, leads to significant improvement over the HPS lower bound, but still leaves considerable room for additional improvement. Much better bounds should be obtainable with additional computing effort, but this has not been done.

Section II develops the integral equation which must be solved to get lower bounds for the model problem, gives the exact energies and HPS lower bounds for comparison with the energies to be obtained from the solution of the integral equation, and discusses the numerical solution of the integral equation. Section III presents a method, based on work of Mysovskih,[3] for rigorously bounding the truncation error incurred by the numerical solution procedure. Estimates of the roundoff error are also given. Section IV presents numerical results and discusses their significance.

II. APPLICATION OF THE METHOD TO THE MODEL PROBLEM

The general form of the identical particle integral equation in one dimension is [from Ref. 1, Eqs. (3.28)–(3.30)]

$$f_i(\rho) = \sum_{j=1}^{n-1} \alpha_{ij} \int_{-\infty}^{\infty} K_{ij}^{(X)}(\rho,\rho') f_j(\rho')\, d\rho', \qquad (2.1)$$

where

$$K_{ij}^{(X)}(\rho,\rho') = (b-a)\phi_i^{(X)}\left[\frac{1}{\sqrt{3}}(\rho+2\rho')\right]\phi_j^{(X)}\left[-\frac{1}{\sqrt{3}}(2\rho+\rho')\right], \qquad (2.2)$$

$$\alpha_{ij} = \frac{4}{\sqrt{3}(b-a)} \frac{(\epsilon_n-\epsilon_j)}{(E-2\epsilon_n-\epsilon_i)}, \qquad (2.3)$$

and X is S or A depending on whether the symmetric or antisymmetric sector of the problem is under consideration (see Sec. III.E of Ref. 1). Here ρ is the third Jacobi coordinate [Ref. 1, Eq. (2.4)], b and a are the well size and core size, respectively, E is the energy eigenvalue of the Schrödinger equation, and the $\phi_i^{(X)}$ and ϵ_i are, respectively, eigenfunctions and eigenvalues of the two-particle internal Hamiltonian. For the model problem that Hamiltonian is [Ref. 1, Eq. (2.10)]

$$H_{(2p)} = -\frac{2\hbar^2}{3m}\frac{\partial^2}{\partial r^2} + V(r), \qquad (2.4)$$

where the two-body potential V is 0 for $a<|r|<b$, and ∞ otherwise. Here r represents the interparticle separation. Normalized even eigenfunctions of $H_{(2p)}$ are

$$\phi_n^{(S)}(r) = \begin{cases} \dfrac{1}{\sqrt{b-a}}\sin\left[\dfrac{n\pi(r-a)}{(b-a)}\right], & a\leqslant r\leqslant b, \\ \dfrac{1}{\sqrt{b-a}}\sin\left[\dfrac{n\pi(r+a)}{(b-a)}\right], & -b\leqslant r\leqslant -a, \\ 0, \quad \text{otherwise.} \end{cases} \qquad (2.5)$$

The odd functions are the same except that the sign of the

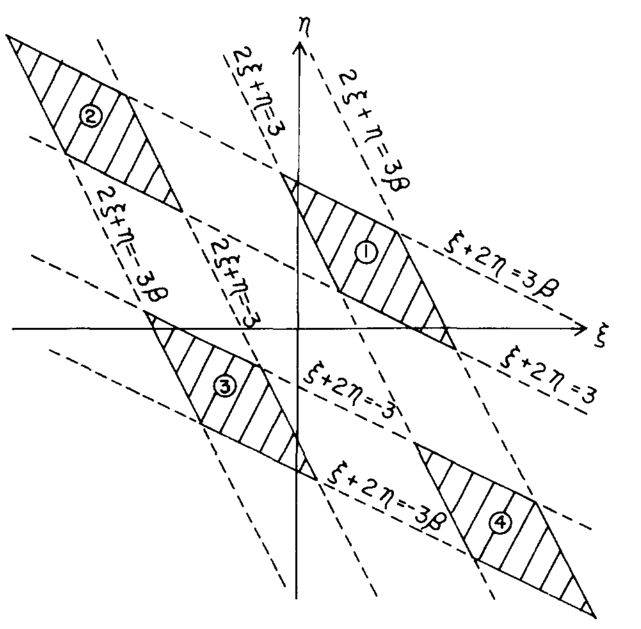

FIG. 1. The kernel is nonvanishing only within the shaded regions. The dimensionless variables used in the figure are defined in Eq. (2.7).

[a]Work done in partial fulfillment of the requirements for the M.S. degree.

 0022-2488/80/051086-06$01.00

second functional form is changed. The eigenvalues are given by

$$\epsilon_n = \frac{2\hbar^2 n^2 \pi^2}{3m(b-a)^2} \tag{2.6}$$

and are the same for both even and odd eigenfunctions with the same index n. It is evident from Eq. (2.5) that the kernel $K_{ij}^{(X)}$ is nonzero only when the absolute values of the arguments of both $\phi_i^{(X)}$ and $\phi_j^{(X)}$ lie between a and b. The four regions in which $K_{ij}^{(X)}(\rho,\rho')$ is nonzero are depicted in Fig. 1 for the case where $b/a = 2.5$.

The following change of variables brings Eqs. (2.1)–(2.3) to dimensionless form: Define

$$\xi = \sqrt{3}\rho/a,$$
$$\eta = \sqrt{3}\rho'/a, \quad \beta = b/a, \quad \gamma = 1/(\beta - 1), \tag{2.7}$$

and

$$\lambda_{ij} = \sqrt{3}/(\alpha_{ij} a).$$

Then the explicit functional forms of the kernel in the four regions are

$$\begin{aligned} K_{ij}^{(X,1)} &= -S^{(X)} \sin\left[\frac{i\gamma\pi}{3}(\xi + 2\eta - 3)\right] \times \sin\left[\frac{j\gamma\pi}{3}(2\xi + \eta - 3)\right], \\ K_{ij}^{(X,2)} &= \sin\left[\frac{i\gamma\pi}{3}(\xi + 2\eta - 3)\right] \sin\left[\frac{j\gamma\pi}{3}(2\xi + \eta + 3)\right], \\ K_{ij}^{(X,3)} &= -S^{(X)} \sin\left[\frac{i\gamma\pi}{3}(\xi + 2\eta + 3)\right] \times \sin\left[\frac{j\gamma\pi}{3}(2\xi + \eta + 3)\right], \\ K_{ij}^{(X,4)} &= \sin\left[\frac{i\gamma\pi}{3}(\xi + 2\eta + 3)\right] \sin\left[\frac{j\gamma\pi}{3}(2\xi + \eta - 3)\right], \end{aligned} \tag{2.8}$$

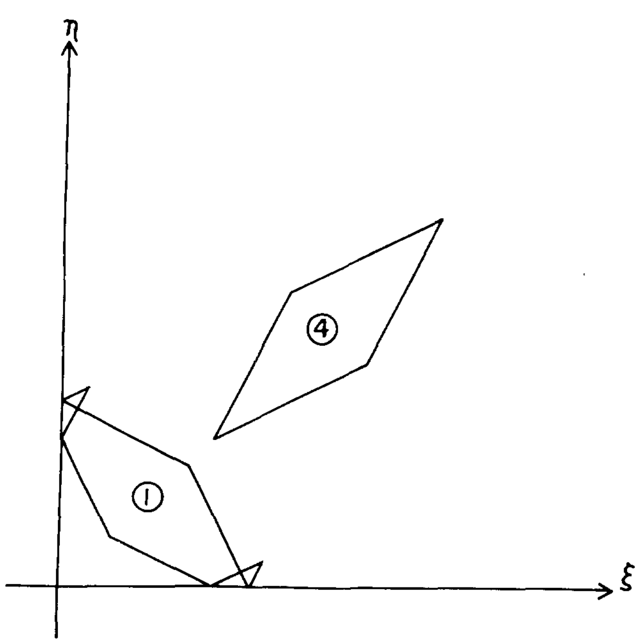

FIG. 2. The kernel $K_{i,j}^{(X,P)}$ is nonzero within the regions shown above and is zero outside these regions.

where

$$S^{(X)} = \begin{cases} +1, & \text{if } X = S \quad \text{(symmetric problem)}, \\ -1, & \text{if } X = A \quad \text{(antisymmetric problem)}. \end{cases} \tag{2.9}$$

The integral equation can now be written

$$\hat{f}_i(\xi) = \sum_{j=1}^{n-1} \frac{1}{\lambda_{ij}} \int_{-\infty}^{\infty} \hat{K}_{ij}^{(X)}(\xi,\eta)\hat{f}_j(\eta)\, d\eta, \tag{2.10}$$

where

$$\frac{1}{\lambda_{ij}} = \frac{4\gamma(\epsilon_n - \epsilon_j)}{3(E - 2\epsilon_n - \epsilon_i)}. \tag{2.11}$$

The symmetries

$$\begin{aligned} \hat{K}_{ij}^{(X,1)}(-\xi,-\eta) &= \hat{K}_{ij}^{(X,3)}(\xi,\eta), \\ \hat{K}_{ij}^{(X,2)}(-\xi,-\eta) &= \hat{K}_{ij}^{(X,4)}(\xi,\eta), \\ \hat{K}_{ij}^{(X,3)}(-\xi,-\eta) &= \hat{K}_{ij}^{(X,1)}(\xi,\eta), \\ \hat{K}_{ij}^{(X,4)}(-\xi,-\eta) &= \hat{K}_{ij}^{(X,2)}(\xi,\eta) \end{aligned} \tag{2.12}$$

of the kernel imply that the eigenfunctions can all be classified as either even or odd. Define

$$f_i^{(e)}(\xi) = \tfrac{1}{2}[\hat{f}_i(\xi) + \hat{f}_i(-\xi)],$$

and (2.13)

$$f_i^{(o)}(\xi) = \tfrac{1}{2}[\hat{f}_i(\xi) - \hat{f}_i(-\xi)] \quad (i = 1,\ldots,n-1),$$

which will satisfy Eq. (2.10) whenever the set $\{\hat{f}_i(\xi)\}$ does. Thus, only the right half-plane need be considered in solving the problem; the left half-plane can be filled in by symmetry. The integral equation can then be reduced to an equation in the first quadrant of the ξ–η plane:

$$f_i^{(P)}(\xi) = \sum_{j=1}^{n-1} \frac{1}{\lambda_{ij}} \int_0^{\infty} K_{ij}^{(X,P)}(\xi,\eta) f_j^{(P)}(\eta)\, d\eta, \tag{2.14}$$

where

$$K_{ij}^{(X,P)}(\xi,\eta) = \hat{K}_{ij}^{(X)}(\xi,\eta) \pm \hat{K}_{ij}^{(X)}(\xi,-\eta). \tag{2.15}$$

Here P is either e or o; the plus sign is associated with $f_j^{(e)}$ and the minus with $f_j^{(o)}$. The regions in which the kernel is nonzero will then consist of region 1 (truncated) and the reflection of region 4 together with reflected pieces of regions 1 and 3 (see Fig. 2). Thus, the upper limit on the integral in Eq. (2.14) can be taken to be 3β.

In the first approximation we keep only the first even (odd) term in the sum in Eq. (2.14) in order to get bounds on the eigenvalues of the symmetric (antisymmetric) problem. Equation (2.14) then becomes

$$\lambda f^{(P)}(\xi) = \int_0^{3\beta} K^{(X,P)}(\xi,\eta) f^{(P)}(\eta)\, d\eta. \tag{2.16}$$

From Eqs. (2.3), (2.6), and (2.7) with $i = j = 1$ and $n = 2$ it can be seen that the energy lower bound is related to the eigenvalue λ by

$$E_L = (4\gamma\lambda + 9)\frac{2\hbar^2\pi^2}{3m(b-a)^2}. \tag{2.17}$$

For comparison, we need the energy eigenvalues and the multiplicities of the exact problem, and the HPS lower bound to the ground state energy. The exact energies are given by

$$E = \frac{4\pi^2\hbar^2}{3m(b-2a)^2}(k^2+kl+l^2), \tag{2.18}$$

[Eq. (4) of Ref. 2]. The energies for which $k = l$ are nondegenerate, and those for which $k \neq l$ are doubly degenerate. This will be of interest later when we examine the performance of the method in finding lower bounds to excited states. The HPS bound $E_{\rm HPS}$ is just three times the lowest eigenvalues of $H_{(2P)}$, i.e., $E_{\rm HPS} = 3\epsilon_1$.

In order to facilitate comparison of the lower bounds with the exact ground state energy, we multiply all three by $3m(b-2a)^2/4\pi^2\hbar^2$ to obtain the dimensionless forms

$$\hat{E} = k^2 + kl + l^2, \tag{2.19a}$$

$$\hat{E}_L = \tfrac{1}{2}(1-\gamma)^2(4\gamma\lambda + 9), \tag{2.19b}$$

$$\hat{E}_{\rm HPS} = \tfrac{3}{2}(1-\gamma)^2 \tag{2.19c}$$

for the exact energy, the present lower bound, and the HPS lower bound, respectively.

The eigenvalue λ can be found by using a numerical integration rule to write Eq. (2.16) as a matrix equation and then diagonalizing the matrix. Written in the discrete form, Eq. (2.16) is

$$\lambda c_i f^{(P)}(\xi_i) = \sum_j c_i c_j K^{(X,P)}(\xi_i, \xi_j) c_j f^{(P)}(\xi_j). \tag{2.20}$$

Here the c_i^2 are the constants prescribed by the numerical integration rule. Now define

$$g_i^{(P)} = c_i f^{(P)}(\xi_i), \tag{2.21}$$

$$K_{ij}^{(X,P)} = c_i c_j K^{(X,P)}(\xi_i, \xi_j). \tag{2.22}$$

Then Eq. (2.20) becomes

$$\lambda g_i^{(P)} = \sum_j K_{ij}^{(X,P)} g_i^{(P)}. \tag{2.23}$$

The eigenvalues λ of Eq. (2.23) can then be used as approximations to the eigenvalues of Eq. (2.16), from which we obtain the lower bounds (2.19b). The eigenfunctions $g_i^{(P)}$ can be used to obtain approximations to the eigenfunctions of Eq. (2.16).

The numerical integration rule used was Simpson's rule. In order to satisfy the hypotheses used to derive the error formula for Simpson's rule it is necessary to adjust the parameters so that evaluation points are located at the vertices of the regions where the kernel is nonzero. This is due to the fact that the derivatives of the kernel are discontinuous at the boundary, and the fourth derivative of the kernel appears in the error formula. This must be kept in mind if rigorous error bounds are desired.

III. ERROR BOUNDS

If we are to maintain our claim that the lower bounds we obtained for the energies of the model problem are rigorous, it is necessary to bound the truncation and roundoff error introduced by the numerical solution procedure. Bounds on the truncation error were obtained via a method developed by Mysovskih[3]; roundoff error was estimated by using a difference table.

Mysovskih's method is based on a theorem of Weyl which gives estimates of the differences of the eigenvalues of two operators. For the case of interest here, where the operators are defined on different Hilbert spaces, it is necessary to use Weyl's theorem twice to get the bounds needed. We will not present a derivation of the error bounds here, but will merely state the theorem upon which the method is based and describe the way in which its use results in the truncation error bounds.

Theorem 3.1 (*Weyl's theorem*): Let B and $A = B + C$ be completely continuous, self-adjoint operators in a Hilbert space X. Let the negative eigenvalues of A be enumerated in nondecreasing order:

$$\lambda_1^{(A,-)} \leqslant \lambda_2^{(A,-)} \leqslant \cdots \leqslant 0, \tag{3.1}$$

and let the positive eigenvalues be enumerated in nonincreasing order:

$$\lambda_1^{(A,+)} \geqslant \lambda_2^{(A,+)} \geqslant \cdots \geqslant 0. \tag{3.2}$$

Here a multiple eigenvalue is repeated a number of times equal to its multiplicity. (Note that A completely continuous implies that zero is the only possible accumulation point of the spectrum.) Let the eigenvalues of B be enumerated in the same manner. Then

$$|\lambda_j^{(A,-)} - \lambda_j^{(B,-)}| \leqslant \|c\|, \tag{3.3}$$

and

$$|\lambda_j^{(A,+)} - \lambda_j^{(B,+)}| \leqslant \|c\|, \tag{3.4}$$

where

$$\|c\| = \sup_{\substack{x\in X \\ x\neq 0}} \frac{\|cx\|}{\|x\|}. \tag{3.5}$$

Application of the numerical integration rule to the second iterated kernel (which has as eigenvalues the squares of the eigenvalues of the kernel) results in

$$K_2(x,y) = B(x,y) + \epsilon(x,y), \tag{3.6}$$

where

$$B(x,y) = \sum_{k=1}^{n} c_k^2 K(x, y_k) K(y_k, y), \tag{3.7}$$

and $\epsilon(x, y)$ is the error term for the numerical integration rule. The c_k^2 are the constants prescribed by the numerical integration rule, and the y_k $(k = 1,\ldots,n)$ are the points at which the integrand is evaluated in computing the numerical approximation to the integral in $K_2(x, y)$. Both B and ϵ are L_2 kernels, so Weyl's theorem can be applied to obtain

$$|[\lambda_j^{(\pm)}]^2 - \tilde{\tilde{\lambda}}_j^{(\pm)}| \leqslant \|\epsilon\|, \tag{3.8}$$

where the $\tilde{\tilde{\lambda}}_j^{(\pm)}$ are eigenvalues of B. Now B is a degenerate kernel. Its eigenvalues can be shown to be the same as those of the matrix Γ whose elements are defined by

$$\Gamma_{ik} = c_i c_k \int_s^t K(y_i, y)\, K(y, y_k)\, dy. \tag{3.9}$$

It is also possible to show[3] that the matrix Γ is related to the matrix L which approximates the kernel (according to the numerical integration rule) by

$$\Gamma = L^2 + E, \tag{3.10}$$

where

$$E_{ik} = c_i c_k E(y_i, y_k), \tag{3.11}$$

and

$$L_{ik} = c_i c_k K(y_i, y_k). \tag{3.12}$$

Thus, we can apply Weyl's theorem again to obtain bounds on the differences between the eigenvalues of Γ and L^2:

$$|\tilde{\tilde{\lambda}}_j^{(\pm)} - [\tilde{\lambda}_j^{(\pm)}]^2| \leqslant \|E\|, \tag{3.13}$$

where the $\tilde{\lambda}_j^{(\pm)}$ (the eigenvalues of L) are the approximations to the exact eigenvalues $\lambda_j^{(\pm)}$.

Now combine the results of the two applications of Weyl's theorem. Some manipulation of the inequalities then yields

$$|\lambda_j^{(\pm)} - \tilde{\lambda}_j^{(\pm)}| \leqslant \frac{\|\epsilon\| + \|E\|}{|\tilde{\lambda}_j^{(\pm)}| + \{[\tilde{\lambda}_j^{(\pm)}]^2 - (\|\epsilon\| + \|E\|)\}^{1/2}}, \tag{3.14}$$

subject to the condition that

$$\|\epsilon\| + \|E\| \leqslant \tilde{\lambda}_j^{(\pm)}. \tag{3.15}$$

In order to obtain the inequality (3.14) it is necessary to assume that the signs of $\lambda_j^{(\pm)}$ and $\tilde{\lambda}_j^{(\pm)}$ are the same, since only their squares appear in Eqs. (3.8) and (3.13).

To implement this method for computing error bounds we must first obtain bounds on the norms $\|\epsilon\|$ and $\|E\|$ [see Eqs. (3.6) and (3.11)]. Since Simpson's rule was used to approximate the kernel of the integral equation (2.16), the $\epsilon(x, y)$ which appears in Eq. (3.6) is the Simpson's rule error term:

$$\epsilon(x, y) = -\frac{(t-s)h^4 f^{(4)}(y';x,y)}{180}, \tag{3.16}$$

for some $y' \in [s,t]$, where

$$f(y';x,y) = K(x,y')K(y',y), \tag{3.17}$$

h is the step size, and the integration limits are

$$s = 0, \quad t = 3\beta. \tag{3.18}$$

As an upper bound on the magnitude of $\epsilon(x, y)$ we use

TABLE I. Main results. Beta is the ratio of well size to core size, $K_{\rm MAX}$ is the size of the matrix kernel, the lower bounds are for the ground state, and the truncation error is given as a percent of the energy lower bound.

Beta	$K_{\rm MAX}$	Small step size	Large step size	Energy lower bound	Hall–Post Lower bound	Truncation error (%)
2.1	67	0.0500	0.1000	0.0249	0.0124	1.7×10^{-1}
2.2	67	0.1000	0.1000	0.0836	0.0417	1.2×10^{-1}
2.3	73	0.0500	0.1000	0.1603	0.0799	8.1×10^{-2}
2.4	73	0.1000	0.1000	0.2458	0.1224	5.8×10^{-2}
2.5	37	0.1250	0.2500	0.3346	0.1667	1.8
2.5	73	0.0625	0.1250	0.3345	0.1667	1.1×10^{-1}
2.5	61	0.1250	0.1250	0.3345	0.1667	1.1×10^{-1}
2.5	31	0.2500	0.2500	0.3345	0.1667	1.8
2.6	49	0.1000	0.2000	0.4234	0.2109	5.4×10^{-1}
2.7	79	0.1000	0.1500	0.5105	0.2543	1.3×10^{-1}
2.8	57	0.1000	0.2000	0.5947	0.2963	3.3×10^{-1}
2.9	77	0.1000	0.1500	0.6756	0.3366	8.1×10^{-2}
3.0	61	0.1000	0.2000	0.7527	0.3750	2.1×10^{-1}
3.1	77	0.1000	0.1500	0.8260	0.4116	5.4×10^{-2}
3.2	77	0.1000	0.1500	0.8957	0.4463	4.5×10^{-2}
3.4	75	0.1000	0.1500	1.0243	0.5104	3.1×10^{-2}
3.6	81	0.1000	0.2000	1.1398	0.5680	7.0×10^{-2}
3.8	85	0.1000	0.2000	1.2434	0.6199	5.2×10^{-2}
4.0	61	0.2000	0.2000	1.3366	0.6667	3.8×10^{-2}
4.2	69	0.1000	0.2000	1.4203	0.7090	2.9×10^{-2}
4.4	67	0.2000	0.2000	1.4959	0.7474	2.3×10^{-2}
4.6	79	0.1000	0.2000	1.5640	0.7824	1.8×10^{-2}
4.8	73	0.2000	0.2000	1.6256	0.8144	1.4×10^{-2}
5.0	61	0.2500	0.2500	1.6812	0.8438	2.8×10^{-2}
5.2	79	0.2000	0.2000	1.7319	0.8707	9.3×10^{-3}
5.4	85	0.1000	0.2000	1.7778	0.8957	7.7×10^{-3}
5.6	85	0.2000	0.2000	1.8197	0.9187	6.3×10^{-3}
5.8	91	0.1500	0.2000	1.8579	0.9401	5.7×10^{-3}
6.0	73	0.2500	0.2500	1.8926	0.9600	1.1×10^{-2}
6.5	79	0.2500	0.2500	1.9681	1.0041	7.3×10^{-3}
7.0	85	0.2500	0.2500	2.0300	1.0417	5.0×10^{-3}
7.5	67	0.2500	0.5000	2.0818	1.0740	5.9×10^{-2}
8.0	49	0.5000	0.5000	2.1254	1.1020	4.2×10^{-2}
8.5	65	0.2500	0.5000	2.1633	1.1267	3.4×10^{-2}
9.0	55	0.5000	0.5000	2.1958	1.1484	2.4×10^{-2}
9.5	71	0.2500	0.5000	2.2247	1.1678	2.0×10^{-2}
10.0	61	0.5000	0.5000	2.2500	1.1852	1.5×10^{-2}
15.0	91	0.5000	0.5000	2.4047	1.2934	2.6×10^{-3}
15.0	181	0.2500	0.2500	2.4047	1.2934	1.7×10^{-4}
20.0	121	0.5000	0.5000	2.4793	1.3463	7.6×10^{-4}

$$\hat{\epsilon}(x, y) = \frac{\beta h^4 \max_{y' \in [s,t]} |f^{(4)}(y';x, y)|}{60}. \quad (3.19)$$

Of course, in the computer program used to compute the error bounds the maximum can only be taken over a discrete set of values, but we assume that the result is quite close to the actual maximum since the function f is relatively smooth. The lower bounds program allows two different step sizes for different regions, but in order to limit complications the truncation error program uses the larger step size as h in Eq. (3.19). The norm of $\epsilon(x, y)$ is bounded by the Hilbert–Schmidt norm of the matrix representing ϵ:

$$\|\epsilon\| \leqslant \left[\int_s^t \int_s^t |\epsilon(x, y)|^2 \, dx \, dy \right]^{1/2} \approx \left[\sum_i \sum_j c_i^2 c_j^2 |\epsilon(y_i, y_j)|^2 \right]^{1/2}, \quad (3.20)$$

and the norm of the matrix E is bounded in similar fashion.

The roundoff error was estimated by constructing difference tables from the ground state energy lower bounds for three different sets of equally spaced values of β. From these tables we estimate that the roundoff error in the ground state energies is less than 2×10^{-4}. For a discussion of the use of difference tables to estimate roundoff error see Hamming.[4]

IV. NUMERICAL RESULTS AND DISCUSSION

This final section presents some of the more significant numerical results together with some discussion of the interpretation and implications of the numbers. In addition, a few methods for improvement of the results will be pointed out. Before listing our table of results, we will make some comments on the cases we chose to solve.

The lower bounds and truncation errors were computed for values of β ranging from 2.1 to 20. This should cover the spectrum of range-to-core radius ratios used in most realistic nuclear and molecular potentials as well as give a fairly complete picture of the dependence of the bounds on β. It was necessary to have the lower limit above 2.0 since the exact energy [Eq. (2.18)] diverges at that point.

For two values of β (15 and 2.5), two or four, respectively, different partitions of the interval of integration were tried in order to get a feeling for the effect on the eigenvalues and on the truncation error bounds. In addition, for 20 of the cases two different step sizes were used in the partition.

TABLE II. Excited state lower bounds and exact energies.

State \ Beta:	2.1	5	20	Exact
1	0.02488	1.6812	2.4793	3
2	0.03100	1.9253	3.0196	7
3	0.03100	2.1620	3.5912	7
4	0.03100	2.2140	3.8655	12
5	0.03392	2.3401	3.8728	13
6	0.03546	2.4772	3.9637	13
7	0.03546	2.4772	3.9922	19
8	0.03546	2.4991	4.0009	19
9	0.03627	2.5024	4.0010	21
10	0.03671	2.5098	4.0043	21

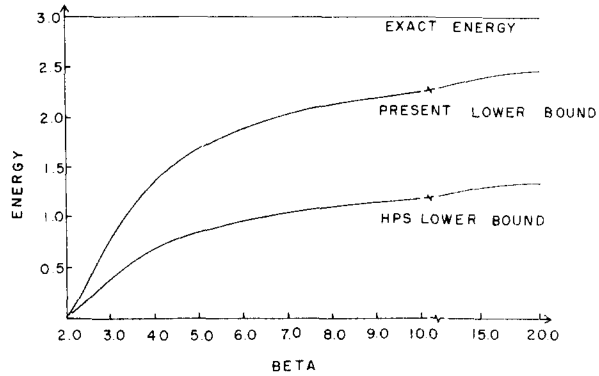

FIG. 3. A comparison of the present HPS lower bounds with the exact ground state energy. (The energy is given in dimensionless units.)

The most significant numerical results are presented in Table I. The first column gives the value of β for each case. K_{MAX} is the order of the approximating matrix and is a measure of the fineness of the mesh used for that case. The next two columns give the step sizes used in the partition. The lowest energy obtained from the eigenvalues of the matrix approximation to the kernel [via Eq. (2.19b)] follows in the next column. The HPS lower bounds are provided for comparison in the sixth column, and the last column contains the upper bounds to the truncation error for each case. The exact ground state energy is 3.

A graphical comparison of the ground state energy lower bounds, the HPS lower bounds, and the exact ground state energy is shown in Fig. 3 for β in the range from 2.1 to 20. The most striking feature is the fact that both bounds are quite poor near $\beta = 2$ and show a gradual improvement with increasing β. From Eq. (2.19c) we see that in the limit as $\beta \to \infty$ ($\gamma \to 0$) the HPS bound approaches a value of 1.5, which is just half of the exact energy. The present lower bound is somewhat better than this, but it is clear that there is still considerable room for improvement. The ratio of the energy lower bound to the HPS bound varies from about 2.0 at $\beta = 2.1$ to about 1.8 at $\beta = 20$. It is not clear whether this ratio will continue to decrease with increasing β, but it is a question of some interest as it suggests the possibility that the bounds eventually will begin to decrease with increasing β.

From the truncation error bounds given in Table I and the graph of the energy lower bounds versus β in Fig. 3 it is evident that the effect of truncation on the energy lower bounds is small compared with the effect due to the fact that only one term was kept in the sum in Eq. (2.14). Even in the worst case ($\beta = 2.5$ with $K_{\mathrm{MAX}} = 37$) the truncation error bound was only about 6.2% (or 0.7% of the exact energy), whereas the total error for that case was about 88.8% of the exact energy. In the case with the least total error ($\beta = 20$) that error was still 17.4% of the exact energy. Clearly, it would be interesting to see the computation done with more terms kept in the sum in Eq. (2.14) in order to see how rapidly the lower bound improves as additional terms are included. One might observe that there is a general trend for the truncation error bounds to decrease with increasing β. This is not just a result of giving the truncation error bounds as a percent of the energy lower bounds, but would be seen in a list of absolute truncation error bounds as well.

Now, we will look briefly at the performance of the method in computing lower bounds to excited states. As one might expect, the results of the method deteriorate as one proceeds to higher states since information about the higher states was discarded in truncating the two-particle Hamiltonians. This can be seen from Table II, which shows the first ten lower bounds and the first ten exact energies for three particular values of β.

Because the bounds resulting from this first order test of the method on a potential with an infinite hard core are only moderately better than those from the HPS method, it may be of interest to look into some of the options available for improvement. One option, which has already been indicated, is to keep more terms in the sum in Eq. (2.14). Such an approach would require the diagonalization of larger matrices, but since the truncation error was so small in the cases we investigated, it would be possible to limit that size by using a coarser mesh. Then, presumably, one could strike a balance such that the truncation error and the error due to the approximation in Eq. (2.14) are of similar magnitude.

A somewhat different approach would be to use the present method either to obtain rough bounds for use in Temple's formula,[5] or as the initial approximation in a recently developed alternative to Temple's formula.[6]

[1]R.N. Hill, J. Math. Phys. **21**, 1070 (1980).
[2]R.N. Hill, J. Math. Phys. **21**, 1083 (1980).
[3]I.P. Mysovskih, Am. Math. Soc. Trans. **35**, 237 (1964).
[4]R.W. Hamming, *Numerical Methods for Scientists and Engineers*, 2nd ed. (McGraw-Hill, New York, 1973), Chap. 10.
[5]G. Temple, Proc. R. Soc. (London) Ser. A **119**, 276 (1928).
[6]R.N. Hill, submitted to J. Math. Phys.

On the application of the Sommerfeld–Maluzhinetz transformation to some one-dimensional three-particle problems

K. Lipszyc
Institute of Physical Chemistry, Polish Academy of Sciences, Warsaw, ul. Kasprzaka 44/52, Poland

(Received 10 March 1978; accepted for publication 9 February 1979)

The reformulation of the one-dimensional three-body problems with boundary condition and delta function interactions, based on the Sommerfeld–Maluzhinetz transformation, is presented. The argumentation is carried out as exemplified by two models–the exactly soluble model of two identical particles interacting through delta potential and each of which interacts with a third one through boundary condition interactions, and a model of two identical particles and a fixed wall, all interactions being of the delta function type. The problems are reduced to those of solving coupled systems of functional equations for the Sommerfeld transforms of the wavefunction. The functional properties of the transforms are then used to derive expressions relating them to the half off-shell extensions of the elastic and exchange probability amplitudes as defined in the Faddeev–Lovelace approach.

I. INTRODUCTION

The number of analytically solvable three-body models is very scarce.[1-9] In most papers dealing with this subject use is made of the analogy between the one-dimensional three-particle Schrödinger equation in the center-of-mass system and the mathematical problems of diffraction of time harmonic waves by plane obstacles spreading out to infinity. In the last field, the most fruitful methods leading to closed form solution are found, namely the integral transform methods supplemented by function theoretic techniques.

The Fourier transformation accompanied by the Wiener–Hopf technique is of limited use in solving three-body problems. Lieb and Koppe[2] employed this method to solve a simple model of the breakup of a pair of particles with delta-function potential. A more complicated model with one two-body delta potential and one hard-core interaction is attempted by the above technique by Jost.[1] The problem is reduced to that of solving the difference equation later solved by Albeviero.[5]

The model of three colinear particles of arbitrary masses with interactions described by the boundary conditions

$$\frac{\partial \Psi}{\partial x_i} + \alpha_i \Psi = 0, \quad \text{for } x_i = 0, \tag{1}$$

where x_i is the distance between the particles j and k, is described by McGuire and Hurst[6] and independently by the author.[7,8] The problem is mathematically identical to the problem of diffraction of the plane or surface waves by an impedance wedge, this analogy originated with Nussenzweig.[3] In Ref. 6 Williams' method[10] is applied, whereas in Refs. 7,8 the author adopted directly the solution of Maluzhinetz.[11-13] In both cases, the solution is represented by integrals of the Sommerfeld type.

The more "realistic" model of three particles with delta function potentials is much more difficult to handle than the boundary condition model, the exception being the models of three identical particles that are solvable in both cases by elementary methods.[4,14] Gaudin and Derrida[9] have reduced the problem of the bound state of three particles with delta function potentials to a system of difference equations and solved it exactly for the nontrivial case of equal masses and two nonzero interactions of equal strength.

This paper treats the reformulation of the problems of three colinear particles with boundary condition and delta function interactions, based on the Sommerfeld–Maluzhinetz transformation. Some of Maluzhinetz results concerning the properties of the above transformation are reviewed in Sec. 2. In Sec. 3 we discuss briefly the model system of two identical particles interacting via the delta potential, each of which interacts with the third one via the impedance-type potential, Eq. (1). For that case, the system of functional equations to which the problem is reduced, with help of the Sommerfeld–Maluzhinetz transformation, decouples into the system of Maluzhinetz equations for the impedance wedge and is thus solvable by this method. In Sec. 4, we study the system of two identical particles and one infinitely heavy particle; all interactions being delta potentials. By the application of the symmetry of the system, a part of the wave function is extracted that cannot be calculated by elementary methods. The nontrivial subproblem is then reduced to two systems of functional equations for the Sommerfeld transforms.

Section 5 illustrates the significance of the Sommerfeld–Maluzhinetz transformation as applied to one-dimensional three-body problems with zero range interactions, irrespective to its eventual significance as the method for solving the Schroedinger equation exactly. As illustrated by the above mentioned models, it is shown here that strikingly simple relations exist between the Sommerfeld transforms of the wave function and the half-off-shell extensions of the probability amplitudes of reactions resulting from the scattering of a particle off a bound state. The properties of these last functions, which are usually studied on the ground of the Faddeev–Lovelace equations, can thus be deduced from the function-theoretic properties demanded for the Sommerfeld

 0022-2488/80/051092-11$01.00

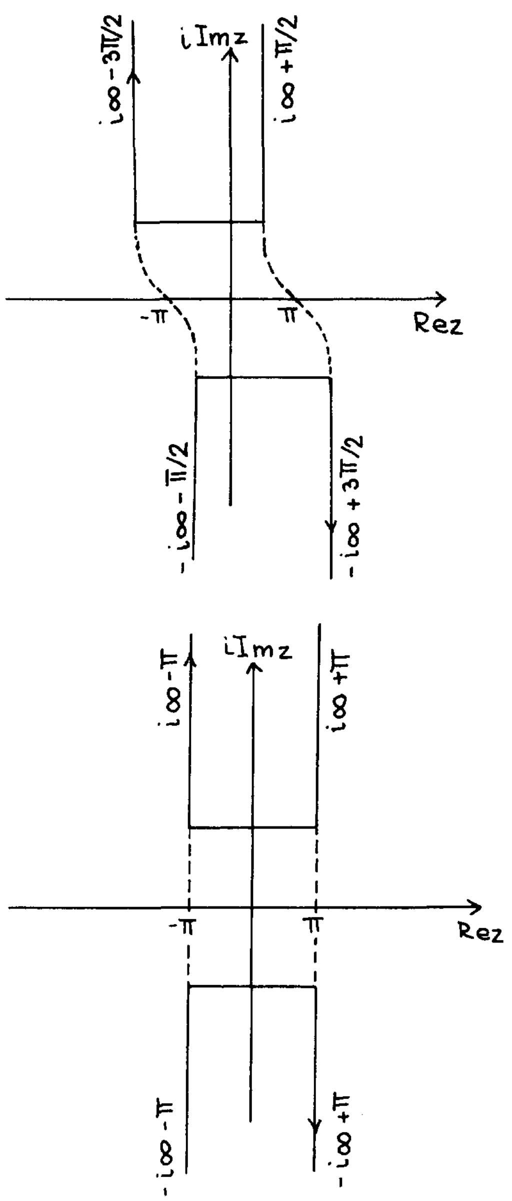

FIG. 1. The Sommerfeld countour γ. (a) $k^2>0$; (b) $k^2<0$, Im $k>0$. ------- steepest descent paths: (a) Re $z = \pm\pi - \operatorname{arc\,cos}(\cosh \operatorname{Im} z)^{-1}$; (b) Re $z = \pm\pi$.

transforms of the wave function.

II. THE SOMMERFELD–MALUZHINETZ TRANSFORMATION

Maluzhinetz[11–13] has shown that the Helmholtz equation

$$\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)+\frac{1}{r^2}\frac{\partial}{\partial\varphi^2}+k^2\right]P(r,\varphi)=0, \tag{2}$$

with the following mixed or impedance boundary conditions:

$$\left[\frac{1}{r}\frac{\partial}{\partial\varphi}\mp ik\sin\Theta_\pm\right]P(r,\varphi)=0,\quad \text{for } \varphi=\pm\Phi, \tag{3}$$

and with the scattering conditions at infinity[11] can be solved in a closed form by representing the solution in the form

$$P(r,\varphi)=\frac{1}{2\pi i}\int_\gamma e^{-ikr\cos z}s(z+\varphi)dz, \tag{4}$$

where γ is known as the Sommerfeld contour[15] and consists of two curves symmetric with respect to the origin in the complex z plane, Fig. 1.

The necessary and sufficient condition for vanishing of the Sommerfeld integral is that the transform be an even function of z. The above statement is true for the class of functions $s(z)$ for which $s(z)=O\{\exp[(1-a)|\operatorname{Im} z|]\}$ with $a>0$, for $|\operatorname{Im} z|\to\infty$. For the corresponding class of original functions (4), we then have: $P(r,\varphi)={}_{r\sim 0}O[r^{a-1}]$ see Ref. 16. The solutions of physical interest are contained in this class. The conditions (3) together with the above mentioned conditions of vanishing of the Sommerfeld integral imply the following system of equations:

$$(\sin z\pm\sin\Theta_\pm)s(z\pm\Phi)$$

$$-(-\sin z\pm\sin\Theta_\pm)s(-z\pm\Phi)=0. \tag{5}$$

In addition, $[s(z)-(z-\varphi_0)^{-1}]$ is said to be regular in the strip $|\operatorname{Re} z|<\Phi$, this condition corresponds to the scattering condition of incidence of the plane or surface wave $\exp\{-ikr\cos(\varphi-\varphi_0)\}$, where φ_0 is real and $|\varphi_0|<\Phi$ or $\varphi_0=\pm(\Phi-\Theta_\pm)$.

The function $s(z)$ can be factorized as follows:

$$s(z)=\sigma(z)F(z)/F(\varphi_0), \tag{6}$$

where $\sigma(z)$ satisfies the equations

$$\sigma(z\pm\Phi)-\sigma(-z\pm\Phi)=0, \tag{7}$$

and has a simple pole at $z=\varphi_0$ with residue equal to one. The function $F(z)$ is the so-called principal solution of Eq. (5) with no poles and no zeros in the strip $|\operatorname{Re} z|<\Phi$. Its logarithmic derivative satisfies the system of inhomogeneous functional equations with constant coefficients, which Maluzhinetz solved by applying the Fourier transformation. Finally, $F(z)$ is expressed as a product of four special meromorphic functions defined by Maluzhinetz[12,13] (also see Appendix in Ref. 17).

The Sommerfeld integral

$$F(r)=\frac{1}{2\pi i}\int_\gamma e^{-ikr\cos z}f(z)\,dz, \tag{8}$$

where $|F(r)|<Mr^{-1+a}e^{br}, M,a,b>0$, has the unique inversion in the class of odd functions that are regular on γ and inside its loops, except possibly for infinitely distant points, and that also satisfy in those regions the inequality $|f(z)|<M\exp[(1-a_1)\operatorname{Im}|z|]$ for some $a_1>0$. For $\operatorname{Re}[-ik\cos z]>b$, this function has the following representation[16]:

$$f(z)=\tfrac{1}{2}ik\sin z\int_0^\infty F(r)e^{ikr\cos z}dr, \tag{9}$$

and for that function $a_1=a$.

A similar pair of transforms was in fact applied to the wedge problem by Senior[18], who started by taking the Laplace transformation of the Helmholtz equation with respect to the radial variable, then followed this by a number of successive transformations. The integral of a type similar to

Eq. (9) was applied also by Gaudin and Derrida[9] with $\cos z$ in the exponent replaced by $\sin z$.

The asymptotic form of the Sommerfeld integral (4) for large values of the radial variable may be found by deforming the original contour into two paths of the steepest descent, shown in Fig. 1. The sum of residues at the poles within the region terminated by the original contour and the steepest descent paths corresponds to the sum of the incident and rescattered waves and to the outgoing surface waves. The integrals along the SD paths, calculated by the saddle point method, give in the first approximation the circular outgoing wave, for $k^2 > 0$, with angle dependent coefficients proportional to $[s(\varphi + \pi) - s(\varphi - \pi)]$.

As shown previously[7,8] the solution of the diffraction problem (2), (3) continued analytically in the $k \equiv \sqrt{E}$ variable is at the same time also the solution of the quantum-mechanical problem of three colinear particles with two-body interactions of type similar to Eq. (1), with the total energy in the center of mass system equal to E.

III. THE BOUNDARY CONDITION MODEL WITH ONE DELTA-TYPE INTERACTION

The one-dimensional potential $v(x) = -2g\delta(x)$ is equivalent to the following junction condition imposed on the solution of the Schroedinger equation for two free particles:

$$\psi'(0^+) - \psi'(0^-) = -2g\,\psi(0), \quad \psi(0^+) = \psi(0^-). \tag{10}$$

If we decompose the wave function into its symmetric and antisymmetric parts

$$\psi_\pm(x) = \tfrac{1}{2}[\psi(x) \pm \psi(-x)] \tag{11}$$

we see that the functions $\psi_\pm$ satisfy the following boundary conditions:

$$\psi_-(0) = 0, \quad \psi_+'(0^+) + g\psi_+(0^+) = 0. \tag{12}$$

When proceeding to the three-particle system with delta function potentials, we find that when each line of the two-body interaction is an axis of symmetry of the total system, the problem can be decomposed by a method similar to that described above into "subproblems" with hard-core or impedance boundary conditions. The only three-body model of the desired symmetry is the model of three identical particles, for which the solution may be found by simple methods.[4] A less trivial model for which the solution can easily be constructed from the known solution for the boundary condition model, described in Sec. II, is the model of two identical particles interacting via a delta-function potential and each of which interacts with the third particle via the impedance-type potential. We describe this below. We define the coordinates of the three-body system in the usual way as follows:

$$s_i = (r_j - r_k)\left(\frac{2m_j m_k}{m_j + m_k}\right)^{1/2},$$

$$t_i = \left(\frac{m_j r_j + m_k r_k}{m_j + m_k} - r_i\right)\left(\frac{2m_i(m_j + m_k)}{m_1 + m_2 + m_3}\right)^{1/2},$$

$$p_i = \frac{m_k k_j - m_j k_k}{[2m_j m_k(m_j + m_k)]^{1/2}},$$

$$q_i = \frac{m_i(k_j + k_k) - (m_j + m_k)k_i}{[2m_i(m_j + m_k)(m_1 + m_2 + m_3)]^{1/2}}, \tag{13}$$

where i,j,k is any cyclic permutation of the numbers 1,2 and 3, and r_i, k_i are the position and momentum coordinates respectively in the center of mass system.

Now let $m_2 = m_3 = m$. The model under consideration satisfies the following Schroedinger equation:

$$\begin{aligned} &-\Delta\,\Psi(s,t) = E\,\Psi(s,t),\\ &\frac{\partial}{\partial s_3}\Psi(s_3,t_3) - \alpha\,\Psi(s_3,t_3) = 0, \quad \text{for } s_3 = 0,\\ &\frac{\partial}{\partial s_2}\Psi(s_2,t_2) + \alpha\,\Psi(s_2,t_2) = 0, \quad \text{for } s_2 = 0,\\ &\frac{\partial}{\partial s_1}\Psi(s_1 = 0^+,t_1) - \frac{\partial}{\partial s_1}\Psi(s_1 = 0^-,t_1)\\ &\quad = -2g\,\Psi(s_1,t_1),\\ &\Psi(s_1 = 0^+,t_1) = \Psi(s_1 = 0^-,t_1), \end{aligned} \tag{14}$$

where (s,t) denotes one of the orthogonally equivalent coordinate systems (s_i,t_i). The initial ordering of particles is assumed to be (1 2 3) or (1 3 2). For positive α and g, the interactions are then purely attractive. The coordinate systems and the regions where the wave function is different from zero are shown in Fig. 2.

We also define the position and momentum polar coordinates by the equalities:

$$\begin{aligned} s_1 &= r\sin\varphi, \quad p_1 = \sqrt{E}\,\sin\psi,\\ t_1 &= r\cos\varphi, \quad q_1 = \sqrt{E}\,\cos\psi. \end{aligned} \tag{15}$$

The lines $s_2 = 0$ and $s_3 = 0$ correspond to the lines $\varphi = \Phi$ and $\varphi = -\Phi$, respectively, with

$$\Phi = \arctan\left[\frac{2m + m_1}{m_1}\right]^{1/2}. \tag{16}$$

It is also convenient to introduce the following complex parameters[7,8]:

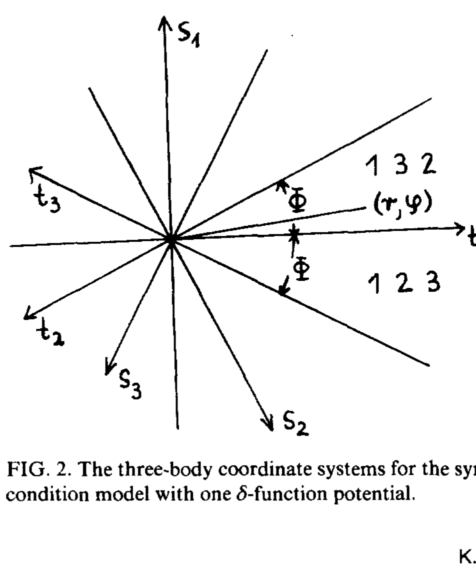

FIG. 2. The three-body coordinate systems for the symmetrical boundary condition model with one δ-function potential.

$$\Theta = \arctan\frac{-i\alpha}{(E+\alpha^2)^{1/2}}, \quad \zeta = \arctan\frac{-ig}{(E+g^2)^{1/2}}, \tag{17}$$

$$\sqrt{E+\alpha^2} = \sqrt{E}\ \cos\Theta, \quad \sqrt{E+g^2} = \sqrt{E}\ \cos\zeta,$$

$$\alpha = i\sqrt{E}\ \sin\Theta, \quad g = i\sqrt{E}\ \sin\zeta. \tag{18}$$

In the above expressions the square root refers to the branch with the positive imaginary part and the arctan denotes the principal branch of the arctangent function. The energy E for the scattering problem may range from $\min(-\alpha^2, -g^2)$ to plus infinity. The discussion in Refs. 7 and 8 concerning the behavior of the functions $\Theta(E)$ can be transferred here with a few changes. The incident wave is of the form $\exp[-iE^{1/2}r\cos(\varphi-\varphi_0)]$, where φ_0 is real and such that $0<|\varphi_0|<\Phi$, which corresponds to the collision of three free particles, or $\varphi_0 = \pm(\Phi-\Theta)$, which corresponds to the initial situations in which particles (1,3) or (1,2) are bounded. Finally, particle 1 may scatter off the bound pair (2,3) and this corresponds to the incident wave $\exp[-iE^{1/2}r\cos(|\varphi|-\zeta)]$.

The boundary conditions (14) can now be written as follows:

$$\left[\frac{1}{r}\frac{\partial}{\partial\varphi} \mp i\sqrt{E}\ \sin\Theta\right]\Psi(r,\varphi)=0, \quad \text{for } \varphi = \pm\Phi,$$

$$\Psi(r,\varphi=0^+) = \Psi(r,\varphi=0^-)$$

$$\frac{1}{r}\frac{\partial}{\partial\varphi}\Psi(r,\varphi=0^+) - \frac{1}{r}\frac{\partial}{\partial\varphi}\Psi(r,\varphi=0^-) = -2i\sqrt{E}\ \sin\zeta\ \Psi(r,\varphi=0). \tag{19}$$

Let us express the solution $\Psi(r,\varphi)$ in the sectors $-\Phi<\varphi<0$ and $0<\varphi<\Phi$ in the form of the Sommerfeld transforms:

$$\Psi(r,\varphi) = \frac{1}{2\pi i}\int_\gamma e^{-i\sqrt{E}\,r\cos z}s_1(z+\varphi)dz, \quad \text{for } -\Phi<\varphi<0.$$

$$\Psi(r,\varphi) = \frac{1}{2\pi i}\int_\gamma e^{-i\sqrt{E}\,r\cos z}s_2(z+\varphi)dz, \quad \text{for } 0<\varphi<\Phi. \tag{20}$$

The boundary conditions (19) combined with the fact that the vanishing of the Sommerfeld integral is equivalent to the vanishing of the odd part of the transform imply the following system of functional equations:

$$(\sin z + \sin\Theta)s_2(z+\Phi) = (-\sin z + \sin\Theta)s_2(-z+\Phi); \tag{21}$$

$$s_1(z) - s_1(-z) = s_2(z) - s_2(-z); \tag{22}$$

$$(\sin z + \sin\zeta)[s_1(z) - s_2(-z)] = (-\sin z + \sin\zeta)[s_1(-z) - s_2(z)]; \tag{23}$$

$$(\sin z - \sin\Theta)s_1(z-\Phi) = (-\sin z - \sin\Theta)s_1(-z-\Phi). \tag{24}$$

According to the incident wave conditions imposed on the functions $\Psi(r,\varphi)$, the functions $s_1(z)$ and $s_2(z)$ are to be regular in the strips $-\Phi<\operatorname{Re} z<0$ and $0<\operatorname{Re} z<\Phi$, respectively, except for the first order pole of one of them at $z=\varphi_0$ with the residuum equal to one, or except for the first order pole of $s_1(z)$ at $z=-\zeta$ and of $s_2(z)$ at $z=\zeta$, both residua being equal to one.

By adding and subtracting from each other Eqs. (21) and (24), we can separate the system (21)–(24) into two pairs of the simultaneous Maluzhinetz equations of type (5) for the functions $[s_1(z) \pm s_2(-z)]$, corresponding to the wedge problem of angle Φ, rotated with respect to the problem (3) by $-\Phi/2$. Making use then of Maluzhinetz solution and also taking into account the regularity conditions imposed on the functions $s_1(z)$ and $s_2(z)$, we obtain the following results:

$$s_1(z) = -\frac{\pi}{2\Phi}\frac{\sin\pi\varphi_0/\Phi}{\cos\pi z/\Phi - \cos\pi\varphi_0/\Phi} \times\left[\mp\frac{F(z+\Phi/2,\Phi/2,\zeta,\Theta)}{F(\mp\varphi_0+\Phi/2,\Phi/2,\zeta,\Theta)} + \frac{F(z+\Phi/2,\Phi/2,-i\infty,\Theta)}{F(\mp\varphi_0+\Phi/2,\Phi/2,-i\infty,\Theta)}\right],$$

$$s_2(z) = -\frac{\pi}{2\Phi}\frac{\sin\pi\varphi_0/\Phi}{\cos\pi z/\Phi - \cos\pi\varphi_0/\Phi} \times\left[\pm\frac{F(-z+\Phi/2,\Phi/2,\zeta,\Theta)}{F(\mp\varphi_0+\Phi/2,\Phi/2,\zeta,\Theta)} + \frac{F(-z+\Phi/2,\Phi/2,-i\infty,\Theta)}{F(\mp\varphi_0+\Phi/2,\Phi/2,-i\infty,\Theta)}\right], \tag{25}$$

where the upper signs correspond to $\operatorname{Re}\varphi_0>0$ and the lower ones to $\operatorname{Re}\varphi_0<0$.

For the case of the incident symmetric "surface" wave, we find

$$s_1(z) = \pi/\Phi\,\frac{\sin\pi\zeta/\Phi}{\cos\pi z/\Phi - \cos\pi\zeta/\Phi} \times\frac{F(z+\Phi/2,\Phi/2,\zeta,\Theta)}{F(-\zeta+\Phi/2,\Phi/2,\zeta,\Theta)},$$

$$s_2(z) = -\pi/\Phi\,\frac{\sin\pi\zeta/\Phi}{\cos\pi z/\Phi - \cos\pi\zeta/\Phi} \times\frac{F(-z+\Phi/2,\Phi/2,\zeta,\Theta)}{F(-\zeta+\Phi/2,\Phi/2,\zeta,\Theta)}, \tag{26}$$

In the above expressions $F(z,\Phi,\Theta_+,\Theta_-)$ denotes the principal solution of Eqs. (5). The rather formal expression $F(z,\Phi,-i\infty,\Theta_-)$ denotes the principal solution of Eqs. (5) with the upper sign equation replaced by $s(z+\Phi) = s(-z+\Phi)$ and corresponding to the boundary condition of a hard core (see Refs. 7 and 8.)

Following the procedure carried out in the papers cited above the scattering probability amplitudes and then the cross sections can be calculated from the asymptotic form of the solution for $r\to\infty$. Alternatively, we can calculate them from the off-shell amplitudes that are studied in Sec. V and expressed there in terms of the functions $s_1(z)$ and $s_2(z)$. We only remark here that contrary to the boundary condition model where only the moduli of the Maluzhinetz special functions were involved, not all of the cross sections are now expressible by elementary functions. The origin of this fact lies in the additive structure of Eqs. (25). The verification of the flux conservation rules is, however, no more difficult to show than in the previous case.

It is interesting that the model solved by Lieb and Koppe[2] with help of the Wiener–Hopf method can also be treated by the method described above. Those authors studied the simplified model of the stripping reaction described by the following Schroedinger equation:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + E\right)\Psi(x,y) = -2C\delta(y)H(x)\Psi(x,y), \qquad (27)$$

where $H(x)$ is the Heaviside function. Only the case corresponding to the symmetric incident wave $\exp[-C|y| - i(E + C^2)^{1/2}x]$ was considered. Equation (27) has the same symmetry as the model described above, namely one axis of symmetry agreeing with the line of the delta-function interaction. Following the procedure outlined above, we can express the solution of Eq. (27) for an arbitrary incident wave as the Sommerfeld integral with the transform composed of Maluzhinetz solutions for $2\Phi = \pi$. We have verified that both methods give the same results for the elastic and break-up cross sections.

IV. THE MODEL OF TWO IDENTICAL PARTICLES INTERACTING WITH A FIXED WALL THROUGH DELTA FUNCTION POTENTIALS

Let $m_1 = \infty$, $m_2 = m_3 = m$. According to definitions (13) and (16) we have

$$s_2 = -t_3, \quad t_2 = s_3, \quad p_2 = -q_3, \quad q_2 = p_3, \quad \Phi = \pi/4. \qquad (28)$$

We consider the following Schroedinger equation:

$$[-\Delta - 2g_1\delta(s_1) - 2g\delta(s_2) - 2g\delta(s_3)]\Psi(s,t) = E\,\Psi(s,t). \qquad (29)$$

For simplicity we study only the case in which all interactions are attractive, i.e., $g_1 > 0$ and $g > 0$. Similar considerations may however be performed for the model of the same symmetry with $g_1 < 0$ and $g > 0$.

The coordinate systems and the lines of interactions of the model under study are shown in Fig. 3. In order to apply the symmetry of the system with respect to the line $s_1 = 0$ we

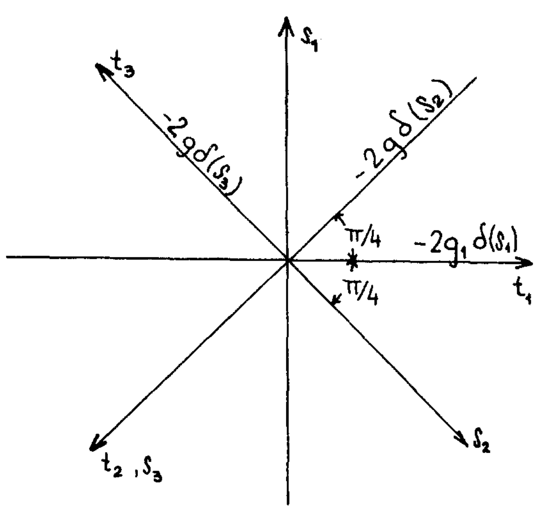

FIG. 3. The coordinate systems and the lines of interaction for the delta potential model with $m_1 = \infty$, $m_2 = m_3$.

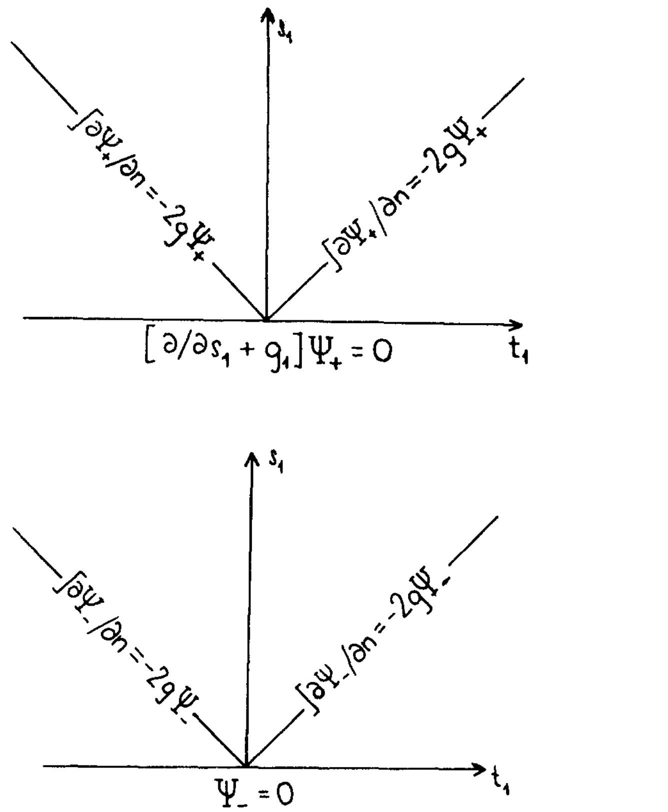

FIG. 4. Boundary condition problems for the functions Ψ_+ and Ψ_-. The symbol "$\int$" denotes the jump of the normal derivative on the lines $s_1 = \pm t_1$.

write

$$\Psi_\pm(s_1,t_1) \equiv \frac{1}{2}[\Psi(s_1,t_1) \pm \Psi(-s_1,t_1)], \qquad (30)$$

with the analogous decomposition of the initial state function. The problems corresponding to the "upper parts" of the functions are demonstrated in Fig. 4. The solution Ψ_- can be easily constructed by adding the appropriate number of rescattering terms to the incident wave in such a way that all the boundary conditions will be fulfilled. Using the diffraction theory language we say that the geometric optics approximation is in that case exact. The essential question is to find the solution Ψ_+. We decompose it further as follows:

$$\Psi_{+\pm}(s_1,t_1) = \frac{1}{2}[\Psi_+(s_1,t_1) \pm \Psi_+(s_1,-t_1)]. \qquad (31)$$

The problem demonstrated in Fig. 4(a) is then reduced to two separate problems in the quarter-plane. It is convenient to formulate them in terms of the polar coordinates defined as follows:

$$s_1 = r\sin(\varphi + \pi/4), \quad t_1 = r\cos(\varphi + \pi/4). \qquad (32)$$

In the region $-\pi/4 < \varphi < \pi/4$ the functions Ψ_{++} and Ψ_{+-} are solutions of the Helmholtz equation with the following conditions:

$$\frac{1}{r}\frac{\partial}{\partial\varphi}\Psi_{++}(r,\varphi) = 0, \quad \Psi_{+-}(r,\varphi) = 0 \text{ for } \varphi = \pi/4,$$

$$\frac{1}{r}\frac{\partial}{\partial\varphi}\Psi_{+\pm}(r,\varphi=0^+)-\frac{1}{r}\frac{\partial}{\partial\varphi}\Psi_{+\pm}(r,\varphi=0^-)$$
$$=-2i\sqrt{E}\,\sin\Theta\Psi_{+\pm}(r,\varphi=0),$$
$$\Psi_{+\pm}(r,\varphi=0^+)=\Psi_{+\pm}(r,\varphi=0^-)$$
$$\frac{1}{r}\frac{\partial}{\partial\varphi}\Psi_{+\pm}(r,\varphi)+i\sqrt{E}\,\sin\Theta_1\Psi_{+\pm}(r,\varphi)$$
$$=0\quad\text{for}\ \varphi=-\pi/4. \tag{33}$$

The parameters Θ and Θ_1 have the same meaning as in the previous cases, i.e.,

$$\Theta=\text{Arctan}\frac{-ig}{(E+g^2)^{1/2}},\quad \Theta_1=\text{Arctan}\frac{-ig_1}{(E+g_1^2)^{1/2}}. \tag{34}$$

For $E<-g^2$ the parameter Θ is a real number in the interval $(-\pi/2,0)$; for $-g^2<E<0$ we have $\text{Re}\,\Theta=-\pi/2$, $\text{Im}\Theta<0$ and for $E>0$, i.e., above the break-up threshold, Θ is purely imaginary with $\text{Im}\,\Theta<0$. The same goes for the function $\Theta_1(E)$ (if g is replaced by g_1). The initial state functions in the quarter-plane $|\varphi|<\pi/4$ are

$$\Phi_{\text{in}}=\exp\big[-i\sqrt{E}\,r\cos(\varphi-\varphi_0)\big]$$

for φ_0 real, $|\varphi_0|<\pi/4$ (three incident particles),

$$\Phi_{\text{in}}=\exp\big[-i\sqrt{E}\,r\cos(\varphi-\Theta_1+\pi/4)\big]$$

(particles 2 and 3 bounded), or

$$\Phi_{\text{in}}=\exp\big[-i\sqrt{E}\,r\cos(|\varphi|-\Theta)\big]$$

(particles 1 and 2 or 2 and 3 bounded).

In the above expressions we have omitted the normalization factors and the factors resulting from the symmetrization of the total initial state function.

Let us express the functions Ψ_{+-} and Ψ_{+-} in the sectors $-\pi/4<\varphi<0$ and $0<\varphi<\pi/4$ in the form of the Sommerfeld transforms

$$\Psi_{+\pm}(r,\varphi)=\frac{1}{2\pi i}\int_\gamma \exp(-i\sqrt{E}\,r\cos z)\,s_1^\pm(z+\varphi)\,dz,\quad \text{for}\ -\pi/4<\varphi<0,$$

$$\Psi_{+\pm}(r,\varphi)=\frac{1}{2\pi i}\int_\gamma \exp(-i\sqrt{E}\,r\cos z)\,s_2^\pm(z+\varphi)\,dz,\quad \text{for}\ 0<\varphi<\pi/4. \tag{35}$$

The scattering conditions require that $s_1^\pm(z)$ and $s_2^\pm(z)$ be regular in the strips $-\pi/4<\text{Re}\,z<0$ and $0<\text{Re}\,z<\pi/4$ respectively except for the first order pole of one of them with the residuum equal to N, where N is the normalization constant. Here we shall set $N=1$. For the case in which a "surface" wave comes from infinity along the line $\varphi=0$, both pairs of functions $s_1^\pm(z)$ and $s_2^\pm(z)$ have first order poles of equal residua at $z=-\Theta$ and $z=\Theta$ respectively. As before we demand that the wave function be bounded for $r\to0$, this condition being a sufficient one for the total particle flux through an infinitely small circle with the center at $r=0$ to remain finite. The above condition implies that $s_1^\pm(z)$ and $s_2^\pm(z)$ tend to finite values for $|\text{Im}\,z|\to\infty$. The boundary conditions (33) imply the following two systems of functional equations:

$$(\sin z-\sin\Theta_1)s_1^\pm(z-\pi/4)=(-\sin z-\sin\Theta_1)s_1^\pm(-z-\pi/4), \tag{36}$$

$$(\sin z+\sin\Theta)[s_1^\pm(z)-s_2^\pm(-z)]=(-\sin z+\sin\Theta)[s_1^\pm(-z)-s_2^\pm(z)], \tag{37}$$

$$s_1^\pm(z)-s_1^\pm(-z)=s_2^\pm(z)-s_2^\pm(-z), \tag{38}$$

$$s_2^\pm(z+\pi/4)=\mp s_2^\pm(-z+\pi/4). \tag{39}$$

In the above equations the upper and lower signs go together. Unlike the system of equations considered in Sec. III, the above systems of equations cannot be decoupled into the system of Maluzhinetz equations. More advanced methods of the theory of functions are needed in order to solve them analytically. Not undertaking this task here, we shall note that each of the systems given by Eqs. (36)–(39) can be reduced to one difference equation of the second order, which is similar to the type of difference equations studied by Jost[1] and Alberviero[5] and as that studied by Gaudin and Derrida,[9] and solved by the authors completely. The problem considered in this paper is however more complicated due to the presence of two, instead of one, different parameters characterizing interactions.

For $\Theta_1=0$ and for $\Theta_1=-i\infty$, which corresponds to the Neumann and Dirichlet boundary conditions, respectively on the line $\varphi=-\pi/4$ in the problems (33), Eq. (36)–(39) are solvable by elementary methods. We shall give here the explicit forms of the functions $s_1^\pm(z)$ and $s_2^\pm(z)$ for $\Theta_1=-i\infty$, this corresponding to replacing Eq. (36) by the equation $s_1^\pm(z-\pi/4)=s_1^\pm(-z-\pi/4)$. With the help of these functions one can construct the "trivial" part $\Psi_-(r,\varphi)$ of the total solution $\Psi(r,\varphi)$. In the coordinate system defined by (32) these functions present the Sommerfeld transforms of the symmetric and antisymmetric parts of $\Psi_-(r,\varphi)$ with respect to the line $t_1=0$ [see Fig. 4(b)]. For $\Theta_1=-i\infty$, we find

$$s_1^-(z)=\tfrac12[F(z)+G(z)],\quad s_2^-(z)=\tfrac12[G(z)-F(-z)],$$

where

$$G(z)=\frac{-4\sin4\varphi_0}{\cos4z-\cos4\varphi_0},$$

$$F(z)=\frac{-4\sin4\varphi_0}{\cos4z-\cos4\varphi_0}\times\frac{(\sin z-\sin\Theta)}{(\sin\varphi_0\pm\sin\Theta)}\frac{(\cos z+\sin\Theta)}{(\cos\varphi_0+\sin\Theta)},$$

for $\text{Re}\,\varphi_0\leqslant0$, and

$$G(z)=0,\quad F(z)=\frac{2\cos\Theta(\cos\Theta-\sin\Theta)}{(\sin z+\sin\Theta)(\cos z-\sin\Theta)}, \tag{40}$$

for the incident surface wave;

$$s_1^+(z)=\tfrac12[f(z)+f(z-\pi/2)],$$

$$s_2^+(z)=\tfrac12[f(z-\pi/2)-f(-z)],$$

where

$$f(z) = \frac{-2\sin 2\varphi_0}{\cos 2z - \cos 2\varphi_0}\frac{\sin z - \sin\Theta}{\sin\varphi_0 \pm \sin\Theta} - \frac{2\sin 2\varphi_0}{\cos 2z + \cos 2\varphi_0}\frac{\sin z - \sin\Theta}{\cos\varphi_0 + \sin\Theta}, \quad \text{for } \mathrm{Re}\varphi_0 \leqslant 0$$

and

$$f(z) = \frac{2\cos\Theta}{\sin z + \sin\Theta} \tag{41}$$

for the incident surface wave.

V. THE RELATIONS BETWEEN THE HALF-OFF-SHELL PROBABILITY AMPLITUDES AND THE SOMMERFELD–MALUZHINETZ TRANSFORMS OF THE WAVE FUNCTION

Here we show that the half off-the-energy shell extensions of the probability amplitudes of the elastic and exchange processes for the three-body models studied in the previous sections are connected in a simple way with the Sommerfeld transforms of the wave function. The method is, in principle, the same as that employed in Ref. 17 for the exactly soluble boundary condition model, with the difference that we make use here exclusively of the functional properties determining the transforms without refering to any analytic solutions.

The half-off-shell amplitudes under consideration present the unknown functions of the Faddeev–Lovelace equations.[19] The numerical study of those equations for the delta potential model was carried out by Dodd.[20] Dodd[21] and Majumdar[22] studied those equations analytically for the exactly soluble model of three identical particles. The analytical structure of the one-dimensional Faddeev equations for a slightly more general class of interactions was also studied by Brayshaw and Peierls.[23] Although we do not touch here on the problems dealing with the structure of the Faddeev–Lovelace equations themselves, we believe that the introduction of the complex polar momentum variable as defined below, which uniformizes the square root functions occurring in the kernel of the above equations, may be helpful in clarifying some of those problems.

The two-body t matrices for the interactions of interest are of the one-term separable form

$$\langle p'|t(z)|p\rangle = \overline{f(p')}\tau(z)f(p), \tag{42}$$

where p is the momentum variable defined by Eq. (13), z is the complex energy parameter, $\tau(z)$ is the two-body propagator, and $f(p) = -\varphi_b(p)(p^2 + E_b)$ is the form factor of the bound state with the eigenfunction φ_b and eigenenergy E_b.

For the delta potential, Eqs. (10)

$$f(p) = -(2g^3/\pi)^{1/2}, \tag{43}$$

$$\tau(z) = -\frac{1}{2g^2}\frac{i\sqrt{z}}{g + i\sqrt{z}}, \tag{44}$$

and for the impedance interactions described by the boundary conditions $\psi'(0) \pm \alpha\,\psi(0) = 0$ with $\alpha > 0$, we have[14]

$$f(p) = -i(\alpha/\pi)^{1/2}(\pm p + i\alpha), \tag{45}$$

$$\tau(z) = -\frac{i\sqrt{z}}{\alpha(z + \alpha^2)}. \tag{46}$$

We consider the functions $h_{ji}(q_j',q_i;E + i0)$, which are the solutions of the following Faddeev–Lovelace equations:

$$h_{ji}(q_j',q_i;E + i0) = b_{ji}(q_j',q_i;E + i0) + \sum_{k=1}^{3}\int_{-\infty}^{\infty} dq''\, b_{jk}(q_j,q'';E + i0) \times \tau_k(E + i0 - q''^2)h_{ki}(q'',q_i;E + i0), \tag{47}$$

where $i,j,k = 1,2,3;q_i$ are defined by Eq. (13),

$$b_{jk}(q_j',q_k;E + i0) = \langle q_j',j|(E + i0 - H_0)^{-1}|k,q_k\rangle \times (1 - \delta_{jk}). \tag{48}$$

$|i\rangle$ denotes the form factor vector of the pair $i = (j,k)$, i.e., $f_i(p_i) = \langle i|p_i\rangle$, H_0 is the kinetic energy operator, and E is the total energy of the three-body syetem. Equations (47) are considered to be "half-on the energy shell" in the meaning that

$$E = q_i^2 - E_{b,i}, \tag{49}$$

where $E_{b,i}$ denotes the bound-state energy of a pair i. As is well known the half-off-shell amplitudes h_{ji} are related in the following way to the scattering wave function:

$$h_{ji}(q_j',q_i;E + i0) = \langle q_j',j|(E + i0 - H_0)^{-1}(V_i + V_k)|\Psi_{i,q_i}\rangle, \tag{50}$$

where $|\Psi_{i,q_i}\rangle$ is the vector of the scattering wave function corresponding to the initial state vector $|\varphi_i,q_i\rangle$ and normalized to a unit current density of incident particles, and V_i, V_k are the operators of interaction of the pair i and k, respectively. In the following, the part $+i0$ at the energy variable is dropped. We shall use Eq. (50) as a starting point to derive the expressions relating the amplitudes h_{ji} to the Sommerfeld transforms of the polar coordinate representation of the state vector $|\Psi_{i,q_i}\rangle$. At the beginning we give some preliminary results that enable us to treat the models under consideration in a uniform way.

Let $\Psi(r,\varphi)$ be the scattering solution of the Helmholtz equation in the angular region $|\varphi| < \Phi$, where $\Phi < \pi$, satisfying the conditions as given by Eq. (19) where eventually two different parameters Θ_+ and Θ_- may appear for $\varphi = \Phi$ and $\varphi = -\Phi$, respectively.[24] Let $s_1(z)$ and $s_2(z)$ denote the Sommerfeld transforms of the solution as defined by Eq. (20). The transforms satisfy the system of equations (21)–(24) with Θ_+ appearing in Eq. (21) and Θ_- appearing in Eq. (24). The operators of the interactions corresponding to imposing the junction or the boundary conditions on the lines $\varphi = 0$, $\varphi = \pm\Phi$ are denoted by $V_{\varphi=0}$ and $V_{\varphi=\pm\Phi}$, respectively. Making use of the inversion formula (9) of the Sommerfeld integral and then of the functional equations satisfied by the functions $s_1(z)$ and $s_2(z)$, we obtain the following result:

$$\int_{-\Phi}^{\Phi} d\varphi \int_0^{\infty} dr\, r e^{i\sqrt{E}\,r\cos(z+\varphi)}V_{\varphi=0}\Psi(r,\varphi) = -2i\sin\zeta\int_0^{\infty} dr\, e^{i\sqrt{E}\,r\cos z}\Psi(r,\varphi = 0)$$

$$= -\frac{2\sin\zeta}{\sin z}[s_1(z)-s_1(-z)] = -2[s_1(z)-s_2(z)]\,, \tag{51}$$

valid for the values of z contained in the region inside the loops of the Sommerfeld contour and such that the integral in (51) converges. With the similar restrictions imposed on $z+\Phi$ and $z-\Phi$, respectively, we also obtain

$$\begin{aligned}\int_{-\Phi}^{\Phi} d\varphi \int_0^\infty r\,dr\, e^{i\sqrt{E}\, r\cos(z+\varphi)} V_{\varphi=\Phi}\Psi(r,\varphi) \\ &= -\int_0^\infty dr\, \Psi(r,\varphi=\Phi)\Big[\Big(\frac{1}{r}\frac{\partial}{\partial\varphi} - i\sqrt{E}\,\sin\Theta_+\Big)e^{i\sqrt{E}\, r\cos(z+\varphi)}\Big]_{\varphi=\Phi} \\ &= \frac{\sin(z+\Phi)+\sin\Theta_+}{\sin(z+\Phi)}[s_2(z+2\Phi)-s_2(-z)] = -2\,s_2(-z)\end{aligned} \tag{52}$$

and

$$\begin{aligned}\int_{-\Phi}^{\Phi} d\varphi \int_0^\infty r\,dr\, e^{i\sqrt{E}\, r\cos(z+\varphi)} V_{\varphi=-\Phi}\,\Psi(r,\varphi) \\ &= -\int_0^\infty dr\, \Psi(r,\varphi=-\Phi)\Big[\Big(\frac{1}{r}\frac{\partial}{\partial\varphi} + i\sqrt{E}\,\sin\Theta_-\Big)e^{i\sqrt{E}\, r\cos(z+\varphi)}\Big]_{\varphi=-\Phi} \\ &= -\frac{\sin(z-\Phi)-\sin\Theta_-}{\sin(z-\Phi)}[s_1(z-2\Phi)-s_1(-z)] = -2\,s_1(-z)\,.\end{aligned} \tag{53}$$

In the left-hand side equalities of (52) and (53), we have applied the rules of action of the distributions $V_{\varphi=\pm\Phi}$ as derived in Ref. 14. Before passing to the calculation of the amplitudes, we define also the "off-shell" momentum polar coordinate η by the following equalities:

$$q=\sqrt{E}\,\cos\eta,\quad \sqrt{E-q^2}=\sqrt{E}\,\sin\eta\,. \tag{54}$$

The domain D of the η variable is determined as follows:

$$D=\{(\eta: -\pi/2<\mathrm{Re}\eta<\pi/2\,,\mathrm{Im}\eta>0)\cup(\pi/2<\mathrm{Re}\eta<3\pi/2,\mathrm{Im}\eta<0)\},\quad \text{for } E>0\,,$$

$$D=\{\eta:0<\mathrm{Re}\eta<\pi\},\quad \text{for } E<0\,.$$

The region D is mapped by the transformation (54) onto the cut plane q on which $\mathrm{Im}\sqrt{E-q_2}>0$.

Let us consider now the scattering off the bound state in the model system described in Sec. III. Applying the Eqs. (15), (18), (51) and (52), and well as the functional equations for $s_1(z)$ and $s_2(z)$, we find:

$$\begin{aligned}h_{3i}(q',q_i) &= \langle q',3\,|(E-H_0)^{-1}(V_1+V_2)\,|\Psi_{i,q_i}\rangle \\ &= \frac{iN_i}{2\pi}\Big(\frac{\alpha}{\pi}\Big)^{1/2}\int_{-\infty}^{\infty} dp_3\,(-p_3-i\alpha)\frac{1}{p_3^2-E+q'^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ds_3dt_3\, e^{-ip_3s_3-iq't_3}(V_1+V_2)\Psi_{i,q_i}(s_3,t_3) \\ &= N_i\Big(\frac{\alpha}{\pi}\Big)^{1/2}\frac{\sqrt{E-q'^2}+i\alpha}{2\sqrt{E-q'^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ds_3dt_3 e^{-i\sqrt{E-q'^2}\,s_3-iqt_3}(V_1+V_2)\Psi_{i,q_i}(s_3,t_3) \\ &= N_i\Big(\frac{\alpha}{\pi}\Big)^{1/2}\frac{\sin\eta'-\sin\Theta}{2\sin\eta'}\int_{-\Phi}^{\Phi} d\varphi\int_0^\infty r\,dr\, e^{i\sqrt{E}\, r\cos(\varphi+\Phi-\eta')}(V_1+V_2)\Psi_{i,q_i}(r,\varphi) \\ &= -\Big(\frac{\alpha}{\pi}\Big)^{1/2}N_i\,\frac{\sin\eta'-\sin\Theta}{\sin\eta'}\,[s_1(\Phi-\eta')-s_2(\Phi-\eta')+s_2(-\Phi+\eta')] \\ &= -\Big(\frac{\alpha}{\pi}\Big)^{1/2}N_i\,\frac{\sin\eta'-\sin\Theta}{\sin\eta'}s_1(-\Phi+\eta')\,,\end{aligned} \tag{55}$$

where $i=1,2,3$ and the normalization constants N_i are $N_1=(g/2\pi)^{1/2}$, $N_2=N_3=(\alpha/\pi)^{1/2}$, and the coordinate η' is related to q' according to (54). In fact, the Sommerfeld transforms should also be supplied by the index determining the initial situation in scattering that, as we recall, determines the regularity conditions imposed on those functions. We, however, drop off such an index.

In a similar manner, we calculate the amplitude $h_{2i}(q',q_i)$ with the help of (51) and (53) and the functional equations (21)–(24):

$$\begin{aligned}h_{2i}(q',q_i) &= \langle q',2\,|(E-H_0)^{-1}(V_1+V_2)\,|\Psi_{i,q_i}\rangle = \cdots \\ &= \Big(\frac{\alpha}{\pi}\Big)^{1/2}N_i\,\frac{\sin\eta'-\sin\Theta}{2\sin\eta'}\int_0^\infty r\,dr\int_{-\Phi}^{\Phi} d\varphi\, e^{i\sqrt{E}\, r\cos(\eta'+\varphi-\Phi)}(V_1+V_2)\Psi_{i,q_i}(r,\varphi) \\ &= \Big(\frac{\alpha}{\pi}\Big)^{1/2}N_i\,\frac{\sin\eta'-\sin\Theta}{\sin\eta'}[s_1(\eta'-\Phi)-s_2(\eta'-\Phi)-s_1(-\eta'+\Phi)] \\ &= -\Big(\frac{\alpha}{\pi}\Big)^{1/2}N_i\,\frac{\sin\eta'-\sin\Theta}{\sin\eta'}s_2(-\eta'+\Phi)\,.\end{aligned} \tag{56}$$

The calculation of the amplitudes $h_{1i}(q',q_i)$ is, in short, as follows:

$$
\begin{aligned}
h_{1i}(q',q_i) &= \langle q,1|(E-H_0)^{-1}(V_2+V_3)|\Psi_{i,q_i}\rangle \\
&= \left(\frac{2g^3}{\pi}\right)^{1/2}\frac{iN_i}{2\sqrt{E-q^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ds_1\,dt_1\,\exp\left(i\sqrt{E-q^2}\,|s_1|-iq't_1\right)(V_2+V_3)\Psi_{i,q_i}(s_i,t_i) \\
&= -\left(\frac{2g}{\pi}\right)^{1/2}N_i\frac{\sin\zeta}{\sin\eta'}\left\{\int_0^{\Phi}d\varphi\int_0^{\infty}r\,dr\exp\left[i\sqrt{E}\,r\cos(\varphi+\eta'-\pi)\right]V_2\Psi(r,\varphi)\right. \\
&\quad\left.+\int_{-\Phi}^{0}d\varphi\int_0^{\infty}r\,dr\exp\left[i\sqrt{E}\,r\cos(\pi+\varphi-\eta')\right]V_3\Psi(r,\varphi)\right\} \\
&= -\left(\frac{2g}{\pi}\right)^{1/2}\frac{\sin\zeta}{\sin\eta'}N_i\left[s_1(\eta'-\pi)-s_2(-\eta'+\pi)\right].
\end{aligned}
\tag{57}
$$

In order to obtain the physical probability amplitudes we put $\eta' = -\Theta$ or $\eta' = \Theta + \pi$, which corresponds to $q' = (E+\alpha^2)^{1/2}$ and $q' = -(E+\alpha^2)^{1/2}$, respectively, in Eqs. (55) and (56), and $\eta' = -\zeta$ or $\eta' = \zeta + \pi$ in Eq. (57). The formal resemblance of the expressions for the amplitudes h_{ji} for different i and the same j is lost if one takes into account the different regularity conditions imposed on the Sommerfeld transforms for different values of i. For instance the functional equation (24) implies that $s_1(-\Phi-\Theta)$ must vanish if $s_1(z)$ is regular for $-\Phi<\operatorname{Re}z<0$ and that $s_1(-\Phi-\Theta) = -\cos\Theta/2\sin\Theta$ if $s_1(z)-(z+\Phi-\Theta)^{-1}$ is regular in the same strip. The above, in turn, implies that $h_{31}(|q_3|,-|q_1|) = h_{32}(|q_3|,|q_2|) = 0$, whereas $h_{33}(|q_3|,|q_3|) = -(i/\pi)|q_3| = -(i/\pi)(E+\alpha^2)^{1/2}$ as demanded by the impenetrability of interactions terminating the region of motion of particles (see Fig. 2).

The amplitude $\tau_{0i}(p',q';q_i)$ of the break-up process can be calculated with help of the Faddeev–Lovelace equation relating this amplitude to the amplitudes h_{ji}[19]

$$
\tau_{0i}(p',q';q_i) = \langle p',q'|(V_1+V_2+V_3)\Psi_{i,q_i}\rangle = \sum_{k=1}^{3}\overline{f_k(p'_k)}\,\tau_k(E+i0-q'^2)h_{ki}(q'_k,q_i),
\tag{58}
$$

where $p'^2_i + q'^2_i = E$ and p',q' denotes any of the orthogonally equivalent systems p'_i, q'_i for $i = 1,2,3$.

It is convenient to consider Eq. (58) in terms of the momentum polar coordinate ψ defined by (15) and the off-shell coordinate η. In the energy region in which the dissociation is possible, we have $0<q^2<E$, which corresponds according to (54) to $\eta\in(0,\pi)$. Taking this into account and comparing the definitions (15) and (54), we find that q'_1, q'_2, and q'_3 in (58) correspond in passing to the polar coordinates to $\eta' = |\psi|$, $\eta' = \psi - \Phi + \pi$, and $\eta' = -\psi - \Phi + \pi$, respectively. Expressing the formfactors and propagators as given by (43)–(46) as trigonometric functions of η' and of the parameters Θ and ζ, and making use of Eqs. (55)–(57) and also of the functional equations for $s_1(z)$ and $s_2(z)$, we can reduce Eq. (58) to the following form:

$$
\begin{aligned}
\tau_{0i}(p',q';q_i) &= \frac{N_i}{\pi}\left[s_2(\psi-\pi)-s_2(\psi+\pi)\right], \quad \text{for } \psi>0, \\
\tau_{0i}(p',q';q_i) &= \frac{N_i}{\pi}\left[s_1(\psi-\pi)-s_1(\psi+\pi)\right], \quad \text{for } \psi<0,
\end{aligned}
\tag{59}
$$

as expected from the asymptotic form of the Sommerfeld integral for $r\to\infty$.[12]

According to these results as well to the results of Sec. III, we find that similarly as for the boundary condition model[17] the half-off-shell elastic and exchange amplitudes for the model under consideration can be expressed in terms of Maluzhinetz special functions.[25]

Let us go now to the delta potential model studied in the previous section, namely to the system with $m_1 = \infty$, $m_2 = m_3$, $V_2 = V_3$. The amplitudes have the following symmetry properties:

$$
\begin{aligned}
&h_{22}(q',q) = h_{33}(q',q), \quad h_{23}(q',q) = h_{32}(q',q), \\
&h_{13}(q',q) = h_{12}(q',q), \quad h_{31}(q',q) = h_{21}(q',q).
\end{aligned}
\tag{60}
$$

Owing to the the symmetricity of the potentials, we get

$$
h_{ji}(-q',q) = h_{ji}(q',-q).
$$

Let us define the following combinations of the amplitudes

$$
\begin{aligned}
h_{vv}(q',q) &= \frac{1}{2}\left[h_{22}(q',q)+h_{32}(q',q)\right] = \frac{1}{2}\left[h_{33}(q',q)+h_{23}(q',q)\right], \\
h_{1v}(q',q) &= h_{13}(q',q) = h_{12}(q',q), \\
h_{v1}(q',q) &= h_{21}(q',q) = h_{31}(q',q), \\
h^{\pm}_{ji}(q',q) &= \frac{1}{2}\left[h_{ji}(q',q)\pm h_{ji}(-q',q)\right], \quad i,j = 1,v.
\end{aligned}
\tag{61}
$$

Making use of Eq. (50), we can easily verify that the amplitudes $h_{ji}^{\pm}(q',q_i)$ are expressible in the following way by the combinations $\Psi_{+\pm}(s,t)$ of the wave function as defined by (30) and (31):

$$h_{vi}^{\pm}(q',q_i) = \langle q',2|(E-H_0)^{-1}(V_1+V_3)|\Psi_{+\pm,q_i,i}\rangle = \langle q',3|(E-H_0)^{-1}(V_1+V_2)|\Psi_{+\pm,q_i,i}\rangle, \tag{62}$$

$$h_{1i}^{\pm}(q',q_i) = \langle q',1|(E-H_0)^{-1}(V_2+V_3)|\Psi_{+\pm,q_i,i}\rangle, \tag{63}$$

where we take $q_i = (E+g^2)^{1/2}$, for $i=v$; and $q_i = -(E+g_1)^{1/2}$, for $i=1$; the above choice corresponding to the directions of the incident wave in the first quarter-plane in the (s,t) plane (Fig. 3). Equations (62) and (63) may be transformed further with help of methods similar to those applied in the previous case if the symmetry properties of the functions $\Psi_{+\pm}$ are also taken into account. For instance, Eq. (62) may be written as follows:

$$\begin{aligned} h_{vi}^{\pm}(q',q_i) &= \frac{N_i}{2\pi}\left(\frac{2g^3}{\pi}\right)^{1/2}\int_{-\infty}^{\infty} dp_2\,\frac{1}{p_2^2-E+q'^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ds_2dt_2 \exp\left[-ip_2s_2-iq't_2\right](V_1+V_3)\Psi_{+\pm,q_i,i} \\ &= \left(\frac{2g^3}{\pi}\right)^{1/2}\frac{iN_i}{2\sqrt{E-q'^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ds_2dt_2 \exp\left[i\sqrt{E-q'^2}\,|s_2|-iq't_2\right](V_1+V_3)\Psi_{+\pm}(s_2,t_s) \\ &= -N_i\left(\frac{2g}{\pi}\right)^{1/2}\frac{\sin\Theta}{2\sin\eta'}\left\{\int_0^{\pi} d\varphi\int_0^{\infty} r\,dr\, e^{i\sqrt{E}\,r\cos(\varphi+\eta')}(V_1+V_3)\Psi_{+\pm}(r,\varphi)\right. \\ &\quad \left.+\int_0^{\pi} d\varphi\int_0^{\infty} r\,dr\, e^{i\sqrt{E}\,r\cos(\varphi-\eta')}(V_1+V_3)\Psi_{+\pm}(r,\varphi)\right\}, \end{aligned} \tag{64}$$

where we have used the same polar coordinates as in Sec. IV, i.e., $s_2 = -r\sin\varphi$, $t_2 = -r\cos\varphi$. For convenience, the indices q_i,i determining the initial state of scattering of the function $\Psi_{+\pm}(r,\varphi)$ have been dropped. Since $\Psi_{+\pm}(r,\varphi)$ are symmetric with respect to the line $s_1=0$, i.e., to the line $\varphi = -\pi/4$, $\varphi = 3\pi/4$, and since

$$\Psi_{+\pm}(r,\varphi) = \pm\Psi_{+\pm}(r,\varphi+\pi), \tag{65}$$

we see that the calculation of the integrals in (64) can be reduced to the calculation of the integrals over the angular regions $-\pi/4<\varphi<0$ and $0<\varphi<\pi/4$ in which the functions $\Psi_{+\pm}$ are solutions of Eqs. (33). By Eqs. (51) and (53) as applied to the Sommerfeld transforms $s_1^{\pm}(z)$ and $s_2^{\pm}$ defined in Sec. IV, and by the functional equations (36)–(39), we obtain the following results:

$$h_{vi}^{\pm}(q',q_i) = -\frac{N_i}{2}\left(\frac{g}{2\pi}\right)^{1/2}\frac{\sin\Theta}{\sin\eta'}\left[s_1^{\pm}(-\eta')-s_2^{\pm}(\eta')\pm s_1^{\pm}(\eta'-\pi)\mp s_2^{\pm}(-\eta'+\pi)\right], \tag{66}$$

where, as before, η' is related to q' according to (54).

The normalization constants are $N_1 = \left(\frac{g_1}{2\pi}\right)^{1/2}$ and $N_v = \left(\frac{g}{2\pi}\right)^{1/2}$.

Performing similar considerations for the amplitude $h_{1i}^{\pm}(q',q_i)$ as defined by (63), we obtain

$$h_{1i}^{\pm}(q',q_i) = -\left(\frac{g_1}{2\pi}\right)^{1/2}N_i\frac{\sin\Theta_1}{\sin\eta'}\left[s_1^{\pm}(\eta'-\pi/4)\pm s_1^{\pm}(-\eta'+3\pi/4)\right]. \tag{67}$$

We recall that all the information concerning the initial state of the scattering process is contained in the functions $s_1^{\pm}(z)$ and $s_2^{\pm}(z)$. For $i=v$, we demand that $[s_1^{\pm}(z)-(z+\Theta)^{-1}]$ be regular for $-\pi/4<\mathrm{Re}\,z<0$ and $[s_2^{\pm}(z)-(z-\Theta)^{-1}]$ be regular for $0<\mathrm{Re}\,z<\pi/4$, whereas for $i=1$ the functions $s_2^{\pm}(z)$ are to be regular for $0<\mathrm{Re}\,z<\pi/4$ and $[s_1^{\pm}(z)-(z-\Theta_1+\pi/4)^{-1}]$ are to be regular for $-\pi/4<\mathrm{Re}\,z<0$.

The on-shell values of (66) and (67) are obtained by putting $\eta' = -\Theta$ or $\eta' = \Theta+\pi$ in (66) and $\eta' = -\Theta_1$ or $\eta' = \Theta_1+\pi$ in (67), the expressions for each of those two values being related to each other according to the symmetry properties of the amplitudes.

For completeness sake, we note that in the decomposition (61), we omitted the following combination of amplitudes:

$$h_{22}(q',q) - h_{23}(q',q) = \langle q',2|(E-H_0)^{-1}(V_1+V_3)\big||\Psi_{2,q}\rangle - |\Psi_{3,q}\rangle\big\rangle. \tag{68}$$

The above combination corresponds to that part of the wave function which, in the (s,t) plane representation, is antisymmetric with respect to the line $s_1=0$, i.e., to the function $\Psi_{-}(s,t)$ that corresponds to the "trivial" part of the scattering problem (see Sec. IV and Fig. 3). By subtracting Eq. (47) as written for the functions h_{22} and h_{33} and making use of the symmetry properties of the Born factors, we find that the amplitude (68) is the solution of one independent Faddeev equation. In a similar way to that described above, the amplitude (68) can be further decomposed into its symmetric and antisymmetric parts and related to their Sommerfeld transforms as given by (40) and (41).

VI. CONCLUSIONS

In this paper we presented some results of the application of the Sommerfeld–Maluzhinetz transformation, as known in the wave diffraction theory, to the one-dimensional three-body problems with zero-range interactions. By this approach, the three-body problem reduces to the problem of solving a system of functional equations for the Sommerfeld transforms of the wave function. In the case of the boundary condition model, and also the related symmetric boundary condition model with a two-body delta potential, the functional equations may be solved by the application of Maluzhinetz methods. However, for more complicated models such as the delta potential model considered in Sec. IV (the author's main interest), we have to look for more advanced function-theoretic methods. Putting aside the questions of solvability of functional equations, it is shown that the applied approach is attractive for the close connections of the Sommerfeld transforms of the wave function with the off-shell extensions of the elements of the three-body scattering matrices. The above connections show (in an indirect way) that the reformulation of the Schroedinger equation through the Sommerfeld transformation can be also considered as the reformulation of the Faddeev–Lovelace equations. Knowledge of the Sommerfeld transforms of the wave functions provides directly all the information expected from the solution of the quantum-mechanical three-body problem; namely, the probability amplitudes for all possible scattering processes as well as the discrete spectrum of the Schroedinger equation. The binding energies of the three-body bound states may be found by investigating those poles of the half-off-shell amplitudes whose position is independent of the final momentum coordinate. Also, the matrix elements of the time delay operator may be expressed in terms of the Sommerfeld transforms by the application of the connections of those elements with the S-matrix elements and their energy derivatives.[26] In all likelihood more could be said about the quantities mentioned above by relying solely on the functional properties demanded by the Sommerfeld transforms without the necessity of referring to the analytic solutions.

ACKNOWLEDGMENT

The author is indebted to Professor J. Stecki for his interest in this work.

[1] R. Jost, Zeitschr. Angew. Math. Phys. **6**, 316 (1955).
[2] E. Lieb and H. Koppe, Phys. Rev. **116**, 367 (1959).
[3] H.M. Nussenzweig, Proc. R. Soc., Ser. A, **26**, 408 (1962).
[4] J.B. McGuire, J. Math. Phys. **5**, 622 (1964).
[5] S. Albeviero, Helv. Phys. Acta **40**, 135 (1967).
[6] J.B. McGuire and C.A. Hurst, J. Math. Phys. **13**, 1595 (1972).
[7] K. Lipszyc, Acta Phys. Polon. A **42**, 571 (1972).
[8] K. Lipszyc, Acta Phys. Polon. A **44**, 115 (1973).
[9] M. Gaudin and B. Derrida, J. Physique **36**, 1183 (1975).
[10] W.E. Williams, Proc. R. Soc. ser.A **252**, 376 (1959).
[11] G.D. Maluzhinetz, Akust. Zhurn. **1**, 144 (1955).
[12] G.D. Maluzhinetz, Dokl. Akad. Nauk SSSR **121**, 436 (1958).
[13] G.D. Maluzhinetz, Ann. Phys. (Germany) **6**, 107 (1960).
[14] K. Lipszyc, J. Math. Phys. **15**, 133 (1974).
[15] *Electromagnetic and Acoustic Scattering by Simple Shapes*, edited by J.J. Bowman, T.B.A. Senior and P.L.E. Uslenghi (North-Holland, Amsterdam, 1969), p. 252.
[16] G.D. Maluzhinetz, Dokl. Akad. Nauk. SSSR **118**, 1099 (1958); **119**, 49 (1958).
[17] K. Lipszyc, Phys. Rev. D **11**, 1649 (1975).
[18] T.B.A. Senior, Commun. Pure Appl. Math. **12**, 337 (1959).
[19] E.W. Schmid and H. Ziegelmann, *The Quantum-Mechanical Three-Body Problem* (Vieweg and Sohn, Braunschweig, 1974), Chap. V.
[20] L.R. Dodd, Australian J. Phys. **25**, 507 (1972).
[21] L.R. Dodd, J. Math. Phys. **11**, 207 (1970); Phys. Rev. D **3**, 2536 (1971).
[22] C.K. Majumdar, J. Math. Phys. **13**, 705 (1972).
[23] D.D. Brayshaw and R.F. Peierls, Phys. Rev. **177**, 2539 (1969).
[24] With obvious changes, we can consider in an analogous way the regions consisting of the sectors $-\Phi_1 < \varphi < 0$ and $0 < \varphi < \Phi_2$, with $\Phi_1 \neq \Phi_2$.
[25] We take the opportunity to supplement the results of Ref. 17 by the functional relations between the amplitudes h_{ij} and the Sommerfeld transform $s(z)$ for the boundary condition model. By the definitions of parameters and variables used in Ref. 17, the said relations read as follows: $h_{3i}(q,p) = N_i(-\alpha_3/\pi)^{1/2}(\sin\eta - \sin\Theta_3)(\sin\eta)^{-1}s(\eta-\Phi)$ and $h_{1i}(q,p) = -N_i(-\alpha_1/\pi)^{1/2}(\sin\eta - \sin\Theta_1)(\sin\eta)^{-1}s(\eta+\Phi-\pi)$, where $N_i = (-\alpha_i/\pi)^{1/2}$, $i = 1,3$.
[26] T.A. Osborn and D. Bollé, J. Math. Phys. **16**, 1533 (1975).

Planar limit for SU(N) symmetric quantum dynamical systems

G. Marchesini and E. Onofri
Istituto di Fisica dell'Università, 43100 Parma, Italy and INFN, Sezione di Milano, Milano, Italy

(Received 8 March 1979; accepted for publication 18 May 1979)

The spectrum of a quantum mechanical Hamiltonian with SU(N) symmetry is studied in the limit $N \to \infty$. A complete description is given for the states which transform according to (i) the scalar representation (singlet states) and (ii) the adjoint representation of SU(N) ("adjoint" states). The eigenvalues of the singlet states are equally spaced with a finite gap $\omega(g)$ as $N \to \infty$. The spectrum of the adjoint states is equally spaced asymptotically for large excitations with the same gap $\omega(g)$. The first excitation of the adjoint states is lower than the first singlet excitation. An accidental degeneracy appears which is removed by $1/N$ corrections. We compute explicitly this splitting for the singlet states. In this formulation $1/N^2$ plays a role similar to $\hbar$ so that all computed quantities are related to the corresponding classical system. All the quantities which enter in the calculation of the spectrum are analytic near $g = 0$ with the same radius of convergence.

I. INTRODUCTION

The topology of graphs emerges as an important concept for Strong Interactions. In fact it has been suggested[1,2] that already the planar amplitudes should describe the main dynamical features. Let us recall some of these indications. For a SU(N) gauge theory, in the limit $N \to \infty$, the Feynman graphs are arranged into sets which have explicitly the same topology of the quantum dual string with quarks at its ends.[1] This suggests that the color confinement may take place already at the planar level. This hope is actually realized in the special case of two space–time dimensions.[3] A second indication of the relevance of the planar amplitudes is given by the recent calculation[4] of hard processes in Q.C.D.: At any given order of perturbation theory the leading contribution is given, in the axial gauge, by sets of planar graphs. Finally, the analogy of planar amplitudes with exchange degenerate Regge poles suggests a way to sum[5] all the terms of the topological expansion with Gribov's Regge Field Theory, thus leading to a unified picture[6] of various attempts in the study of strong interactions.

Due to these promising features one would like to develop nonperturbative methods for dealing with the large N limit of SU(N) gauge theory. This has been done[3] only in the special case of two space–time dimensions where in the axial gauge, the topology of diagrams is trivial. In order to learn some methods which could be applied in general, attempts have been made to study the limit $N \to \infty$ of SU(N) symmetric scalar theories, defined by a Lagrangian of the form

$$\mathscr{L} = \mathrm{Tr}\,(\tfrac{1}{2}\partial_\mu \varphi \partial^\mu \varphi - (N/g)\,V((g/N)\,\varphi^2))$$

φ being a Hermitian $N \times N$ matrix.

A preliminary problem was the study of the multiplicity of graphs[7-9] (zero space–time dimension). A lesson one learns here is that, since $1/N^2$ plays a role similar to $\hbar$, the planar amplitudes can be studied in terms of equations of the classical type.

The second step is the study of quantum mechanical systems with SU(N) symmetry. The ground state energy has been computed[9] and again one finds that only classical equations are involved. This calculation however does not involve the full structure of SU(N) since the "angular variables" of SU(N) enter in a trivial way.

In this paper we want to continue the analysis of the quantum mechanical system and study the spectrum of nonsinglet states, where the full structure of SU(N) is involved. Actually we will limit ourselves to study the spectrum of states which belong to the adjoint representation of SU(N), but the method we develop can be applied, in principle, to higher representations. Essentially the problem is to make a partial wave expansion with respect to SU(N) in order to reduce the Schrödinger equation to a simpler "radial" form. For singlet states one obtains a separable equation,[9] equivalent to that describing N noninteracting fermions in a common anharmonic potential. For $N \to \infty$ the spectrum is then equally spaced with a finite gap $\omega(g)$ which can be interpreted as the frequency of the classical orbit at the Fermi energy. An accidential degeneracy appears which is removed by the first $1/N$ correction: The splitting is again expressed in terms of classical quantities.

For the adjoint representation, the reduced equation is actually given by a system of coupled differential equations which are not separable. We find however that $1/N^2$ is a natural expansion parameter, since the nonseparable part can be treated as a perturbation, where the leading term reduces to the singlet state equation. For large N the spectrum of the adjoint states is then given by ordinary perturbation theory of degenerate eigenvalues. This problem can be cast in the form of a singular integral equation whose eigenvalues give directly the spectrum of the adjoint states up to $O(1/N)$. The result of this analysis is that the gaps of these states with the ground state are finite as $N \to \infty$ and the first excited state of the adjoint representation is lower than the first excited singlet state. For large excitations, the adjoint states become equally spaced with the same gap $\omega(g)$ of the singlet states. The results we present here are for the potential $V(\sigma^2) = \frac{1}{2}\sigma^2 + \sigma^4$ but the general structure of the spectrum is the same provided that the classical system has only periodic orbits.

After defining in Sec. 2 our SU(N) quantum mechanical system, we discuss the full spectrum of the singlet states (Sec.

 0022-2488/80/051103-08$01.00

3) and of the adjoint representation (Sec. 4). Sec. 5 contains the summary and some comments.

2. THE QUANTUM MECHANICAL SYSTEM

We want to study the spectrum of the Hamiltonian

$$H = \mathrm{Tr}(\tfrac{1}{2}\pi^2 + \tfrac{1}{2}\varphi^2 + (g/N)\,\varphi^4), \tag{2.1}$$

where φ is a Hermitian $N \times N$ matrix and $\mathrm{Tr}(\frac{1}{2}\pi^2)$ is the kinetic energy, namely

$$\mathrm{Tr}\,\pi^2 = -\sum_{ij} \frac{\partial^2}{\partial\varphi_{ij}\partial\varphi_{ji}}. \tag{2.2}$$

H is an operator in the Hilbert space L_2 of square integrable functions with respect to the invariant volume element. $[d\varphi] = \Pi\, d\varphi_{ii}\, \Pi(2i)^{-1} d\varphi_{ij} d\varphi_{ji}$. The Hamiltonian is invariant under the SU(N) transformations $\varphi \rightarrow U\varphi U^\dagger$. This implies that to every energy level there corresponds a unitary irreducible representation of SU(N)/Z_N. The most natural choice of coordinates to study the Schroedinger equation for the Hamiltonian (2.1) is therefore given by the "polar" representation of φ

$$\varphi \equiv U \Lambda U^\dagger, \tag{2.3}$$

where $\Lambda = \mathrm{diag}[\lambda_1,...,\lambda_N]$ and U belongs to the coset space SU(N)/**H**, **H** being the stability subgroup of Λ under SU(N). Except for a set of measure zero, **H** is the Cartan subgroup of diagonal unitary matrices with determinant one. The kinetic energy (2.2), i.e., the Laplace operator, can be easily obtained in polar coordinates starting from the line element

$$ds^2 = \mathrm{Tr} d\varphi^2 = \sum_i^N d\lambda_i^2 + \sum_{i,j} (\lambda_i - \lambda_j)^2 \, |(U^\dagger dU)_{ij}|^2 \tag{2.4}$$

and it is given by

$$-\Delta = -\frac{1}{V(\lambda)} \sum_1^N \left(\frac{\partial}{\partial\lambda_i}\right)^2 V(\lambda) + \sum_{i<j} (\lambda_i - \lambda_j)^{-2} M_{ij} \left(\frac{\partial}{\partial\vartheta_\alpha}\right), \tag{2.5}$$

where $V(\lambda) = \Pi_{i<j}(\lambda_i - \lambda_j)$ and M_{ij} is a set of noncommuting differential operators with respect to the angular coordinates ϑ_α ($\alpha = 1,..., N^2 - N$) which parametrize the coset space SU(N)/**H**. For $N = 2$, M_{12} is the ordinary total angular momentum squared. We shall not need an explicit characterization of the parameters ϑ_α, since all our calculations will be coordinate-free.

The invariant volume element $[d\varphi]$ is given by $[d\varphi] = V(\lambda)^2 d\lambda_1 \cdots d\lambda_N \cdot [dU]$, where $[dU]$ is the invariant volume element over the coset space SU(N)/**H**. Since in the following we shall integrate over **H**-invariant functions, $[dU]$ will be extended to the invariant volume over SU(N).

3. SINGLET STATES

The wave functions for SU(N)-invariant states are given by

$$\Psi_s(\varphi) = \psi(\lambda_1,...,\lambda_N), \tag{3.1}$$

where ψ is totally symmetric under permutation of the eigenvalues λ_i. From (2.5) the angular part in the kinetic term vanishes thus simplifying the Schrödinger equation which turns out to be separable[9]

$$\sum_1^N \left(-\frac{1}{2}\left(\frac{\partial}{\partial\lambda_i}\right)^2 + \frac{1}{2}\lambda_i^2 + \frac{g}{N}\lambda_i^4 \right) \Phi(\lambda) = E_s \Phi(\lambda), \tag{3.2}$$

where

$$\Phi(\lambda) = V(\lambda)\psi(\lambda). \tag{3.3}$$

Since Φ is totally antisymmetric, the problem is reduced to finding the energy levels for a system of N noninteracting fermions in the anharmonic potential $\frac{1}{2}\lambda^2 + g/N\,\lambda^4$. Let $e_n(g)$ be the nth energy level of the anharmonic oscillator $\frac{1}{2}(p^2 + x^2) + gx^4$ and $u_n(g,x)$ the corresponding eigenfunction.

The ground state is obtained by filling up the first N levels

$$\Phi_0(\lambda) = \frac{1}{\sqrt{N!}} \det\|u_n(g/N, \lambda_i)\|_1^N \tag{3.4}$$

with ground state energy $E_0(g,N)$ given by

$$E_0(g,N) = \sum_1^N e_i(g/N). \tag{3.5}$$

The leading term of the $1/N$ expansion of (3.5)

$$E_0(g,N) = N^2 \epsilon_0(g) + \eta_0(g) + O(1/N^2) \tag{3.6}$$

was obtained in Ref. (9) by the semiclassical method. For latter convenience we rederive their result. For large N the leading term in (3.6) is given by the eigenvalues $e_n(g/N)$ with large n for which we can apply the semiclassical approximation

$$(1/2\pi) \oint dx \sqrt{2e_n(g/N) - x^2 - 2gx^4/N} = n - \tfrac{1}{2} \quad (n = 1,2,\cdots). \tag{3.7}$$

It follows that

$$e_n(g/N) \simeq \frac{N}{g}\, \Omega\,((n - \tfrac{1}{2})g/N), \tag{3.8}$$

where $\Omega(\tilde{g})$ is obtained by inverting the equation

$$\tilde{g} = (1/2\pi) \oint dx \sqrt{2\Omega(\tilde{g}) - x^2 - 2x^4}. \tag{3.9}$$

The sum in Eq. (3.5) can be approximated with an integral (via Euler–MacLaurin expansion) yielding

$$E_0(g,N) \sim \int_1^N dn\,(N/g)\,\Omega\,(ng/N) = (N^2/g^2) \int_0^g \Omega(g')\, dg'. \tag{3.10}$$

Considering the first singlet excitation, we have

$$E_1(g,N) = E_0(g,N) + e_{N+1}(g/N) - e_N(g/N) \tag{3.11}$$

obtained by exciting the last level of our N-fermion system. It follows from Eq. (3.8)

$$E_1(g,N) - E_0(g,N) \simeq \frac{d\Omega(g)}{dg} \equiv \omega(g). \tag{3.12}$$

Notice that $\omega(g)$ is just the frequency of the classical orbit in the anharmonic potential at the Fermi energy $e_N(g/N)$:

$$\omega(g) = \{(1/2\pi) \oint dx (2\Omega(g) - x^2 - 2x^4)^{-1/2}\}^{-1}. \tag{3.13}$$

TABLE I. $1/N$ fine structure of singlet states
$\omega(1000) = 18.37130$; $\omega'(1000) = .00612$; $N = 5$, $g = 1000$

$\{k,q\}$	$E_{\{k,q\}} - E_0$ (Eq. 3.19)	Exact value[15] (Eq. 3.15)
$\{1,0\}$	18.37130	18.35163
$\{2,0\}$	37.96497	37.85734
$\{1,1\}$	35.52023	35.38035
$\{3,0\}$	58.78101	58.39384
$\{2,1\}$	55.11390	54.88606
$\{1,2\}$	51.44685	50.83442

A general singlet state is characterized by a set of integers

$$0 \leqslant q_1 < q_2 < \cdots < q_r, \qquad 1 \leqslant k_1 < k_2 < \cdots < k_r \tag{3.14}$$

$$\begin{aligned} E_{\{q,k\}}(g,N) &= E_0 + \sum_1^r (e_{N+k_i} - e_{N-q_i}) \\ &\simeq E_0 + \omega(g) \sum_1^r (k_i + q_i), \end{aligned} \tag{3.15}$$

which shows that the singlet spectrum is equally spaced with a gap given by the classical frequency $\omega(g)$ and a degeneracy given by the partitions of $\omega(g)^{-1}(E - E_0)$ into integers satisfying the inequalities (3.14). This degeneracy is "accidental" and it is broken by the $1/N$ corrections.

To calculate higher order corrections in the $1/N$ expansion of the energy gaps, one has to calculate higher order WKB approximations to the eigenvalues $e_n(g)$. The first nonleading term is given by

$$e_n(g/N) = (N/g)\,\Omega\left((n - \tfrac{1}{2})g/N\right) + \tfrac{1}{2}(g/N)\, e_n^{(2)} + O(g/N). \tag{3.16}$$

$(ng/N \sim \text{constant})$,

where $e_n^{(2)}$ is obtained through the well-known relation[10]

$$e_n^{(2)} = -\frac{1}{12}\,\omega(\tilde{g}) \left(\frac{d}{d\Omega}\right)^2 \oint \frac{dx}{2\pi}(1 + 12x^2) \times \sqrt{2\Omega - x^2 - 2x^4}\,\Big|_{\Omega = \Omega(\tilde{g})} \quad (\tilde{g} = (n - \tfrac{1}{2})g/N) \tag{3.17}$$

which gives

$$e_n(g/N) \simeq (N/g)\,\Omega(\tilde{g}) + \frac{1}{2}\frac{g}{N}\left[\frac{1}{12}\,\omega(\tilde{g})^{-2}\omega'(\tilde{g}) - \omega(\tilde{g}) + 3\tilde{g}\omega'(\tilde{g})\right]. \tag{3.18}$$

Equation (3.15) is thus modified into the following expression which exhibits the splitting of degenerate singlet levels

$$E_{\{q,k\}} - E_0 = \omega(g) \sum_{i=1}^r (k_i + q_i) + \frac{1}{2}\frac{g}{N}\,\omega'(g) \times \sum_{i=1}^r (k_i + q_i)(k_i - q_i - 1). \tag{3.19}$$

Notice that the correction is of order $1/N$ instead of $1/N^2$ as expected (except for the first, nondegenerate excited state). This may be understood as a consequence of the accidental degeneracy and *it does not spoil* the correct $1/N^2$ expansion of Green functions, since levels with $1/N$ corrections appear in pairs with opposite sign ($E_{\{q,k\}}$ and $E_{\{k-1,q+1\}}$). The eigenvalue formula (3.19) gives a good approximation to the eigenvalues of H even for N small. See Table I for a numerical example with $N = 5$.

To calculate the $1/N^2$ correction one should consider the third order WKB term.[10] From the expression of the general k th term in the WKB expansion of $e_n(g)$, only the function $\Omega(g)$ and its derivatives appear (see for instance Eq. 3.18). This fact entails that all the terms in the $1/N$ expansion have a singularity in the complex g plane where $\Omega(g)$ is singular, namely at $g = g_c = -2^{1/2}/3\pi$.[9] (Notice that $\omega(g_c) = 0$, i.e., the classical motion has infinite period, which corresponds to the instanton of the anharmonic oscillator.)

Since $e_n(g/N)$ is nonanalytic in $g = 0$ (at fixed N), one infers that the $1/N$ expansion is an asymptotic expansion[11] with coefficients which have the same radius of convergence in g.

4. STATES BELONGING TO THE ADJOINT REPRESENTATION

The spectrum for nonsinglet states involves the angular operators M_{ij} in Eq. 2.6. For $N > 2$, M_{ij} are noncommuting operators, thus the partial wave decomposition of the Schrödinger equation becomes highly nontrivial. For the adjoint representation, however, the angular dependence of the wave function can be explicity written down and the Schroedinger equation can be reduced to radial variables.

A. Wave function angular dependence

Let us first express the angular dependence of the general wave function for the adjoint representation. In this subsection we shall prove that the general form $\Psi_A(\varphi)$ of a $N^2 - 1$ multiplet is given by

$$\Psi_A(\varphi) = \sum_{n=2}^N \mathrm{Tr}(C\varphi^{n-1})\,\psi_n(\lambda_1,\ldots,\lambda_N), \tag{4.1}$$

where $\psi_n(\lambda)$ are symmetric functions of the eigenvalues of the matrix φ, and C is any traceless matrix which gives the $N^2 - 1$ multiplicity. In Eq. (4.1) the angular dependence is explicit in the factor $\mathrm{Tr}(C\varphi^{n-1})$.

Let us come to the proof of Eq. (4.1). A general element $\Theta_{ij}(\varphi)$ of the $N^2 - 1$ multiplet must transform according to

$$\Theta_{ij}(V^\dagger\varphi V) = V^\dagger{}_{ik}\,\Theta_{kl}(\varphi)\,V_{lj}, \qquad V \in \mathrm{SU}(N), \tag{4.2}$$

$$\sum \Theta_{ii}(\varphi) \equiv 0$$

and the general linear combination of Θ_{ij} is then given by

$$\Psi_A(\varphi) = \mathrm{Tr}(C\Theta(\varphi)), \qquad \mathrm{Tr}C = 0. \tag{4.3}$$

For the special case of $\varphi = \Lambda$ (diagonal), Eq. (4.2) implies that also $\Theta(\Lambda)$ is diagonal:

$$\Theta_{ij}(\Lambda) = \delta_{ij}\,\chi_i(\lambda_1,\ldots,\lambda_N). \tag{4.4}$$

Going into polar coordinates for φ, we may write Eq. (4.3) in the form

$$\Psi_A(\varphi) = \sum_{k=1}^N (U^\dagger C U)_{kk}\,\chi_k(\lambda_1,\ldots,\lambda_N), \qquad \varphi = U\Lambda U^\dagger. \tag{4.5}$$

Introducing $\psi_n(\lambda_1,...,\lambda_N)$ with

$$\chi_k(\lambda) = \sum_{n=1}^{N} \lambda_k^{n-1} \psi_n(\lambda) \tag{4.6}$$

we obtain the result in Eq. (4.1). Notice that ψ_1 is not independent of $\psi_2,...,\psi_N$; by $\mathrm{Tr}C = 0$, however, ψ_1 does not contribute to $\psi_A(\varphi)$. The fact that $\psi_n(\lambda)$ are symmetric functions is a consequence of Eq. (4.4). In fact under a permutation P of the eigenvalues we have

$$\chi_k(P\lambda) = \chi_{Pk}(\lambda) \Rightarrow \psi_n(P\lambda) = \psi_n(\lambda) \tag{4.7}$$

B. Radial Schrödinger equation

Let us consider the Schrödinger equation for the adjoint states characterized by Eq. (4.1)

$$(-\tfrac{1}{2}\Delta + V(g/N,\varphi) - E)\Psi_A(\varphi) = 0, \tag{4.8}$$

where the potential is given in 2.1 and is invariant with respect to SU(N). We shall now reduce Eq. (4.8) to a system of coupled equations in the unknown functions $\psi_n(\lambda)$. For the kinetic term we have

$$\Delta\Psi_A = \sum_{ij} \frac{\partial}{\partial\varphi_{ij}} \frac{\partial}{\partial\varphi_{ji}} \Psi_A = \sum_{n=2}^{N} \Big\{ \mathrm{Tr}C\varphi^{n-1} \Delta\psi_n + \psi_n \Delta \,\mathrm{Tr}\, C\varphi^{n-1} + 2 \frac{\partial}{\partial\varphi_{ij}} \mathrm{Tr}\, C\varphi^{n-1} \frac{\partial\psi_n}{\partial\varphi_{ji}} \Big\} \tag{4.9}$$

and we have to extract the explicit angular dependence. The first term is simple, since only the radial part of the Laplace operator is involved in $\Delta\psi_n(\lambda)$. For the second term we have

$$\Delta\, \mathrm{Tr}C\varphi^{n-1} = 2 \sum_{h=0}^{n-2} (n-h-2)\, \mathrm{Tr}\varphi^h\, \mathrm{Tr}C\varphi^{n-h-3}. \tag{4.10}$$

The last term is evaluated as follows:

$$\frac{\partial\psi_n(\lambda)}{\partial\varphi_{ij}} = \sum_{k,p=1}^{N} \frac{\partial\psi_n}{\partial\lambda_k} \frac{\partial\lambda_k}{\partial\tau_p} \frac{\partial\tau_p}{\partial\varphi_{ij}}, \qquad \tau_p \equiv \mathrm{Tr}\varphi^p. \tag{4.11}$$

Introducting the Vandermonde matrix

$$V_{pk}(\lambda) \equiv \lambda_k^{p-1}, \tag{4.12}$$

we have

$$\frac{\partial\lambda_k}{\partial\tau_p} = p^{-1}(V^{-1})_{kp}, \tag{4.13}$$

and finally

$$\frac{\partial}{\partial\varphi_{ij}} \mathrm{Tr}C\varphi^{n-1} \frac{\partial\psi_n}{\partial\varphi_{ji}} = (n-1) \sum_{k=1}^{N} (U^\dagger CU)_{kk}\, \lambda_k^{n-1} \frac{\partial\psi_n}{\partial\lambda_k}. \tag{4.14}$$

Schrödinger equation is then reduced to the form

$$\sum_{k=1}^{N} (U^\dagger CU)_{kk} A_k[\psi] = 0, \tag{4.15}$$

where

$$A_k[\psi] \equiv \sum_{n=2}^{N} \Big\{ \lambda_k^{n-1} [-\tfrac{1}{2}\Delta + V(g/N,\lambda) - E]\, \psi_n(\lambda) - (n-1)\, \lambda_k^{n-3} \frac{\partial\psi_n}{\partial\lambda_k} - \sum_{h=0}^{n-3} (n-h-2) \times \lambda_k^{n-h-3}\, \mathrm{Tr}\varphi^h\, \psi_n(\lambda) \Big\}. \tag{4.16}$$

Since $\Sigma(U^\dagger CU)_{kk} = \mathrm{Tr}C = 0$, the most general solution of Eq. (4.15) is given by

$$A_1[\psi] = A_2[\psi] = \cdots = A_N[\psi]. \tag{4.17}$$

A more convenient form of these $N-1$ equations is given by

$$\sum_{k=1}^{N} A_k[\psi] V^{-1}{}_{kn}(\lambda) = 0, \qquad n = 2,3,...,N, \tag{4.18}$$

which explicitly reads

$$[-\tfrac{1}{2}\Delta + V(g/N,\lambda) - E]\, \psi_n(\lambda) = \sum_{m=2}^{N} G_{nm}\Big(\lambda, \frac{\partial}{\partial\lambda}\Big) \psi_m(\lambda), \tag{4.19}$$

where

$$G_{mn}\Big(\lambda, \frac{\partial}{\partial\lambda}\Big) = \sum_{k=1}^{N} \Big\{ n\lambda_k^{m-n-2} + (m-1)\, \lambda_k^{m-2} \times V_{kn}^{-1}(\lambda) \frac{\partial}{\partial\lambda_k} \Big\} \tag{4.20}$$

(it is understood that the first term in the bracket is absent for $m < n+2$).

Even in the simplest case ($N = 2$) the reduced (radial) Schrödinger equation is nonseparable, unlike the singlet case.

C. Perturbation theory in *1/N*

We have just shown that the Schrödinger equation (4.19)–(4.20) for the adjont states is nonseparable. We can overcome this difficulty by treating the operator $G_{n,m}$ as a perturbation in $1/N$. In fact, while the ground state energy $E_0(g,N)$ grows like N^2, we are able to prove that the gaps for the adjoint states remain finite as $N \to \infty$. For this reason, to leading order in $1/N$, the spectrum of the states $\Psi_A(\varphi)$ can be obtained by perturbation theory of degenerate levels starting from the unperturbed states

$$\Psi_A^{(0)} = \psi_s(\lambda) \sum_{n=2}^{N} \alpha_n\, \mathrm{Tr}(C\varphi^{n-1}), \tag{4.21}$$

where all the radial wave functions $\psi_n(\lambda)$ are proportional to the same wavefunction $\psi_s(\lambda)$ of any singlet state, with energy E_s. In fact the left-hand side of Eq. (4.9) vanishes for $\psi_n = \psi_s$, $E = E_s$. To first order in $1/N^2$ the spectrum is then obtained by diagonalizing the expectation value of the Hamiltonian with respect to the $N-1$ parameters α_n. The expectation value is calculated in the appendix; it holds:

$$\frac{\langle \Psi_A^{(0)} | H | \Psi_A^{(0)} \rangle}{\langle \Psi_A^{(0)} | \Psi_A^{(0)} \rangle} = E_s + \frac{1}{2} \frac{\Sigma_{n,n'}\, \alpha_n \bar{\alpha}_{n'}\, \mathcal{N}_{nn'}}{\Sigma_{n,n'}\, \alpha_n \bar{\alpha}_{n'}\, \mathcal{D}_{nn'}}, \tag{4.22}$$

where

$$\begin{aligned} \mathcal{N}_{nn'} &= \sum_{r=0}^{n-2} \sum_{r'=0}^{n'-2} \langle \psi_s | \{ \mathrm{Tr}\varphi^{n-r+r'-2}\, \mathrm{Tr}\varphi^{n'-r'+r-2} - (1/N)\, \mathrm{Tr}\varphi^{n+n'-4} \} | \psi_s \rangle, \\ \mathcal{D}_{nn'} &= \langle \psi_s | \{ \mathrm{Tr}\varphi^{n+n'-2} - (1/N) \mathrm{Tr}\varphi^{n-1}\, \mathrm{Tr}\varphi^{n'-1} \} | \psi_s \rangle. \end{aligned} \tag{4.23}$$

Now we have to compute the asymptotic behavior of $\mathcal{N}$ and $\mathcal{D}$ for N large and prove that $E_A - E_s$ is finite. For simplicity, we take for ψ_s the ground state Ψ_0 (Eq. 3.4). $\mathcal{N}$ and $\mathcal{D}$ are given in terms of expectation values of $\mathrm{Tr}\varphi^{2p}$ and

$\mathrm{Tr}\varphi^p\ \mathrm{Tr}\varphi^q$ over Slater–Fock wave functions. The expectation value of $\mathrm{Tr}\varphi^{2p}$ is given by

$$\langle \Psi_0 | \mathrm{Tr}\varphi^{2p} | \Psi_0 \rangle = \sum_{n=1}^{N} \int_{-\infty}^{\infty} dx\, x^{2p} u_n (g/N,x)^2 \equiv \sum_{n=1}^{N} \langle n | x^{2p} | n \rangle, \quad (4.24)$$

where $u_n(g/N,x)$ are the anharmonic oscillator eigenfunctions considered in Sec. 3. The leading contribution in N comes from large n terms, which can be estimated by the semiclassical approximation[12]

$$\langle n | x^{2p} | n \rangle \simeq \frac{1}{T(e_n)} \int_0^{T(e_n)} dt\, x_{cl}^{2p}(e_n,t), \qquad (4.25)$$

where $e_n = e_n(g/N) \simeq N\Omega(ng/N)/g$ is the nth eigenvalue and $\Omega(g)$ was introduced in Eq. (3.9). $T(e_n)$ and $x_{cl}(e_n,t)$ are the period and the trajectory of the classical motion at the energy e_n. We find then

$$\langle n | x^{2p} | n \rangle \simeq \frac{d}{dn} \oint \frac{dx}{2\pi} x^{2p} \sqrt{2e_N - x^2 - 2gx^4/N}. \quad (4.26)$$

By scaling the variable x as suggested by Eq. (4.26), we find

$$\langle n | x^{2p} | n \rangle \simeq \left(\frac{N}{g}\right)^p \omega(\bar{g}) \times \oint \frac{dx}{2\pi} x^{2p} (2\Omega(\bar{g}) - x^2 - 2x^4)^{-1/2} \Big|_{\bar{g} = ng/N}, \quad (4.27)$$

which shows that this matrix element scales as N^p times a function of n/N. Finally we obtain

$$\langle \Psi_0 | \mathrm{Tr}\varphi^{2p} | \Psi_0 \rangle \simeq \int_1^N dn\, \langle n | x^{2p} | n \rangle \simeq \left(\frac{N}{g}\right)^{p+1} \oint \frac{dx}{2\pi} x^{2p} \times \sqrt{2\Omega(g) - x^2 - 2x^4}. \quad (4.28)$$

The calculation of the leading behavior of the ground state expectation value of the product $\mathrm{Tr}\varphi^p\ \mathrm{Tr}\varphi^q$ can be reduced to the previous case. In fact we have

$$\langle \mathrm{Tr}\varphi^p \mathrm{Tr}\varphi^q \rangle = \langle \mathrm{Tr}\varphi^p \rangle \langle \mathrm{Tr}\varphi^q \rangle + \sum_{n=1}^{N} \sum_{m=N+1}^{\infty} \langle n | x^p | m \rangle \langle m | x^q | n \rangle. \quad (4.29)$$

For p and q even, the first term is of order $N^{(1/2)(p+q)+2}$ [by Eq. (4.28)]. The correlation term in Eq. (4.29) is nonleading in N, as follows from elementary inequalities:

$$\begin{aligned} \Big| \sum_{n=1}^{N} \sum_{m=n+1}^{\infty} \langle n | x^p | m \rangle \langle m | x^q | n \rangle \Big| &\leqslant \sum_{n=1}^{N} \Big| \sum_{m=N+1}^{\infty} \langle n | x^p | m \rangle \langle m | x^q | n \rangle \Big| \\ &\leqslant \sum_{n=1}^{N} \sqrt{\langle n | x^{2p} | n \rangle \langle n | x^{2q} | n \rangle} \\ &\leqslant \sqrt{\sum_{n=1}^{N} \langle n | x^{2p} | n \rangle \sum_{m=1}^{N} \langle m | x^{2q} | m \rangle} \\ &= O(N^{1/2(p+q)+1}). \end{aligned} \quad (4.30)$$

Presumably one can do better and prove that the correlations are of order $N^{(1/2)(p+q')}$ as for the pure harmonic oscillator. In fact the matrix elements $\langle n | x^p | m \rangle$ vanish very rapidly for large $|n-m|$ so that the sum in Eq. (4.29) is restricted to a narrow strip near $n \sim m \sim N$.

In conclusion we have that the leading contributions in N to $\mathscr{N}$ and $\mathscr{D}$ come from the terms containing traces of even powers of φ and the correlation terms can be neglected:

$$\begin{aligned} \mathscr{N}_{nn'} &\sim \sum_{r=0}^{n-2} \sum_{r'=0}^{n'-2} \langle \mathrm{Tr}\varphi^{n-r+r'-2} \rangle \langle \mathrm{Tr}\varphi^{n'-r'+r-2} \rangle \\ &= O(N^{(1/2)(n+n')}), \\ \mathscr{D}_{nn'} &\sim \langle \mathrm{Tr}\varphi^{n+n'-2} \rangle - (1/N) \langle \mathrm{Tr}\varphi^{n'-1} \rangle \langle \mathrm{Tr}\varphi^{n'-1} \rangle \\ &= O(N^{(1/2)(n+n')}). \end{aligned} \quad (4.31)$$

Actually this asymptotic form, proven for the ground state expectation values, holds for any wave function ψ_s.

The asymptotic estimate (4.31) greatly simplifies the structure of the expectation value of the Hamiltonian and shows that the gaps for the adjoint states are finite as $N \to \infty$.

D. Diagonalization of the secular determinant

Now we can proceed to the actual calculation of the spectrum of the adjoint representation. Let us start from the states which are built perturbatively over the gound state. From Eqs. (4.28)–(31) we have

$$\mathscr{N}_{nn'} = \left(\frac{N}{g}\right)^{(1/2)(n+n')} \int \frac{dx}{\pi} \sigma(g,x) \times \int \frac{dy}{\pi} \sigma(g,y) \frac{(x^{n-1} - y^{n-1})(x^{n'-1} - y^{n'-1})}{(x-y)^2}, \quad (4.32)$$

$$\mathscr{D}_{nn'} = \left(\frac{N}{g}\right)^{(1/2)(n+n')} \left(\int \frac{dx}{\pi} \sigma(g,x)\, x^{n+n'-2} - \frac{1}{g} \int \frac{dx}{\pi} \sigma(g,x)\, x^{n-1} \int \frac{dy}{\pi} \sigma(g,y)\, y^{n'-1} \right), \quad (4.33)$$

where we have introduced

$$\sigma(g,x) = \sqrt{2\Omega(g) - x^2 - 2x^4}\, \theta(2\Omega(g) - x^2 - 2x^4). \quad (4.33)$$

The sums over n,n' in Eq. (4.22) can now be performed explicitly by introducing the function

$$f(x) \equiv \sum_{n=2}^{N} (N/g)^{n/2} \alpha_n x^{n-1}. \quad (4.34)$$

We obtain

$$\frac{\langle \Psi_A^{(0)} | H | \Psi_A^{(0)} \rangle}{\langle \Psi_A^{(0)} | \Psi_A^{(0)} \rangle} \simeq E_0 + \mathscr{H}[f] \equiv E_0 + \frac{1}{2} \times \frac{\int (dx/\pi) \sigma(g,x) \int (dy/\pi) \sigma(g,y) [(f(x) - f(y))/(x-y)]^2}{\int (dx/\pi) \sigma(g,x) f(x)^2 - (1/g)(\int (dx/\pi)\, \sigma(g,x) f(x))^2} \quad (4.35)$$

(we can assume α_n to be real, without losing in generality). Before solving the variational problem for $\mathscr{H}[f]$, let us observe that a variational estimate of the first level can be easily obtained by putting $f(x) = x$; we obtain

$$E_{A1} - E_0 \leqslant [OV{:}[PW{:}g{:}2]{:}2\textstyle\int (dx/\pi)\, \sigma(g,x) [PW{:}x{:}2]] = \frac{(1/2)g^2}{7\int_0^g \Omega(g')dg' - 3g\Omega(g)}. \quad (4.36)$$

For example the asymptotic behavior for $g \to \infty$ is simply

TABLE II. The energy levels for the adjoint representation for various values of g (leading trajectory in Fig. 1).

g	n	$E_{A(n,0)}$	$(n+1)\omega(g)+q(g)$	
2	1	2.1281936	2.1636227	
	2	4.6201131	4.6176809	$\omega(g) =$ 2.4540582
	3	7.0716122	7.0717391	
	4	9.5258038	9.5257973	$q(g) = -2.7444937$
	5	11.9798556	11.9798555	
50	1	5.7594935	5.8946344	
	2	12.7149752	12.7041452	$\omega(g) =$ 6.8095108
	3	19.5129917	19.5136560	
	4	26.3232060	26.3231668	$q(g) = -7.7243871$
	5	33.1326773	33.1326775	
200	1	9.0832498	9.3025234	
	2	20.0835820	20.0657875	$\omega(g) =$ 10.763264
	3	30.8279452	30.8290517	
	4	41.5923818	41.5923159	$q(g) = -12.224005$
	5	52.3555793	52.3555801	
1000	1	15.4894625	15.8679888	
	2	34.2712440	34.2403381	$\omega(g) =$ 18.3723493
	3	52.6107716	52.6126875	
	4	70.9851687	70.9850368	$q(g) = -20.8767099$
	5	89.3574010	89.3573861	

$E_{A(n,0)}$: numerical results from Ref. 13.

given by

$$E_{A1} - E_0 \leqslant \frac{5\Gamma(1/4)^{16/3}}{489^{1/3}\pi^3} 9^{1/3} = 1.55440\, g^{1/3}. \tag{4.37}$$

[Notice that this estimate is already sufficient to conclude that E_{A1} is lower than the first singlet excited state, e.g., $\omega(g) \sim 1.83534\, g^{1/3}$ for $g \to \infty$.]

To calculate the energy levels in the adjoint representation we have now to solve the variational problem $\delta\mathscr{H}[f] = 0$, where the variation is with respect to the parameters α_n, or, equivalently, with respect to $f(x)$. Although $f(x)$ is a polynomial of degree $N-1$, in the limit $N \to \infty$ we can effectively consider $f(x)$ as a general analytic function. A property of the functional $\mathscr{H}[f]$ which simplifies the problem is the following:

$$\mathscr{H}[f + \text{constant}] = \mathscr{H}[f]. \tag{4.38}$$

Let us consider the modified variational problem $\delta\tilde{\mathscr{H}}[f] = 0$, where

$$\tilde{\mathscr{H}}[f] = \frac{1}{2} \times \frac{\int (dx/\pi)\, \sigma(g,x) \int (dy/\pi)\, \sigma(g,y)\, [(f(x) - f(y))/(x-y)]^2}{\int (dx/\pi)\, \sigma(g,x) f(x)^2}. \tag{4.39}$$

We can always choose an additive constant in order to have $\int dx\, \sigma(g,x) f(x) = 0$. Actually this property is always satisfied by the (nonconstant) eigenfunctions of $\tilde{\mathscr{H}}[f]$, since $\tilde{\mathscr{H}}[f]$ is a Hermitian quadratic form and $f = 1$ is the first eigenfunction. Hence the critical values (eigenvalues) of $\mathscr{H}$ coincide with those of $\tilde{\mathscr{H}}$ (except for the eigenvalue zero of $\tilde{\mathscr{H}}$ which is not in the spectrum of $\mathscr{H}$). The variation with respect to $f(x)$ leads to the following singular integral equation

$$\int_{-a}^{a} \frac{dy}{\pi} \sigma(g,y) \frac{f(x) - f(y)}{(x-y)^2} = \epsilon f(x), \tag{4.40}$$

the limits of integration being given by

$$a = \tfrac{1}{2}(-1 + \sqrt{1 + 16\Omega(g)}\,)^{1/2}. \tag{4.41}$$

It was shown in Ref. 13 how one can reduce Eq. (4.40) to a more manageable form, namely

$$-\frac{1}{\pi} \sigma(g,x) \frac{d}{dx} \int_{-a}^{a} \frac{\tilde{f}(y)}{y-x} dy + q(g,x) \tilde{f}(x) = \epsilon \tilde{f}(x), \tag{4.42}$$

where $\tilde{f}(x) = \sigma(g,x) f(x)$, with boundary conditions $\tilde{f}(\pm a) = 0$, and

$$q(g,x) = \frac{1}{\pi} \frac{d}{dx} \int_{-a}^{a} \frac{\sigma(g,y)}{y-x} dy. \tag{4.43}$$

The "potential" $q(g,x)$ is negative and bounded from below for every $g \geqslant 0$. Actually, it can be easily proven that $q(g,x)$ is analytic in g with the same nearest singularity g_c of $\omega(g)$. This entails that also the eigenvalues of Eq. (4.42) are analytic for $|g| < -g_c$.

Equation (4.42) can be studied by standard methods[14] and we find that the spectrum is discrete with an asymptotic behavior

$$\epsilon_n = (n+1)\omega(g) + q(g) + O(1/n), \tag{4.44}$$

where

$$q(g) = \frac{1}{\pi} \omega(g) \int_{-a}^{a} \sigma(g,x)^{-1} q(g,x) dx, \tag{4.45}$$

and $\omega(g)$ is the frequency defined by Eq. (3.13). Notice that this is the semiclassical estimate for the Hamiltonian

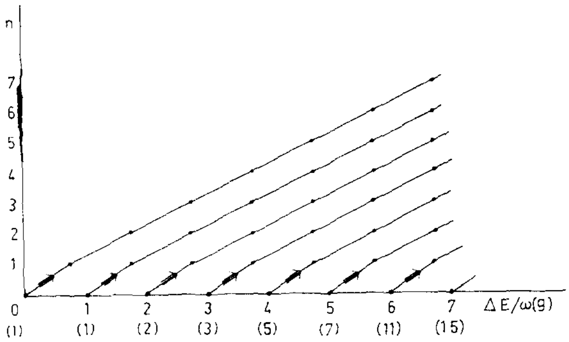

FIG. 1 The full spectrum of the adjoint representation up of $O(1/N)$: ΔE is the gap from the ground state; the arrows indicate the singlet eigenvalue over which the trajectory is built by perturbation theory in $1/N$. The numbers in parenthesis denote the multiplicity of the singlet state and of the originating trajectory.

$\mathscr{K}(x,p) = \sigma(g,x)|p| + q(g,x)$; $q(g)$ is the time average of the potential $q(g,x)$ on any orbit of $\mathscr{K}$. $\mathscr{K}(x,p)$ is the symbol of the operator of Eq. 4.42 up to subdominant terms of order $|p|^{-1}$. In the limit $g\to 0$ we have $\omega(g)\to 1$, $q(g)\to -1$ and $\epsilon_n \to n$, which is the obvious result for the N^2-dimensional harmonic oscillator. The asymptotic formula (4.44) gives a good approximation (5 to 8-figure accuracy depending on g) for $n \gtrsim 5$, while it violates the bound (4.36) for $n = 1$. To explore the low lying states, we made a numerical calculation directly on Eq. (4.40). Some results are reported in Table II. For more details see Ref. 13.

The full spectrum of the adjoint states is now obtained by making the same calculation starting from any singlet excited state ψ_s with energy $E_s = E_0(g,N) + s\omega(g) + O(1/N)$. Since the leading behavior of $\mathscr{H}[f]$ is not changed by the substitution $\psi_0 \to \psi_s$, the only difference is in the unperturbed eigenvalue E_s, and we obtain

$$E_A(n,s)(g,N) = E_0(g,N) + s\omega(g) + \epsilon_n(g) + O(N^{-1}). \tag{4.46}$$

In particular for n large an accidental degeneracy appears, due to the asymptotic behavior (4.44).

We conclude this section by recalling that a singular integral equation similar to Eq. (4.40) was found by 't Hooft[3] in the study of the mass spectrum for two-dimensional SU(N) gauge theory in the planar limit. In that case however, the potential $q(g,x)$ was singular and consequently a logarithmic correction to the asymptotic eigenvalue formula appeared.

5. SUMMARY

We have studied the spectrum of a SU(N) symmetric quantum mechanical system in the limit $N\to\infty$. For the states invariant under SU(N) we find, in this limit, an equally spaced spectrum

$$E_s - E_0 = s\omega(g) + O(N^{-1}), \quad s = 1,2,\cdots. \tag{5.1}$$

For $s \geqslant 2$ there is an accidental degeneracy given by the partition of s into integers

$$s = \sum_{i=1}^{r} (k_i + q_i) \tag{5.2}$$

such that $0 \leqslant q_1 < q_2 < \cdots < q_r$; $1 \leqslant k_1 < k_2 < \cdots < k_r$. This degeneracy is splitted by the $1/N$ correction given in Eq. (3.19). The fact that the first correction of (5.1) is of order $1/N$ does not spoil the correct $1/N^2$ expansion for the Green's function since the splitting appears with opposite sign.

We have computed also the spectrum of the states which transform as the adjoint representation of SU(N) which involves the angular variables of SU(N) in a nontrivial way. For $N\to\infty$ the spectrum is equally spaced only for high excitation and we find ($N\to\infty$)

$$E_{A(n,s)} = E_s + (n+1)\omega(g) + q(g) + O(n^{-1}), \quad n = 1,2,\cdots$$
$$(E_s = E_0 + s\omega(g), \quad s = 0,1,2,\cdots), \tag{5.3}$$

which shows that each $N^2 - 1$ multiplet of the adjoint state have the same accidental degeneracy of the corresponding singlets with energy E_s. This degeneracy is removed by $1/N$ corrections. The spectrum of $E_A(n,s)$ is given in Table II and presented in Fig. 1 which show that the asymptotic regime in n is quite precocious: In the drawing the trajectories are indistinguishable from the straight line even for $n = 1$. Equation (5.3) and Fig. 1 show that the eigenvalues $E_A(n,s)$ with equal $n + s$ are almost degenerate. Since $\omega(g) + q(g) < 0$ for any g positive we find

$$E_{A(1,0)} < E_{s=1}, \quad g > 0, \tag{5.4}$$

which shows that the first adjoint state excitation is lower than the first singlet excitation. For $g = 0$ these two states become degenerate since $\omega(0) = -q(0) = 1$.

Before concluding let us observe that the first two computed terms of the $1/N$ expansion for the singlet gaps are given by the WKB function $\omega(g)$ (Eq. 3.13) whose perturbative expansion in g has a radius of convergence given by $|g| < -g_c = 2^{1/2}/3\pi$. It is very tempting to conjecture that all the coefficients of the $1/N$ expansion have the same radius of convergence in g; in fact that $1/N$ expansion was cast in the form of a WKB expansion involving $\omega(g)$ and its derivatives. Similarly we found that the leading $1/N$ term of the gaps for the adjoint states have the same radius of convergence: The "potential" in Eq. (4.44) becomes singular at $g = g_c$. This would entail that the $1/N$ expansion is an asymptotic expansion[11] in order to reproduce the singularity of the eigenvalues at $g = 0$ for N fixed.

APPENDIX

In this appendix we compute the expectation value of the Hamiltonian on the unperturbed state

$$\Psi_A^{(0)}(\varphi) = \psi_s(\lambda) \sum_{n=2}^{N} \alpha_n \operatorname{Tr}(C\varphi^{n-1}), \tag{A1}$$

with the norm

$$\|\Psi_A^{(0)}\|^2 = \sum_{n,n'} \alpha_n \bar{\alpha}_{n'} \int V(\lambda)^2 \, d\lambda \, [dU] \, \psi_s(\lambda)^2 \times \operatorname{Tr} C\varphi^{n-1} \operatorname{Tr} C^\dagger \varphi^{n'-1}, \tag{A2}$$

where $V(\lambda)$ is the Vandermonde determinant, $d\lambda$ is the integration over the eigenvalues $\lambda_1,\cdots,\lambda_N$ of φ and $[dU]$ is the Haar measure over SU(N). The gradient of $\Psi_A^{(0)}$ is given by

$$\frac{\partial \Psi_A^{(0)}}{\partial \varphi_{ij}} = \sum_{n=2}^{N} \alpha_n \left(\sum_{k,p=1}^{N} V_{kp}^{-1}(\lambda)(\varphi^{p-1})_{ji} \operatorname{Tr}C\varphi^{n-1} \frac{\partial \psi_s}{\partial \lambda_k} + \sum_{r=0}^{n-2} (\varphi^r C \varphi^{n-r-2})_{ji} \psi_s \right), \tag{A3}$$

where $V_{kp}(\lambda)$ is the Vandermonde matrix in Eq. (4.12). This allows the calculate the kinetic energy density (notice that ψ_s is taken to be real)

$$\frac{1}{2}\sum_{i,j} \frac{\partial \Psi_A^{(0)}}{\partial \varphi_{ij}} \frac{\partial \bar{\Psi}_A^{(0)}}{\partial \varphi_{ji}}$$
$$= \frac{1}{2}\sum_{n,n'} \alpha_n \bar{\alpha}_{n'} \Bigg\{ \operatorname{Tr}C\varphi^{n-1} \operatorname{Tr}C^\dagger \varphi^{n'-1} \sum_k \left(\frac{\partial \psi_s}{\partial \lambda_k}\right)^2 + \frac{1}{2}\sum_{k,p} V_{kp}^{-1}[(n'-1)\operatorname{Tr}C\varphi^{n-1}\operatorname{Tr}C^\dagger\varphi^{n'+p-3} + (n-1)\operatorname{Tr}C^\dagger\varphi^{n'-1}\operatorname{Tr}C\varphi^{n+p-3}] \frac{\partial \psi_s^2}{\partial \lambda_k} + \sum_{r=0}^{n-2}\sum_{r'=0}^{n'-2} \operatorname{Tr}(C\varphi^{n-r+r'-2} C^\dagger \varphi^{n'-r'+r-2}) \psi_s^2 \Bigg\}. \tag{A4}$$

The kinetic energy is obtained by integrating the density over $V^2(\lambda) d\lambda\,[dU]$. The integration over SU(N) can be readily performed by introducing the polar coordinates $\varphi = U\Lambda U^\dagger$ and using the identities

$$\int [dU]\,(U^\dagger C U)_{ii}\,(U^\dagger C^\dagger U)_{jj} = (N^2-1)^{-1}\operatorname{Tr}C^\dagger C\left(\delta_{ij} - \frac{1}{N}\right), \tag{A5}$$

$$\int [dU]\,(U^\dagger C U)_{ij}\,(U^\dagger C^\dagger U)_{ji} = (N^2-1)^{-1}\operatorname{Tr}C^\dagger C\left(1 - \frac{1}{N}\delta_{ij}\right), \tag{A6}$$

which follow from the orthogonality relation of SU(N). The expectation value of the kinetic energy is then given by

$$\langle T\rangle = \frac{1}{2}\|\Psi_A^{(0)}\|^{-2} \sum_{n,n'} \alpha_n \bar{\alpha}_{n'} \int V(\lambda)^2 d\lambda \left\{ \sum_k \frac{\partial \psi_s}{\partial \lambda_k} \times \frac{\partial}{\partial \lambda_k}(W_{nn'}\psi_s) + \psi_s^2 \sum_{r=0}^{n-2}\sum_{r'=0}^{n'-2} \left[\operatorname{Tr}\varphi^{n-r+r'-2} \times \operatorname{Tr}\varphi^{n'-r'+r-2} - \frac{1}{N}\operatorname{Tr}\varphi^{n+n'-4} \right] \right\}, \tag{A7}$$

where

$$W_{nn'} \equiv \operatorname{Tr}\varphi^{n+n'-2} - \frac{1}{N}\operatorname{Tr}\varphi^{n-1}\operatorname{Tr}\varphi^{n'-1}. \tag{A8}$$

The norm of $\Psi_A^{(0)}$ is also given in terms of $W_{nn'}$:

$$\|\Psi_A^{(0)}\|^2 = \sum_{n,n'} \alpha_n \bar{\alpha}_{n'} \int V(\lambda)^2 d\lambda\, \psi_s^2\, W_{nn'}. \tag{A9}$$

The final expression in Eq. (4.22) is obtained by integrating by parts the first integral in Eq. (A7) and using the fact that ψ_s is an eigenstate of the Hamiltonian with eigenvalue E_s.

[1]G.'t Hooft, Nucl. Phys. B **72**, 461 (1974).
[2]G. Veneziano, Phys. Lett. B **52**, 220 (1974); Nucl. Phys. B **74**, 365 (1974).
[3]G. t'Hooft, Nucl. Phys. B **75**, 461 (1974).
[4]Yu. L. Dokshitser, D.I.D'Yakonov, and S.I. Troyan, Proc. 13th Leningrad Winter School, Konstantinov Institute of Nuclear Physics Leningrad (1978); C.H. Llewellyn Smith, Acta Physica Austriaca, Suppl. XIX, 331 (1978); D. Amati, R. Petronzio, and G. Veneziano; Nucl. Phys. B **147** 29 (1978).
[5]M. Ciafaloni, G. Marchesini, and G. Veneziano, Nucl. Phys. B **98**, 472, 493 (1975).
[6]G. Veneziano, Nucl. Phys. B **117**, 519 (1976).
[7]W.T. Tutte, Can. J. Math. **14**, 21 (1962).
[8]J. Koplik, A. Neveu, and S. Nussinov, Nucl. Phys. B **123**, 109 (1977).
[9]E. Brezin, C. Itzykson, G. Parisi, and J.B. Zuber, Commun. Math. Phys. **59**, 35 (1978).
[10]V.P. Maslov, *Théorie des Perturbations et Méthodes Asymptotiques* (Dunod, Paris 1972).
[11]B. De Wit and G.'t Hooft, Phys. Lett. B **69**, 61 (1977).
[12]L. Landau and E. Lifschitz, *Mécanique Quantique* (MIR, Moscow, 1966).
[13]M. Casartelli, G. Marchesini, and E. Onofri, Univ. of Parma preprint (February 1979).
[14]N.I. Muskhelishvili, *Singular Integral Equations* (Nordhoff, Groningen, Holland, 1953).
[15]F.T. Hioe and E.W. Montroll, J. Math. Phys. **16**, 1945 (1975).

On the local structure of the Euclidean Dirac field[a)]

Diego de Falco and Francesco Guerra
Istituto di Fisica dell'Università di Salerno, Italy

(Received 28 November 1978)

A simple Green's formula for the Euclidean Dirac operator in Schwinger's real formalism allows us to study some localization properties of the Euclidean Dirac field. A rather complete analysis follows of the connection between the Grassmann structure of the Euclidean field and the Clifford structure of the field at sharp time, analyzed in terms of its independent degrees of freedom.

1. INTRODUCTION

The use of methods of Euclidean quantum field theory has already provided deep technical results and insights in the theory of interacting Dirac fields (Refs. 1–6 are a first guide to the existing literature without, of course, any pretension of completeness). In this note we wish to add a few comments on the conceptual structure of the theory of Euclidean Dirac fields.

Our aim is to contribute to the understanding of the mathematical structure of the methods of functional integration in QFT, for a long time a precious heuristic instrument[7,8] (for a recent review see Ref. 9) whose deep mathematical soundness and relevance to a rigorous approach to QFT we have begun to understand since the work of Segal and Nelson.[10,11]

In this note we concentrate on the local (Markov) structure of the Euclidean Dirac field and on the problem of imbedding the sharp time Minkowski theory into the Euclidean scheme.

2. NOTATION AND BASIC DEFINITIONS

We find it expedient to use, throughout this paper, the eight-component real formalism suggested by Schwinger for the Dirac field in four space–time dimensions.[12]

In our opinion it is in this real setting that the algebraic structure of the theory becomes most transparent.

Referring to Refs. 12 and 13 for the proof of the equivalence of Schwinger's formalism to the ordinary one, here we merely recall the basic notations.

The Dirac algebra is described here in terms of the 8×8 Hermitian matrices $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ (real) and $\alpha_5, \alpha_6, \alpha_7$ (imaginary) satisfying the anticommutation relations

$$\{\alpha_i,\alpha_j\} = 2\delta_{ij}, \quad i,j = 1,\ldots,7.$$

In this formalism the Dirac equation reads, in terms of the eight-component real field Ψ,

$$i\frac{\partial}{\partial x^0}\Psi = (\boldsymbol{\alpha}\cdot\mathbf{p} + M\alpha_5)\Psi \quad \left(\mathbf{p} = \frac{1}{i}\nabla, \quad \boldsymbol{\alpha} = (\alpha_1,\alpha_2,\alpha_3)\right),$$

while the canonical anticommutation relations are

$$\{\Psi_a(x^0,\mathbf{x}), \Psi_b(x^0,\mathbf{x})\} = \delta_{a,b}\delta(\mathbf{x}-\mathbf{y}), \quad a,b, = 1,\ldots,8.$$

The matrix α_4 appears, as we shall see in the following, in the description of time reversal, while α_6 and α_7 are needed in the construction of the charge matrix.

In these notations the two-point Wightman function for the free field is

$$W(x-y) = \frac{1}{(2\pi)^3}\int d^3\mathbf{p}\,\frac{E(\mathbf{p}) + \boldsymbol{\alpha}\cdot\mathbf{p} + M\alpha_5}{2E(\mathbf{p})} \times \exp i[\mathbf{p}\cdot(\mathbf{x}-\mathbf{y}) - E(\mathbf{p})(x^0-y^0)]$$

$$[E(\mathbf{p}) = \sqrt{\mathbf{p}^2 + M^2}].$$

Analytical continuation to the Schwinger points

$$x^0 = -ix_4, \quad y^0 = -iy_4, \quad x_4 \geqslant y_4$$

leads to the Wightman function at imaginary times

$$\mathscr{W}(x-y) = \frac{1}{(2\pi)^4}\int \frac{-ip_4 + \boldsymbol{\alpha}\cdot\mathbf{p} + M\alpha_5}{p^2+M^2} \times \exp ip\cdot(x-y)\, d^4p$$

$$(p\cdot x = p_1x_1 + \cdots + p_4x_4).$$

It is only after a matrix transformation on the spinor indices generated by the matrix

$$T = \exp i(\pi/4)(1+\alpha_4),$$

according to the scheme

$$S_{ab}(x-y) = \sum_{a',b'} T_{aa'}T_{bb'}\,\mathscr{W}_{a'b'}(x-y),$$

that we get a manifestly Euclidean covariant two-point function

$$S(x-y) = \frac{i}{(2\pi)^4}\int d^4p\,\frac{\alpha\cdot p + M\alpha_5}{p^2+M^2}\exp ip\cdot(x-y)$$

$$(\alpha\cdot p = \alpha_1p_1 + \cdots + \alpha_4p_4),$$

which we will call the two-point Schwinger function and will take as the basic object of our discussion.

S is the Green's functions of the Euclidean Dirac operator

$$D = -(\alpha\cdot\partial + iM\alpha_5).$$

Notice that S is real and is antisymmetric in the sense

[a)]Research supported by Consiglio Nazionale delle Ricerche.

 0022-2488/80/051111-04$01.00

that

$$S_{ab}(x-y) = -S_{ba}(y-x).$$

2. THE LOCAL STRUCTURE

In the theory of the free Euclidean scalar field[10,14] the fact that the two-point Schwinger function

$$S_B(x-y) = \frac{1}{(2\pi)^4}\int \frac{\exp ip(x-y)}{p^2+m^2}\, d^4p$$

is the kernel of the scalar product in the Sobolev Hilbert space $\mathscr{H}_{-1}$ of the real tempered distributions with Fourier transform in $L^2[d^4p/(p^2+m^2)]$ and the Hafnian form[15] of the many point functions immediately point to the relevant functional integration theory, that of the unit Gaussian process over $\mathscr{H}_{-1}$.

The fact that $\mathscr{H}_{-1}$ contains elements strictly localized on each $x_4 =$ constant plane is the root of the possibility, through the imbedding theory, of reconstructing the sharp time Minkowski fields from the Euclidean ones.

As we wish to stay as close as possible to this line of thought also in the Euclidean Dirac case, we must confront ourselves with the difficulty that the best we can do along this line is to consider S as the kernel of an antisymmetric bilinear form on the Sobolev Hilbert space $\mathscr{H}_{-1/2}$ of the eight component real tempered distributions with Fourier transforms in $L^2[d^4p/(p^2+M^2)^{1/2}]$.

This perspective leads to the mathematically as yet unexplored (with the notable exceptions of Refs. 6 and 16), but heuristically widely used theory of the unit Pfaffian process over the simplectic space $\{\mathscr{H}_{-1/2}, S\}$ as the relevant functional integration scheme.

Here we wish to concentrate on the following problem: how to reconstruct the sharp time Minkowski fields from the Euclidean scheme in spite of the fact that $\mathscr{H}_{-1/2}$ does not contain elements localized on $x_4 =$ constant planes.

Our main tool is a Green's formula for the Euclidean Dirac operator D.

Let σ be a smooth suface dividing $\mathbb{R}^4$ into two open disjoint regions Λ_+, Λ_-. At each point $z\in\sigma$ choose the positive normal $n(z)$ as the one pointing, say, towards Λ_+. Then, for every choice of $F\in C_0^\infty(\Lambda_+)$, $G\in C_0^\infty(\Lambda_-)$ the value at F, G of the bilinear form generated by the two-point Schwinger function can be written as

$$S(F,G) = \int_\sigma d\sigma(z)\,(D^{-1}F)(z)\alpha\cdot n(z)(D^{-1}G)(z).$$

Namely, S correlates $F\in C_0^\infty(\Lambda_+)$, and $G\in C_0^\infty(\Lambda_-)$ only through the values on σ of the solutions of the inhomogeneous problems for D having F and G, respectively, as sources.

In particular, if σ is any $x_4 =$ constant plane, say $x_4 = 0$, and $F(G)$ is localized in $x_4 > 0$ (<0), then

$$S(F,G) = \langle f_0, \alpha_4 g_0\rangle_{L^2(d^3\mathbf{x})},$$

where we have set

$$f_0(\mathbf{x}) = (D^{-1}F)(\mathbf{x},0), \quad g_0(\mathbf{x}) = (D^{-1}G)(\mathbf{x},0).$$

Thus motivated, we introduce the real Hilbert spaces

$$\mathscr{H}_\pm = \{ f\in L^2_{\mathbb{R}}(d^3\mathbf{x}) : f(\mathbf{x}) = (D^{-1}F)(\mathbf{x},0)\}^-$$

with

$$\operatorname{supp} F \subseteq \{x\in\mathbb{R}^4 : x_4 \lessgtr 0\}\}$$

[where the closure is in the $L^2(d^3\mathbf{x})$ norm].

After checking that $\mathscr{H}_\pm$ are mutually orthogonal in $L^2(d^3\mathbf{x})$ we set

$$\mathscr{H} = \mathscr{H}_+ \oplus \mathscr{H}_-.$$

We will agree to say that functions in $\mathscr{H}_+$ ($\mathscr{H}_-$) are localized on the upper (lower) face of the $x_4 = 0$ plane.

This convention, obviously motivated by the definitions on $\mathscr{H}_\pm$, is also coherent with the fact that the time reversal operator

$$\tau : f\in\mathscr{H} \to \tau f = \alpha_4 f$$

exchanges $\mathscr{H}_+$ with $\mathscr{H}_-$:

$$f\in\mathscr{H}_\pm \Rightarrow \tau f\in\mathscr{H}_\mp.$$

τ is a symmetric orthogonal operator on the real Hilbert space $\mathscr{H}$:

$$\tau = \tau^T, \quad \tau^T\tau = \tau\tau^T = 1.$$

The projection $P_\pm$ onto $\mathscr{H}_\pm$ are easily seen to be

$$P_\pm = (1 \pm k\tau)/2,$$

where k is the antisymmetric orthogonal operator on $\mathscr{H}$, anticommuting with τ

$$k^T = -k, \quad k^2 = -1, \quad k\tau + \tau k = 0,$$

best described in terms of its action on Fourier transforms as multiplication by

$$k(\mathbf{p}) = i\,\frac{\boldsymbol{\alpha}\cdot\mathbf{p} + M\alpha_5}{(\mathbf{p}^2 + M^2)^{1/2}}.$$

Observe that $kP_\pm = \mp\tau P_\pm$.

It is convenient to define an antisymmetric bilinear form s on $\mathscr{H}$ by the position

$$s(f,g) = \langle f, kg\rangle_{L^2(d^3\mathbf{x})}, \quad f,g\in\mathscr{H}.$$

In terms of the objects just introduced we can summarized the discussion up to this point into the following pre-Markov property for the Euclidean Dirac field: given any $x_4 = \sigma$ plane, we have exhibited a canonical prescription to construct

(i) a simplectic space $\{\mathscr{H}, s\}$ of functions on the $x_4 = \sigma$ plane;

(ii) a map from functions F localized above the plane $x_4 = \sigma$ into functions f_σ localized on its upper face;

(iii) a map from functions G localized below the plane into functions g_σ localized on its lower face; such that, for any such F and G:

$$S(F,G) = s(f_\sigma, g_\sigma).$$

More generally, if $\sigma_1 < \sigma_2$ and F is localized above the plane $x_4 = \sigma_2$ and G is localized below the plane $x_4 = \sigma_1$

$$S(F,G) = s(f_{\sigma_2}, \exp[-(\sigma_2 - \sigma_1)\mu]\, g_{\sigma_1}).$$

where μ is the operator described by its action on Fourier transforms as multiplication by

$$\mu(\mathbf{p}) = (\mathbf{p}^2 + M^2)^{1/2}.$$

3. Q SPACE FOR THE DIRAC FIELD

As for second quantization, the triple $\{\mathscr{H},k,\tau\}$ lends itself very nicely to the Fock realization of the sharp time field and to the analysis of its degree of freedom.

After introducing the orthogonal projectors

$$\Lambda_{\pm} = (1 \mp ik)/2$$

in $L^2(d^3\mathbf{x})$ (which are not any more reality preserving), Fock space is constructed as

$$\mathscr{F} = \bigoplus_{n=0}^{\infty} F \textcircled{A}^n,$$

where F is the range of Λ_+.

The vacuum state is $\Omega_0 = (1,0,0,\cdots)$ while the field operators are constructed through the real linear map

$$f \in \mathscr{H} \to \Psi(f) = C(\Lambda_+ f) + A(\Lambda_+ f),$$

where C and A are creation and annihilation operators.

Observe that the two point function can be written as

$$W_0(f,g) \equiv \langle \Omega_0, \Psi(f)\Psi(g)\Omega_0\rangle_{\mathscr{F}}$$

$$= \tfrac{1}{2}[\langle f,g\rangle_{\mathscr{H}} - i\langle f,kg\rangle_{\mathscr{H}}].$$

Together with this conventional realization of the free field structure, we wish to suggest here an equivalent alternative which, being closely related to the simplectic structure of $\mathscr{H}$, as opposed to its orthogonal structure, is better suited to a comparison with the Euclidean theory.

We start by observing that the fields $\Psi(f)$ split into two mutually anticommuting sets, those with $f \in \mathscr{H}_+$ and those with $f \in \mathscr{H}_-$, and that, due to the simple observation that if f is, say, in $\mathscr{H}_+$, then

$$\Psi(f)\Omega_0 = -i\Psi(kf)\Omega_0,$$

the vacuum, is cyclic under the fields of each of the two sets.

Let us focus out attention on the dense set $\mathscr{F}_-$ in $\mathscr{F}$ generated by the Wick polynomials of the fields $\Psi(f)$ with $f \in \mathscr{H}_-$ applied to Ω_0.

Here we define Wick ordering by

$$:c: = c,$$

$$:\Psi(f_1): = \Psi(f_1),$$

$$:\Psi(f_1)\Psi(f_2): = \Psi(f_1)\Psi(f_2) - W_0(f_1,f_2),$$

$$\cdots\cdots\cdots\cdots$$

$$:\Psi(f_1)\cdots\Psi(f_n): = \Psi(f_1):\Psi(f_2)\cdots\Psi(f_n): - \sum_{j=2}^{n}(-1)^j W_0(f_1,f_j)$$

$$\vdots$$

Call $\mathscr{G}(\mathscr{H})$ ($\mathscr{G}(\mathscr{H}_-)$) the Grassmann algebra over $\mathscr{H}$ ($\mathscr{H}_-$) and denote its generators by $\psi(f)$, $f \in \mathscr{H}$ ($\mathscr{H}_-$).

Define the linear map (henceforth called duality map)

$$\mathscr{D}: :\Psi(f_1)\cdots\Psi(f_n):\Omega_0 \in \mathscr{F}_-$$
$$\to 2^{-n/2}\psi(f_1)\cdots\psi(f_n) \in \mathscr{G}(\mathscr{H}_-).$$

The point is that $\mathscr{G}(\mathscr{H}_-)$ can be made into a Hilbert space $L^2(\mathscr{H}_-,\omega,K)$ in such a way that $\mathscr{D}$ can be lifted to a unitary transformation.

Here $\{\mathscr{H},\omega\}$ is the unit Pfaffian process over the simplectic space $\{\mathscr{H},s\}$, namely ω is the linear functional on the Grassmann algebra $\mathscr{G}(\mathscr{H})$ defined by

$$\omega(\mathbb{1}) = 1,$$

$$\omega(\psi(f_1)\psi(f_2)) := s(f_1,f_2),$$

$$\omega(\psi(f_1)\cdots\psi(f_{2k})) = \sum_{j=2}^{2k}(-1)^j s(f_1,f_j)$$

$$\omega(\psi(f_1)\cdots\psi(f_{2k+1})) = 0.$$

K is the antilinear map on $\mathscr{G}(\mathscr{H})$ defined on the monomials as

$$K\lambda\psi(f_1)\cdots\psi(f_n) = \lambda^*\psi(kf_n)\cdots\psi(kf_1).$$

$L^2(\mathscr{H}_-,\omega,K)$ is the completion of $\mathscr{G}(\mathscr{H}_-)$ in the scalar product

$$\langle\ ,\ \rangle_{\mathscr{G}(\mathscr{H}_-)}: (\mathscr{A},\mathscr{B}) \in \mathscr{G}(\mathscr{H}_-) \times \mathscr{G}(\mathscr{H}_-)$$

$$\to \omega((K\mathscr{A})\mathscr{B}).$$

The image of the field Ψ under the duality transform is easily checked to be

$$\mathscr{D}\Psi(f)\mathscr{D}^{-1} = \begin{cases} \dfrac{1}{\sqrt{2}}(\psi(f) + \partial_f), & f \in \mathscr{H}_-, \\[2ex] \dfrac{i}{\sqrt{2}}(\partial_{kf} - \psi(kf)), & f \in \mathscr{H}_+. \end{cases}$$

Here the linear operator of functional derivation in the direction f is defined through its action on the monomials as:

$$\partial_f\psi(f_1)\cdots\psi(f_n) = \sum_{j=1}^{n}\langle f_1,f_j\rangle \frac{\partial}{\partial\psi(f_j)}\psi(f_1)\cdots\psi(f_n)$$

$$= \sum_{j=1}^{n}(-1)^{j+1}\langle f,f_j\rangle_{\mathscr{H}}$$

$$\times \psi(f_1)\cdots\widehat{\psi(f_j)}\cdots\psi(f_n).$$

The foregoing decomposition of the dual image of the time zero Dirac operator into a multiplication part (nicely imbedded into the Euclidean scheme) and a differentiation part is the key to understanding how the Grassmann structure of the Euclidean field goes into the Clifford structure of the time zero Minkowski field, an apparent difficulty sharply focused in the lattice approximation of Ref. 17.

[1]K. Osterwalder and R. Schrader, "Euclidean Fermi Fields and a Feynman–Kac Formula for Boson–Fermion Models," Helv. Phys. Acta **46**, 277 (1973).

[2]E. Seiler, "Schwinger Functions for the Yukawa Model in Two Dimensons with Space–Time Cutoff," Commun. Math. Phys. **42**, 163 (1975).
[3]E. Seiler and B. Simon, "Nelson's Symmetry and All That in the Yukawa$_2$ and $(\varphi^4)_3$ Field Theories," Ann. Phys. **97**, 470 (1976).
[4]A. Cooper and L. Rosen, "The Weakly coupled Yukawa$_2$ Field Theory: Cluster expansion and Wightman Axioms," Trans. Am. Math. Soc. (to appear).
[5]O. McBrian, "Volume Dependence of Schwinger Functions in the Yukawa$_2$ Quantum Field Theory," Commun. Math. Phys. **45**, 279 (1975).
[6]L. Gross, "Euclidean Fermion Fields," Report to the Marseille Colloqium on Mathematical Methods of Quantum Field Theory, 23–27 June 1975.
[7]P.T. Matthews, and A. Salam, "The Green's Functions of Quantized Fields," Nuovo Cimento **12**, 563 (1954).
[8]P.T. Matthews and A. Salam, "Propagators of Quantized Fields," Nuovo Cimento **2**, 120 (1955).
[9]V.N. Popov, "Functional Integrals in Quantum Field Theory," CERN Report TH. 2424 (1977).
[10]E. Nelson, "Construction of Quantum Fields from Markoff Fields," J. Func. Anal. **12**, 97 (1973).
[11]I. Segal, "Tensor Algebras over Hilbert Spaces, I," Trans. Am. Math. Soc. **81**, 106 (1956).
[12]J. Schwinger, "Euclidean Quantum Electrodynamics," Phys. Rev. **115**, 721 (1959).
[13]D. de Falco, "Metodi di Integrazione Funzionale in Teoria dei Campi," Salerno, Ph. D. thesis, 1976 (unpublished).
[14]F. Guerra, L. Rosen, and B. Simon, "The $P(\varphi)_2$ Euclidean Quantum Field Theory as Classical Statistical Mechanics," Ann. Math. **101**, 111 (1975).
[15]E. Caianiello, *Combinatorics and Renormalization* (Benjamin, New York, 1973).
[16]F.A. Berezin, *The Method of Second Quantization* (Academic, New York, 1966).
[17]K. Osterwalder and E. Seiler, "Gauge Field Theories on the Lattice," Ann. Phys. **110**, 440 (1978).

Quantum Inverse Method for two-dimensional ice and ferroelectric lattice models

H.B. Thacker
Fermi National Accelerator Laboratory, Batavia, Illinois 60510

Received 17 December 1979; accepted for publication 18 January 1980)

The quantum inverse scattering transform method previously developed for continuum field theories is applied to the exactly soluble symmetric six-vertex (ice or ferroelectric) lattice model. Operators analogous to those which appear in the quantum inverse treatment of the nonlinear Schrödinger and sine–Gordon equations are constructed on the lattice by forming strings of vertices contracted over horizontal arrows. From the commutation relations for these operators, exact formulas for the eigenstates and eigenvalues of the transfer matrix are obtained without making an explicit ansatz for the wave functions. These results illustrate the connection between the quantum inverse method and the transfer matrix formalism for lattice models.

The inverse scattering method was developed as a means of solving certain classical nonlinear field equations.[1] The possibility that this technique might be generalized to provide a method for solving quantum field theory was suggested by studies of the nonlinear Schrödinger equation.[2,3] In its classical form, this equation had been solved via the 2×2 matrix inverse problem of Zakharov and Shabat.[4] The quantum nonlinear Schrödinger equation (also known as the delta-function gas) had also been solved by the Bethe ansatz of Lieb and Liniger.[5] The connection between these two methods was established by constructing quantum operators analogous to the classical Jost functions and scattering data of the Zakharov–Shabat eigenvalue problem.[3,6–8] An operator $B(k)$ thus constructed was found to create the Bethe ansatz eigenstates of the delta-function gas. Recently, the quantum inverse method has been applied to the sine-Gordon equation[9] and shown to reproduce the results of the Bethe ansatz solution of the massive Thirring model.[10] The elegant formulation of this method by Faddeev, Skylanin, and Takhtajan[9] exhibits a striking connection with the transfer matrix formalism developed in the treatment of solvable lattice statistical models.[11] In this paper we explore this connection by applying the quantum inverse method to the ice and ferroelectric lattice models of Lieb and Baxter[12,13] which were originally solved by writing a Bethe ansatz for the eigenvectors of the transfer matrix.[14]

We find a very compact derivation of the known results by constructing operators on the lattice which are analogous to the A and B operators used in the quantum nonlinear Schrödinger[6–8] and sine–Gordon[9] equations. This formulation illustrates a profound connection between the 2×2 matrix structure of the inverse scattering eigenvalue problem used in continuum field theories, and the matrix structure represented by the horizontal arrows of the lattice theory. The vertical arrows are associated with the operators of the field theory. The transfer matrix T is related to the A operator, while the B operator creates eigenstates of T. The path-ordered exponential expression which describes solutions of the eigenvalue problem in the inverse method arises on the lattice as a string of vertices contracted over horizontal indices. It is remarkable that the inverse method, which originated in classical field theory, is so closely related (in its quantum field version) to the transfer matrix formalism for lattice models.

The general ice or ferroelectric model (symmetric six-vertex model) is constructed by placing arrows on the bonds of a square lattice in all possible way which obey the "ice rule," i.e., that there are two arrows in and two arrows out at each vertex. It is a special case of the Baxter eight-vertex model[16] with Baxter's parameter $d=0$. This eliminates the two vertices with four arrows in or four arrows out. The symmetric model is then described by three vertex weights, a, b, and c in Baxter's notation. The elementary vertex can be written as

$$L(\alpha,\beta;\lambda,\mu)=\sum_{i=1}^{4} w_i\sigma^i_{\alpha\beta}\sigma^i_{\lambda\mu}, \tag{1}$$

where σ^i, $i=1,2,3$, are Pauli matrices, $\sigma^4=1$, and the indices α,β and λ,μ refer to horizontal and vertical arrows, respectively. The parameters w_i are related to the vertex weights by

$$w_1=w_2=\tfrac{1}{2}c, \tag{2a}$$

$$w_3=\tfrac{1}{2}(a-b), \tag{2b}$$

$$w_4=\tfrac{1}{2}(a+b), \tag{2c}$$

For our considerations, it is convenient to regard the vertex (1) as an explicit 2×2 matrix in the horizontal indices, each element of which is a spin operator in the space of vertical indices. Thus, we write

$$L_n=\begin{pmatrix} w_3\sigma_n^3+w_4\sigma_n^4 & 2w_1\sigma_n^- \\ 2w_1\sigma_n^+ & -w_3\sigma_n^3+w_4\sigma_n^4\end{pmatrix}, \tag{3}$$

where $\sigma^{\pm}=\frac{1}{2}(\sigma^1\pm i\sigma^2)$ and the subscript n indicates that the σ-matrices act on the vertical arrow at site n.

In the usual quantum inverse method for continuum field theories,[6–9] one considers solutions to a linear problem of the form

$$\left[\frac{\partial}{\partial x}+iQ\right]\psi=0, \tag{4}$$

 0022-2488/80/051115-03$01.00

where ψ is a two-component column vector, and $Q(x)$ is a 2×2 matrix, each element of which is a function of the field (e.g., nonlinear Schrödinger or sine–Gordon) at the point x. A solution of Eq. (4) can be written as a path-ordered exponential,

$$\psi(y) = P\exp\left\{-i\int_x^y Q(x')dx'\right\}\psi(x). \tag{5}$$

The observation which leads to the present application of the inverse method is that the path-ordered exponential in (5) has a precise analog in the lattice theory. It is a string of elementary vertices formed by contracting on the horizontal arrows, i.e., by multiplying matrices of the form (3) along adjacent sites in a row.

For a lattice with N sites in a row, the quantities which correspond to the scattering data in the continuum inverse method are obtained by multiplying over the whole row, leaving the end arrows uncontracted,

$$\mathscr{T} = L_1L_2\cdots L_N. \tag{6}$$

Henceforth, we will adopt Baxter's parametrization of the vertex weights[14] (specialized to the six-vertex case),

$$w_1 = w_2 = \rho\sin 2\eta, \tag{7a}$$

$$w_3 = \rho\sin\eta\cos v, \tag{7b}$$

$$w_4 = \rho\cos\eta\sin v. \tag{7c}$$

For the discussion to follow, η is regarded as a real constant and v as a variable. (They are related to coupling constant and rapidity, respectively, in field theory.[10]) Without loss of generality, we can take the overall normalization $\rho = 1$.

The elements of $\mathscr{T}$ given by (6) are the "scattering data" operators of the theory,

$$\mathscr{T}(v) = \begin{pmatrix} A(v) & B(v) \\ C(v) & D(v)\end{pmatrix}. \tag{8}$$

The transfer matrix is just the trace of (8),

$$T(v) = \mathrm{Tr}\,\mathscr{T}(v) = A(v) + D(v). \tag{9}$$

Let us define the direct product of two matrices as follows:

$$M\otimes N = \begin{pmatrix} M_{11}N_{11} & M_{11}N_{12} & M_{12}N_{11} & M_{12}N_{12} \\ M_{11}N_{21} & M_{11}N_{22} & M_{12}N_{21} & M_{12}N_{22} \\ M_{21}N_{11} & M_{21}N_{12} & M_{22}N_{11} & M_{22}N_{12} \\ M_{21}N_{21} & M_{21}N_{22} & M_{22}N_{21} & M_{22}N_{22}\end{pmatrix}. \tag{10}$$

Here, each element is a product of operators, and must be written in the specified order. As in other applications of the quantum inverse method,[6,9] we find that the direct products of two elementary vertices $L_n(v)$ and $L_n(v')$, taken in different order, are related by a similarity transformation,

$$L(v')\otimes L(v) = RL(v)\otimes L(v')R^{-1}. \tag{11}$$

Here R is a c-number matrix of the form

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \beta & \alpha & 0 \\ 0 & \alpha & \beta & 0 \\ 0 & 0 & 0 & 1\end{pmatrix}, \tag{12}$$

where

$$\alpha(v,v') = \frac{\sin(v-v')}{\sin(v-v'-2\eta)}, \tag{13a}$$

$$\beta(v,v') = \frac{-\sin(2\eta)}{\sin(v-v'-2\eta)}. \tag{13b}$$

Equation (11) may be verified by direct calculation. Formation of the direct products in Eq. (11) may be visualized as the contraction of two vertices along a vertical arrow (represented by an operator product in field theory). The matrix R in Eq. (12) is the same as one constructed for the sine–Gordon theory by Faddeev, *et al.*[9] It is also the $d = 0$ limit of a matrix constructed by Baxter, who used an equation of the form (11) in his derivation of commuting transfer matrices for the eight-vertex model.[14]

The fundamental relation (11) provides all the commutation relations needed to construct the eigenvectors of the transfer matrix and to calculate its eigenvalues. The scattering data matrix $\mathscr{T}(v)$, by its definition, Eq. (6), satisfies a similar equation,

$$\mathscr{T}(v')\otimes\mathscr{T}(v) = R\,[\,\mathscr{T}(v)\otimes\mathscr{T}(v')]R^{-1}, \tag{14}$$

which specifies the commutation relations among the operators A, B, C, and D. Just as in the sine–Gordon case, Eq. (14) leads to the following results:

$$[A(v), A(v')] = [B(v), B(v')] = 0, \tag{15a}$$

$$A(v)B(v') = \frac{1}{\alpha(v',v)}B(v')A(v) - \frac{\beta(v',v)}{\alpha(v',v)}B(v)A(v'), \tag{15b}$$

$$D(v)B(v') = \frac{1}{\alpha(v,v')}B(v')D(v) + \frac{\beta(v,v')}{\alpha(v,v')}B(v)D(v'), \tag{15c}$$

$$[A(v)+D(v), A(v')+D(v')] = 0. \tag{15d}$$

As in the usual Bethe ansatz formulation,[12,13] the eigenstates of the transfer matrix $T(v) = A(v) + D(v)$ are constructed upon one of the two direct product eigenstates, e.g., the state with all spins up,

$$|\Omega_0\rangle = |\uparrow\rangle_1\otimes|\uparrow\rangle_2\otimes\cdots\otimes|\uparrow\rangle_N. \tag{16}$$

Notice that L_n, Eq. (3), when acting on an up spin at site n, becomes a triangular matrix,

$$L_n|\uparrow\rangle_n = \begin{pmatrix}\sin(v+\eta) & \sin 2\eta\ \sigma_- \\ 0 & \sin(v-\eta)\end{pmatrix}|\uparrow\rangle_n. \tag{17}$$

From (16), (17), and (6), we conclude that $|\Omega_0\rangle$ is an eigenstate of $A(v)$ and $D(v)$ separately,

$$A(v)|\Omega_0\rangle = [\sin(v+\eta)]^N|\Omega_0\rangle, \tag{18a}$$

$$D(v)|\Omega_0\rangle = [\sin(v-\eta)]^N|\Omega_0\rangle. \tag{18b}$$

Eigenstates of $T(v)$ with n reversed arrows are constructed by applying operators $B(v_i)$, $i = 1,\ldots,n$ [where $B(v)$ is defined by (3), (6), and (8)] to the state $|\Omega_0\rangle$,

$$|v_1,\ldots,v_n\rangle = \prod_{i=1}^n B(v_i)|\Omega_0\rangle. \tag{19}$$

Conditions on the v_i's emerge in the course of verifying that (19) is an eigenstate of $T(v)$. Using the relations (18) and (15), the following result can be shown:

$$T(v)|v_1,\ldots,v_n\rangle = \Lambda(v;v_1,\ldots,v_n)|v_1,\ldots,v_n\rangle, \tag{20}$$

where

$$A(v;v_1,\dots,v_n) = [\sin(v+\eta)]^N \prod_{i=1}^{n} \left[\frac{\sin(v-v_i-2\eta)}{\sin(v-v_i)}\right] - [\sin(v-\eta)]^N \prod_{i=1}^{n} \left[\frac{\sin(v-v_i+2\eta)}{\sin(v-v_i)}\right]. \tag{21}$$

To show (20), we write $T(v) = A(v) + D(v)$ and commute A and D past all of the B operators in (19) using (15b) and (15c). When such a procedure is carried out, for example, on $A(v)$, it produces 2^n terms. One of these terms comes entirely from the first term in (15b) and, along with the corresponding term from $D(v)$, yields directly the right-hand side of (20) with the eigenvalue (21). The remaining terms involve states in which one of the v_i's is replaced by v, and these terms must be made to cancel if Eq. (20) is to be satisfied. The first such term, where v_1 is replaced by v, is easily found to be

$$\frac{\beta(v,v_1)}{\alpha(v,v_1)}\left\{[\sin(v_1+\eta)]^N \prod_{l=2}^{n} \frac{1}{\alpha(v_l,v_1)} - [\sin(v_1-\eta)]^N \prod_{l=2}^{n} \frac{1}{\alpha(v_1,v_l)}\right\}|v,v_2,\dots,v_n\rangle. \tag{22}$$

Other terms involving the states in which v_j is replaced by v, with $j > 1$, may also be calculated directly, but such a calculation is unnecessary. From the symmetry of the state (19), which follows from the second commutator in (15a), we see that each of the remaining terms may be obtained from (22) simply by interchanging v_1 and v_j. The requirement that all such terms vanish leads to the conditions

$$[\sin(v_j-\eta)]^N \prod_{\substack{l=1\\ l\neq j}}^{n} [\sin(v_j-v_l+2\eta)] = [\sin(v_j+\eta)]^N \prod_{\substack{l=1\\ l\neq j}}^{n} [\sin(v_j-v_l-2\eta)]. \tag{23}$$

Equations (21) and (23) are the familiar transfer matrix eigenvalues and periodic boundary conditions for the ice models.[12,13] Thus, we have constructed the eigenstates and eigenvalues of the transfer matrix by a method which is considerably more transparent than the original Bethe ansatz treatment and which clearly demonstrates the connection between soluble lattice models and the quantum inverse formalism.

From the examples of the nonlinear Schrödinger equation, the sine–Gordon/massive Thirring model, and the ice models discussed here, it is apparent that the quantum generalization of the classical inverse scattering technique provides an elegant formulation of exact results for soluble quantum field theories and lattice statistical models. Further refinement and extension of this method may provide additional insight into the nature of conservation laws and exact integrability in quantum field theory.

ACKNOWLEDGMENTS

I would like to thank W.A. Bardeen, H. Bergknoff, D.B. Creamer, and D. Wilkinson for many stimulating conversations.

[1]C.S. Gardner, J.M. Greene, M.D. Kruskal, and R.M. Miura, Phys. Rev. Lett. **19**, 1095 (1967); for a review, see A.C. Scott, F.Y.F. Chu, and D.W. McLaughlin, Proc. IEEE **61**, 1443 (1973).

[2]H.B. Thacker, Phys. Rev. D **17**, 1031 (1978).

[3]E.K. Sklyanin and L.D. Faddeev, Akad. Nauk SSR Dokl. **243**, 1430 (1978).

[4]V.E. Zakharov and A.B. Shabat, Zh. Eksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys. JETP **34**, 62 (1972)].

[5]E.H. Lieb and W. Liniger, Phys. Rev. **130**, 1605 (1963).

[6]E.K. Sklyanin, Akad. Nauk SSR Dokl. **244**, 1337 (1979).

[7]H.B. Thacker and D. Wilkinson, Phys. Rev. D **19**, 3660 (1979).

[8]J. Honerkamp, P. Weber, and A. Wiesler, Nucl. Phys. B **152**, 266 (1979).

[9]L.D. Faddeev, E.K. Sklyanin, and L.A. Takhtajan, Leningrad preprint LOMI P-1-79.

[10]H. Bergknoff and H.B. Thacker, Phys. Rev. Lett. **42**, 135 (1979); Phys. Rev. D **19**, 3666 (1979).

[11]For a review, see P.W. Kasteleyn, *Fundamental Problems in Satistical Mechanics*, Vol. 3, edited by E.D.G. Cohen (North-Holland, Amsterdam, 1975).

[12]E.H. Lieb, Phys. Rev. Lett. **18**, 1046 (1967); **19**, 108 (1967); Phys. Rev. **126**, 162 (1967).

[13]R.J. Baxter, Stud. Appl. Math. **50**, 51 (1971).

[14]The solution of the six-vertex model described here is closely related to a recent treatment of the *XXZ* Heisenberg spin chain by P.P. Kulish and E.K. Sklyanin, Phys. Lett. **70A**, 461 (1979). I would like to thank L.D. Faddeev for bringing this reference to my attention. The correspondence between the *XXZ* spin chain and the six-vertex model is a special case of the general correspondence between the *XYZ* chain and the eight-vertex model (see Ref. 15). These authors consider only the case of an infinitely long spin chain and, hence, only treat the case $\Delta > 1$ (η = pure imaginary). The operator albegra constructed here [Eq. (15)] is valid for a finite size lattice, giving both the transfer matrix eigenvalues and the periodic boundary conditions. Thus, the present treatment is valid for all values of the η parameter, including the case η real ($|\Delta| < 1$) which is relevant to continuum field theory.

[15]R.J. Baxter, Ann. Phys. **70**, 323 (1972).

[16]R.J. Baxter, Phys. Rev. Lett. **26**, 832 (1971); Ann. Phys. N.Y. **70**, 193 (1972).

A variational method in two-group transport theory

Tomasz Bleński

Computing Center CYFRONET, Institute of Nuclear Research, Świerk, 05—400 Otwock, Poland

(Received 18 July 1979; accepted for publication 23 August 1979)

The two-group integral transport equation in subcritical homogeneous isotropically scattering medium in arbitrary geometry is considered. A functional is constructed which is stationary for the solution to the integral equation. If the determinant of the transfer matrix is positive, then the solution gives the minimum of the functional. The applications to the cylindrical Milne problem is shown and examples of the numerical results of the extrapolation distances are presented.

1. INTRODUCTION

The extension of the Case method of elementary solutions for the multigroup and especially two-group transport theory has been the subject of a number of papers.[1-10] Most of them were devoted to the equations with the plane symmetry where the H-matrix technique has been developed and successfully applied in so-called half-range problems.[6-8] In this technique solutions are superpositions of the elementary solutions with coefficients obtained in terms of the scalar products from the orthogonality relations on the basis of the completeness theorem. For practical calculations of these scalar products the numerical evaluation of the H matrix from a nonlinear integral equation is needed.

In other geometries the transport equations, for which the transformations into planelike equations exist, can be treated in principle with the same method. There are two methods of the transformation: the replication property[10,11] or the transform function technique.[12,13] Nevertheless, these methods seem to be attractive from the numerical point of view in the limited number of "inner" problems, while in the "outer" serious numerical difficulties appear.

For these reasons in the present work the extension of the classical variational method[14] for two-group transport equation is proposed. The basic equation is the integral transport equation in the homogeneous isotropically scattering medium in the arbitrary geometry.

For this equation it is possible to construct a functional which is stationary for the solution. However, the proof that this is also the extremum requires the assumption of the positivity of the determinant of the transfer matrix $\hat{C}$. If this if the case, then the convergence of the Ritz (or other known) method is guranteed. As the example of application the cylindrical Milne problem[15,16] is considered in detail. It is shown that the extrapolation distances λ_i depend in fact on certain integrals from the bounded part of the solutions. These integrals can be obtained as the combination of the minima of the variational functionals; it makes the method especially attractive when applied to calculate λ_i.

The method works with the sufficient accuracy for the range of R (the radius of the black cylinder) from about 0.1 to 15. (in mean free path). The most interesting is the intermediate region of $R\sim 1$ while for $R\to 0$ or $R\to\infty$ there exist the asymptotic expansions.[16,17] Two numerical examples presented in the paper are for $R=1$, and another one with $R=10$ is for comparison with the results from the asymptotic expansion.

The plane or spherical geometry yields much more simple matrix kernels of the integral equations, and one may expect that the corresponding calculations in these symmetries can be performed with great numerical accuracy.

2. THE INTEGRAL TRANSPORT EQUATION

We write the two-group, stationary transport equation in homogeneous isotropically scattering medium in region V in the form:

$$(\mathbf{\Omega}\cdot\nabla+\hat{\Sigma})\bar{\Psi}(\mathbf{r},\mathbf{\Omega})=\hat{C}\frac{1}{4\pi}\int d\,\mathbf{\Omega}'\bar{\Psi}(\mathbf{r},\mathbf{\Omega}')+\bar{q}(\mathbf{r},\mathbf{\Omega}),\quad(1)$$

where the angular flux

$$\bar{\Psi}(\mathbf{r},\mathbf{\Omega})=\begin{pmatrix}\Psi^{(1)}(\mathbf{r},\mathbf{\Omega})\\ \Psi^{(2)}(\mathbf{r},\mathbf{\Omega})\end{pmatrix};$$

with $\Psi^{(i)}(\mathbf{r},\mathbf{\Omega})$, the angular flux in the ith energy group, depends on position $\mathbf{r}\in V$ and velocity directional vector $\mathbf{\Omega}$;

$$\hat{\Sigma}=\begin{pmatrix}\sigma & 0\\ 0 & 1\end{pmatrix}$$

is the total cross section matrix where σ and 1 are the cross sections in the first and second group, respectively; $\hat{C}$ with nonnegative elements c_{ij}, $i,j=1,2$, is the transfer matrix. It is assumed that $c_{12}\cdot c_{21}\neq 0$ since if one of the off-diagonal terms vanishes the equations for each group are in fact uncoupled. Hence one can obtain the equation with the symmetric transfer matrix by the transformation[6]

$$\bar{\Psi}'=\hat{A}\bar{\Psi},\quad \hat{C}'=\hat{A}\hat{C}\hat{A}^{-1},\qquad(2)$$

with

$$\hat{A}=\begin{pmatrix}(c_{21}/c_{12})^{1/2} & 0\\ 0 & 1\end{pmatrix}.$$

The vector $\bar{q}(\mathbf{r},\mathbf{\Omega})$ denotes sources.

Integration of (1) along the characteristics[14] leads to the integral equation for the flux

$$\bar{n}(\mathbf{r})=\int d\,\mathbf{\Omega}\;\bar{\Psi}(\mathbf{r},\mathbf{\Omega}),$$

$$\bar{n}(\mathbf{r})=(\hat{P}\hat{C}\bar{n})(\mathbf{r})+\bar{f}(\mathbf{r}),\qquad(3)$$

where $\hat{P}$ is the integral operator:

 0022-2488/80/051118-08$01.00

$$(\widehat{P}\widehat{C}\bar{n})(\mathbf{r}) = \int_V d\mathbf{r}'\, \widehat{\mathscr{P}}(\mathbf{r},\mathbf{r}')\widehat{C}\bar{n}(\mathbf{r}'), \tag{4}$$

with

$$\widehat{\mathscr{P}}(\mathbf{r},\mathbf{r}') = \frac{1}{4\pi}\frac{1}{|\mathbf{r}-\mathbf{r}'|^2}\widehat{\mathrm{E}}\mathrm{xp}(-|\mathbf{r}-\mathbf{r}'|)\chi(\mathbf{r},\mathbf{r}'),$$

$$\widehat{\mathrm{E}}\mathrm{xp}(x) = \begin{pmatrix} \exp(\sigma x) & 0 \\ 0 & \exp(x) \end{pmatrix};$$

$$\chi(\mathbf{r},\mathbf{r}') = \begin{cases} 1, & \text{if}\{\mathbf{r}-s(\mathbf{r}'-\mathbf{r}); s\in(0,1)\}\subset V, \\ 0, & \text{otherwise.} \end{cases}$$

The vector $\bar{f}(\mathbf{r})$ is obtained from the boundary condition and the source vector $\bar{q}(\mathbf{r},\mathbf{\Omega})$. Redefining $\bar{n}' = \bar{n} - \bar{f}$, one gets

$$\bar{n}'(\mathbf{r}) = [\widehat{P}(\widehat{C}\bar{n}' + \widehat{C}\bar{f})](\mathbf{r}) \tag{5}$$

and now $\widehat{C}\bar{f}(\mathbf{r})$ corresponds to isotropic sources in (1) with $\bar{\Psi}'(\mathbf{r},\mathbf{\Omega})$, $\int d\,\mathbf{\Omega}\bar{\Psi}(\mathbf{r},\mathbf{\Omega}) = \bar{n}'(\mathbf{r})$, instead of $\bar{\Psi}(\mathbf{r},\mathbf{\Omega})$.

If $V = \mathbb{R}^3$, then the straightforward integration gives for any $\mathbf{r}$

$$\int_{\mathbb{R}^3} d\mathbf{r}'\, \widehat{\mathscr{P}}(\mathbf{r},\mathbf{r}') = \widehat{\Sigma}^{-1}. \tag{6a}$$

From (6a) and the positivity of $\widehat{\mathscr{P}}(\mathbf{r},\mathbf{r}')$ it follows for $V\subset\mathbb{R}^3$

$$\int_V d\mathbf{r}'\, \widehat{\mathscr{P}}(\mathbf{r},\mathbf{r}') < \widehat{\Sigma}^{-1}, \quad V\neq\mathbb{R}^3, \tag{6b}$$

where the inequality here and in all matrix inequalities in the remaining part of the paper denotes the inequalities for the corresponding nonzero elements of the matrices on both sides.

Since we want to consider also equations in one or two dimensions as well as in curvilinear geometries suppose that we change the variables from $\mathbf{r}$ to (x,η), where by x we denote k ($k\leqslant 3$) variables upon which the flux $\bar{n}$ depends, $\bar{n} = \bar{n}(x)$, while η stands for the remaining $3-k$ variables.

Let us suppose further that

$$d\mathbf{r} = \omega(x)dx\,d\eta, \quad \omega(x) > 0.$$

Now our equation of interest is of the following form:

$$\bar{n}(x) = (\widehat{K}\widehat{C}\bar{n})(x) + \bar{f}(x), \quad x\in V, \tag{7a}$$

$$(\widehat{K}\widehat{C}\bar{n})(x) = \int_V dx'\, \omega(x')\widehat{\mathscr{K}}(x,x')\widehat{C}\bar{n}(x'), \tag{7b}$$

where

$$\widehat{\mathscr{K}}(x,x') = \int d\eta'\, \widehat{\mathscr{P}}(x;x',\eta'). \tag{8}$$

For convenience we retain V to denote the region of x and the integration in (8) is over the whole range of η' in $\mathbb{R}^3$ [$(3-k)$-dimensional subspace].

The inequality (6) immediately gives

$$\int_V dy\, \omega(y)\widehat{\mathscr{K}}(x,y) \leqslant \widehat{\Sigma}^{-1}, \tag{9}$$

where equality holds only when V covers the k-dimenisional subspace in $\mathbb{R}^3$.

We consider Eq. (7) in the Hilbert space $L^2_\omega(V)$ with the scalar product $\langle\,|\,\rangle_L$;

$$\bar{h},\bar{g}\in L^2_\omega(V), \quad \langle\bar{h}|\bar{g}\rangle_L = \int_V dx\, \omega(x)\langle\bar{h}(x)|\bar{g}(x)\rangle, \tag{10}$$

where $\langle\,|\,\rangle$ denotes a scalar product in $\mathbb{R}^2$, and with the norm $\|\;\|_L$:

$$\bar{h}\in L^2_\omega(V), \quad \|\bar{h}\|_L = (\langle\bar{h}|\bar{h}\rangle_L)^{1/2}.$$

First, let us assume that $\bar{f}\in L^2_\omega(V)$ and estimate the spectral radius of $\widehat{K}\widehat{C}$ in $L^2_\omega(V)$ [18];

$$\|\widehat{K}\widehat{C}\|_{\mathrm{sp}} = \lim_{m\to\infty}(\|(\widehat{K}\widehat{C})^m\|_L)^{1/m}.$$

From the definition

$$\|\widehat{K}\widehat{C}\bar{n}\|_L = \bigg[\int_V dx\, \omega(x)\int_V dy\, \omega(y)\int_V dz\, \omega(z) \times\langle\widehat{\mathscr{K}}(x,y)\widehat{C}\bar{n}(y)|\widehat{\mathscr{K}}(x,z)\widehat{C}\bar{n}(z)\rangle\bigg]^{1/2}.$$

$\widehat{\mathscr{K}}(x,y)$ is diagonal and positive. Let us denote

$$\bar{n}^{(1)}(x,y,z) = [\widehat{\mathscr{K}}(x,y)]^{1/2}[\widehat{\mathscr{K}}(x,z)]^{1/2}\widehat{C}\bar{n}(y),$$

$$\bar{n}^{(2)}(x,y,z) = [\widehat{\mathscr{K}}(x,y)]^{1/2}[\widehat{\mathscr{K}}(x,z)]^{1/2}\widehat{C}\bar{n}(z).$$

The Hölder inequality and (9) gives

$$\begin{aligned}
&\langle\widehat{K}\widehat{C}\bar{n}|\widehat{K}\widehat{C}\bar{n}\rangle_L \\
&= \int_V dx\, \omega(x)\int_V dy\, \omega(y)\int_V dz\, \omega(z)\langle\bar{n}^{(1)}(x,y,z)|\bar{n}^{(2)}(x,y,z)\rangle \\
&\leqslant\bigg\{\int_V dx\, \omega(x)\int_V dy\, \omega(y)\int_V dz\, \omega(z) \times\langle\bar{n}^{(1)}(x,y,z)|\bar{n}^{(1)}(x,y,z)\rangle\bigg\}^{1/2} \\
&\quad\times\bigg\{\int_V dx\, \omega(x)\int_V dy\, \omega(y)\int_V dz\, \omega(z) \times\langle\bar{n}^{(2)}(x,y,z)|\bar{n}^{(2)}(x,y,z)\rangle\bigg\}^{1/2} \\
&= \int_V dx\, \omega(x)\int_V dy\, \omega(y)\int_V dz\, \omega(z) \times\langle\bar{n}^{(1)}(x,y,z)|\bar{n}^{(1)}(x,y,z)\rangle \\
&= \int_V dy\, \omega(y)\bigg\langle\int_V dx\, \omega(x)\widehat{\mathscr{K}}(x,y) \times\int_V dz\, \omega(z)\widehat{\mathscr{K}}(x,z)\widehat{C}\bar{n}(y)|\widehat{C}\bar{n}(y)\bigg\rangle \\
&\leqslant\int_V dy\, \omega(y)\langle\widehat{\Sigma}^{-1}\widehat{C}\bar{n}(y)|\widehat{\Sigma}^{-1}\widehat{C}\bar{n}(y)\rangle \\
&\leqslant\|\widehat{\Sigma}^{-1}\widehat{C}\|\;\|\bar{n}\|_L,
\end{aligned} \tag{11}$$

where $\|\;\|$ is the norm of 2×2 matrices introduced by the scalar product $\langle\,|\,\rangle$.

Iteration for any natural m yields

$$\|(\widehat{K}\widehat{C})^m\bar{n}\|_L \leqslant\|(\widehat{\Sigma}^{-1}\widehat{C})^m\|\cdot\|\bar{n}\|_L,$$

and the conclusion is that $\|\widehat{K}\widehat{C}\|_{\mathrm{sp}}$ is less or equal to the spectral radius of the matrix $\widehat{\Sigma}^{-1}\widehat{C}$, the largest of absolute values of its eigenvalues which we denote after Ref. 7 by k_{BMS}.

$$\|\widehat{K}\widehat{C}\|_{sp}\leqslant|\lambda_{max}| = k_{\mathrm{BMS}}. \tag{12}$$

It is clear that in (12) we have equality only when V corresponds to the k-dimensional subspace in $\mathbb{R}^3$. In this paper we consider subcritical media so we assume $k_{\rm BMS} \leqslant 1$ or $k_{\rm BMS} < 1$ when in (12) is equality. For subcritical media $\|\widehat{K}\widehat{C}\|_{\rm sp} < 1$, and Eq. (7) has in $L^2_\omega(V)$ one and only one solution $\bar{n}_f$ which can be found as the series[18]

$$\bar{n}_f = \bar{f} + \widehat{K}\widehat{C}\bar{f} + (\widehat{K}\widehat{C})^2\bar{f} + \cdots.$$

3. THE VARIATIONAL THEOREM

The vector $\bar{n}_f^* = \widehat{C}\bar{n}_f$, $\bar{n}_f^* \in L^2_\omega(V)$, where $\bar{n}_f$ is the solution to (7) satisfies the equation adjoint to (7):

$$\bar{n} = (\widehat{K}\widehat{C})^*\bar{n} + \bar{f}^*, \tag{13}$$

where $(\widehat{K}\widehat{C})^* = \widehat{C}\widehat{K}$ is adjoint to $\widehat{K}\widehat{C}$ in $L^2_\omega(V)$ and $\bar{f}^* = \widehat{C}\bar{f}$. Let us introduce the functional

$$\mathscr{J}_f(\bar{n}) = \tfrac{1}{2}\langle \widehat{C}\bar{n} | (\widehat{I} - \widehat{K}\widehat{C})\bar{n}\rangle_L - \langle \widehat{C}\bar{n} | \bar{f}\rangle_L \tag{14}$$

and check whether it has the required properties.

The first term in (14) can be also written as $\langle \bar{n} | \widehat{C}(\widehat{I} - \widehat{K}\widehat{C})\bar{n}\rangle_L$ since we have $\widehat{C}$ symmetric. The operator $\widehat{C}(\widehat{I} - \widehat{K}\widehat{C})$ is self-adjoint in $L^2_\omega(V)$, and from this fact it follows that

$$\overline{\rm grad}\, \mathscr{J}_f(\bar{n}) = \widehat{C}\,[(\widehat{I} - \widehat{K}\widehat{C})\bar{n} - \bar{f}\,], \tag{15}$$

where $\overline{\rm grad}\, \mathscr{J}_f(\bar{n})$ denotes the gradient of $\mathscr{J}_f(\bar{n})$ introduced in the usual way[19]. For instance, let $F(\bar{n})$ be a functional

$$\bar{n} \in L^2_\omega(V), \quad F(\bar{n}) : L^2_\omega(V) \to \mathbb{R}^1;$$

then for $\bar{h} \in L^2_\omega(V)$:

$$\langle\, \overline{\rm grad}\, F(\bar{n}) | \bar{h}\,\rangle_L = \frac{d}{dt} F(\bar{n} + t\bar{h})\Big|_{t=0} = \lim_{t\to 0} \frac{F(\bar{n} + t\bar{h}) - F(\bar{n})}{t}.$$

Equation (15) is obtained directly from the above definition.

If $\det\widehat{C} \neq 0$, then the functional $\mathscr{J}_f(\bar{n})$ is stationary only when $\bar{n} = \bar{n}_f$

$$\overline{\rm grad}\, \mathscr{J}_f(\bar{n}_f) = 0. \tag{16}$$

Equation (16) is the necessary condition that $\bar{n}_f$ be the point of extremum of $\mathscr{J}_f(\bar{n})$.[19] Our next step is to investigate the sign of $\Delta\mathscr{J}_f(\bar{n}) = \mathscr{J}_f(\bar{n}) - \mathscr{J}_f(\bar{n}_f)$ to find whether $\mathscr{J}_f(\bar{n})$ has in $\bar{n} = \bar{n}_f$ minimum or maximum. From the definition (14)

$$\Delta\mathscr{J}_f(\bar{n}) = \tfrac{1}{2}\langle \delta\bar{n}_f | \widehat{C}(\widehat{I} - \widehat{K}\widehat{C})\delta\bar{n}_f\rangle_L, \tag{17}$$

where

$$\delta\bar{n}_f = \bar{n} - \bar{n}_f,$$

which means that our question is whether or not the scalar product $\langle \delta\bar{n}_f | \widehat{C}(\widehat{I} - \widehat{K}\widehat{C})\delta\bar{n}_f\rangle_L$ is always positive (or negative) for all $\delta\bar{n}_f \in L^2_\omega(V)$, $\delta\bar{n}_f \neq 0$.

Let us consider the first possibility. Due to the positive matrix kernel of $\widehat{K}$ the term $\langle \delta\bar{n}_f | \widehat{C}\widehat{K}\widehat{C}\delta\bar{n}_f\rangle_L$ is always positive. Thus $\det\widehat{C} > 0$ is the necessary condition for minimum; $\widehat{C}$ has nonnegative elements and with $\det\widehat{C} > 0$ is positive definite; hence the first term $\langle \delta\bar{n}_f | \widehat{C}\delta\bar{n}_f\rangle_L$ in $\Delta\mathscr{J}_f(\bar{n})$ is positive. Let us assume that $\det\widehat{C} > 0$ and try to estimate $\langle \bar{n} | \widehat{C}\widehat{K}\widehat{C}\bar{n}\rangle_L$, $\bar{n} \in L^2_\omega(V)$. We repeat the steps leading to (11):

$$\langle \bar{n} | \widehat{C}\widehat{K}\widehat{C}\bar{n}\rangle_L = \int_V dx\, \omega(x) \int_V dy\, \omega(y) \langle \bar{g}^{(1)}(x,y) | \bar{g}^{(2)}(x,y)\rangle,$$

where

$$\bar{g}^{(1)}(x,y) = [\widehat{\mathscr{K}}(x,y)]^{1/2}\widehat{C}\bar{n}(x),$$
$$\bar{g}^{(2)}(x,y) = [\widehat{\mathscr{K}}(x,y)]^{1/2}\widehat{C}\bar{n}(y),$$

$$\begin{aligned}\langle \bar{n} | \widehat{C}\widehat{K}\widehat{C}\bar{n}\rangle_L &\leqslant \int_V dx\, \omega(x) \int_V dy\, \omega(y) \langle \bar{g}^{(1)}(x,y) | \bar{g}^{(1)}(x,y)\rangle \\ &\leqslant \int_V dx\, \omega(x) \Big\langle \int_V dy\, \omega(y) \widehat{\mathscr{K}}(x,y) \widehat{C}\bar{n}(x) \Big| \widehat{C}\bar{n}(x)\Big\rangle \\ &\leqslant \int_V dx\, \omega(x) \langle \widehat{C}\widehat{\Sigma}^{-1}\widehat{C}\bar{n}(x) | \bar{n}(x)\rangle. \end{aligned} \tag{18}$$

With $\widehat{C}$ positive definite and symmetric, we can introduce in $\mathbb{R}^2$ a new scalar product $[\;|\;]$:

$$\bar{u}, \bar{w} \in \mathbb{R}^2, \quad [\bar{u} | \bar{w}] = \langle \widehat{C}\bar{u} | \bar{w}\rangle,$$

in which the matrix $\widehat{\Sigma}^{-1}\widehat{C}$ is symmetric:

$$[\bar{u} | \widehat{\Sigma}^{-1}\widehat{C}\bar{w}] = \langle \widehat{C}\bar{u} | \widehat{\Sigma}^{-1}\widehat{C}\bar{w}\rangle = [\widehat{\Sigma}^{-1}\widehat{C}\bar{u} | \bar{w}].$$

Now the Rayleigh principle[20] can be applied,

$$\begin{aligned}\langle \widehat{C}\widehat{\Sigma}^{-1}\widehat{C}\bar{n}(x) | \bar{n}(x)\rangle &= [\widehat{\Sigma}^{-1}\widehat{C}\bar{n}(x) | \bar{n}(x)] \\ &\leqslant k_{\rm BMS}\,[\bar{n}(x) | \bar{n}(x)] = k_{\rm BMS}\,\langle \widehat{C}\bar{n}(x) | \bar{n}(x)\rangle, \end{aligned} \tag{19}$$

since $k_{\rm BMS}$ denotes the largest eigenvalue of the matrix $\widehat{\Sigma}^{-1} \times \widehat{C}$. In case of subcritical media (19) gives

$$\langle \widehat{C}\bar{n} | \bar{n}\rangle_L \geqslant \langle \widehat{C}\widehat{\Sigma}^{-1}\widehat{C}\bar{n} | \bar{n}\rangle_L, \quad \bar{n} \in L^2_\omega(V), \quad \bar{n} \neq 0. \tag{20}$$

Finally (18), (19), and (20) yield

$$\langle \widehat{C}\bar{n} | \bar{n}\rangle_L > \langle \widehat{C}\widehat{K}\widehat{C}\bar{n} | \bar{n}\rangle_L, \quad \bar{n} \in L^2_\omega(V), \quad \bar{n} \neq 0, \tag{21}$$

the inequality is sharp since either $k_{\rm BMS} < 1$ or there is the sharp inequality in (18). From (21) we conclude that the requirement $\det\widehat{C} > 0$ appears to be also the sufficient condition for the minimum.

In the opposite case, i.e., $\det\widehat{C} < 0$ the sign of $\langle \widehat{C}\bar{n} | (\widehat{I} - \widehat{K}\widehat{C})\bar{n}\rangle_L$ depends on $\bar{n} \in L^2_\omega(V)$. For instance, let us consider the subspace $H_1 \subset L^2_\omega(V)$ generated by $\bar{h}_1 \in \mathbb{R}^2$, an eigenvector corresponding to the largest eigenvalue $k_{\rm BMS}$ of the matrix $\widehat{\Sigma}^{-1}\widehat{C}$. The vector $\bar{h} \in H_1$ is of the form

$$\bar{h}(x) = h(x)\bar{h}_1,$$

where $h(x)$ is a scalar function. Using (18) and the definition of $\bar{h}$, one obtains

$$\begin{aligned}\langle \widehat{C}\bar{h} | (\widehat{I} - \widehat{K}\widehat{C})\bar{h}\rangle_L &= \langle \widehat{C}\bar{h} | \bar{h}\rangle_L - \langle \widehat{C}\bar{h} | \widehat{K}\widehat{C}\bar{h}\rangle_L \\ &\geqslant \langle \widehat{C}\bar{h} | \bar{h}\rangle_L - \int_V dx\, \omega(x) h^2(x) \langle \widehat{C}\bar{h}_1 | \widehat{\Sigma}^{-1}\widehat{C}\bar{h}_1\rangle \\ &= \langle \widehat{C}\bar{h} | \bar{h}\rangle_L - k_{\rm BMS}\langle \widehat{C}\bar{h} | \bar{h}\rangle_L. \end{aligned} \tag{22}$$

The elements of $\widehat{\Sigma}^{-1}\widehat{C}$ are nonnegative. From the Perron–Frobenius theorem[20] it follows that there exists an $\bar{h}_1 \in \mathbb{R}^2$

which is nonnegative. As $\widehat{C}$ is nonnegative, $\widehat{C}\bar{h}_1$ is also nonnegative, and

$$\langle \widehat{C}\bar{h} | \bar{h} \rangle_L = \int_V dx\, \omega(x) \langle \widehat{C}\bar{h}_1 | \bar{h}_1 \rangle h^2(x) \geqslant 0.$$

Finally from (22)

$$\langle \widehat{C}\bar{h} | (\widehat{I} - \widehat{K}\widehat{C})\bar{h} \rangle_L \geqslant 0, \quad \bar{h} \in H_1.$$

On the other hand, choosing another subspace H_2, $\bar{h}' \in H_2$, $\bar{h}'(x) = h'(x)\bar{h}_2$, where $h'(x)$ is a scalar function and $\bar{h}_2 \in \mathbb{R}^2$ is the eigenvector of $\widehat{C}$ with the negative eigenvalue, one obtains immediately

$$\langle \widehat{C}\bar{h}' | (\widehat{I} - \widehat{K}\widehat{C})\bar{h}' \rangle_L \leqslant 0, \quad \bar{h}' \in H_2.$$

These examples show that in case of $\det\widehat{C} < 0$ none of the operators $\widehat{C}(\widehat{I} - \widehat{K}\widehat{C})$ or $-\widehat{C}(\widehat{I} - \widehat{K}\widehat{C})$ is positive definite.

The following theorem results from the above considerations:

Theorem: If $\det\widehat{C} \neq 0$ and $\widehat{C}$ is symmetric, then Eq. (7) is equivalent to the equation

$$\overline{\text{grad}}\, \mathscr{J}_f(\bar{n}) = 0, \tag{23}$$

where the functional $\mathscr{J}_f(\bar{n})$ is defined in (14). If, moreover, $\det\widehat{C} > 0$ then the functional $\mathscr{J}_f(\bar{n})$ has minimum in $\bar{n} = \bar{n}_f$, where $\bar{n}_f$ denotes the solution to (7) or (23). In case of $\det\widehat{C} < 0$, $\mathscr{J}_f(\bar{n})$ has no extremum in $\bar{n} = \bar{n}_f$.

We turn again to the case $\det\widehat{C} > 0$. Slightly changing the steps leading to (21), one can obtain the relation

$$\langle \widehat{C}\bar{n} | (\widehat{I} - \widehat{K}\widehat{C})\bar{n} \rangle_L \geqslant m^2 \langle \widehat{C}\bar{n} | \bar{n} \rangle_L, \quad \bar{n} \in L^2_\omega(V), \tag{24a}$$

where in the subcritical media

$$m^2 = 1 - k_{\text{BMS}} + \delta > 0;$$

δ denotes the smallest eigenvalue of the matrix $(\widehat{\Sigma}^{-1} - \widehat{S})\widehat{C}$ and $\widehat{S}$ is the diagonal matrix with elements

$$S_{ii} = \sup_{x \in V} \int_V dy\, \omega(y) \mathscr{K}_{ii}(x,y), \quad i = 1,2.$$

The inequality (24a) means that the operator $\widehat{I} - \widehat{K}\widehat{C}$ is self-adjoint and positive definite[21,22] in $L^2_\omega(V)$, but in the sense of the new scalar product $[\ |\ ]_L$:

$$\bar{h}, \bar{g} \in L^2_\omega(V), \quad [\bar{h} | \bar{g}]_L = \langle \widehat{C}\bar{h} | \bar{g} \rangle_L.$$

The norm corresponding to $[\ |\ ]_L$ we denote by $[|\ \ |]_L$:

$$\bar{h} \in L^2_\omega(V), \quad [|\bar{h}|]_L = ([\bar{h} | \bar{h}]_L)^{1/2}.$$

We now apply the known procedure for the self-adjoint and positive definite operators.[21] First, these properties and the fact that $\widehat{I} - \widehat{K}\widehat{C}$ is bounded in $L^2_\omega(V)$ allows us to introduce in $L^2_\omega(V)$ the third scalar product $\{\ |\ \}_L$:

$$\bar{h}, \bar{g} \in L^2_\omega(V), \quad \{\bar{h} | \bar{g}\}_L = [\bar{h} | (\widehat{I} - \widehat{K}\widehat{C})\bar{g}]_L,$$

with the norm $\{|\ \ |\}_L$:

$$\bar{h} \in L^2_\omega(V), \quad \{|\bar{h}|\}_L = (\{\bar{h} | \bar{h}\}_L)^{1/2}.$$

The space $L^2_\omega(V)$ with the scalar product $\{\ |\ \}_L$ will be denoted as $\widetilde{L}^2_\omega(V)$.

Using the new definitions, we can express (24a) in the shorter form:

$$\bar{h} \in L^2_\omega(V), \quad \frac{1}{m}\{|\bar{h}|\}_L \geqslant [|\bar{h}|]_L, \quad m > 0. \tag{24b}$$

The fixed element $\bar{f} \in L^2_\omega(V)$ defines a bounded functional in $\widetilde{L}^2_\omega(V)$:

$$\bar{h} \in L^2_\omega(V), \quad [\bar{f} | \bar{h}]_L \leqslant [|\bar{f}|]_L [|\bar{h}|]_L \leqslant \frac{[|\bar{f}|]_L}{m} \{|\bar{h}|\}_L.$$

From the Riesz theorem[23] there exists $\bar{n}_f \in L^2_\omega(V)$ such that for all $\bar{h}$

$$[\bar{f} | \bar{h}]_L = \{\bar{n}_f | \bar{h}\}_L; \tag{25}$$

and the functional $\mathscr{J}_f(\bar{n})$ takes the form

$$\begin{aligned} \mathscr{J}_f(\bar{n}) &= \tfrac{1}{2}\{\bar{n}|\bar{n}\}_L - [\bar{n}|\bar{f}]_L \\ &= \tfrac{1}{2}\{\bar{n}|\bar{n}\}_L - \{\bar{n}|\bar{n}_f\}_L = \tfrac{1}{2}\{|\bar{n} - \bar{n}_f|\}^2_L - \tfrac{1}{2}\{|\bar{n}_f|\}^2_L, \end{aligned} \tag{26}$$

with $\bar{n}_f$ being the solution to (7) or (23). In our case since the domain of $\widehat{I} - \widehat{K}\widehat{C}$ is the whole space $L^2_\omega(V)$ the vector $\bar{n}_f \in L^2_\omega(V)$. The new form of $\mathscr{J}_f(\bar{n})$ in the case $\det\widehat{C} > 0$ given in (26) is useful when the approximate methods are to be applied. The space $L^2_\omega(V)$ is separable. Any minimization series

$$\bar{n}^{(k)}, \quad k = 1,2,\cdots, \quad \text{for } \mathscr{J}_f(\bar{n}):$$

$$\lim_{k \to \infty} \mathscr{J}_f(\bar{n}^{(k)}) = \inf_{\bar{n} \in L^2_\omega(V)} \mathscr{J}_f(\bar{n}) = -\tfrac{1}{2}\{|\bar{n}_f|\}^2_L;$$

converges to $\bar{n}_f$ in the norm $\{|\ \ |\}_L$ and from (24b) also in the norm $[|\ \ |]_L$.[21,22] Let $\{\bar{\varphi}_k\}$ be a basis in $L^2_\omega(V)$. In the Ritz method one looks for the minimum in the form

$$\sum_{i=1}^{k} c_i^{(k)} \bar{\varphi}_i, \quad k = 1,2,3,\cdots,$$

and obtains immediately from (26) the system of linear equations

$$\sum_{i=1}^{k} c_i^{(k)} \{\bar{\varphi}_i | \bar{\varphi}_j\}_L = [\bar{f} | \bar{\varphi}_j]_L, \quad j = 1,2,\ldots,k.$$

The fact that $\widehat{I} - \widehat{K}\widehat{C}$ is positive definite gurantees that the determinant of the matrix $(\{\bar{\varphi}_i | \bar{\varphi}_j\}_L)$, $i,j = 1,\ldots,k$, is different from zero.

4. REMARK ON THE n-GROUP APPROXIMATION

All results obtained in the previous section remain valid in the n-group approximation, $n > 2$, provided that the assumptions about the matrix $\widehat{C}$ are fulfilled. In the two-group case $\widehat{C}$ can be symmetrized by the diagonal matrix $\widehat{A}$ if only $c_{12} \cdot c_{21} \neq 0$. The fact that $\widehat{A}$ is diagonal is essential since the symmetrization should preserve $\widehat{\Sigma}$ diagonal. Generally, in the case of n groups, $n > 2$, the symmetrization can be done only if the detailed balance exists[24] and, moreover, is not violated by the discretization in energy. In the continuous dependence upon energy the detailed balance condition gives for the scattering kernel $K(v,v')$ the relation

$$m_0(v',T)K(v,v') = m_0(v,T)K(v',v), \tag{27}$$

where $m_0(v,T)$ denotes the Maxwellian with the temperature T. This equality leads to the symmetrized kernel $\widetilde{K}(v,v')$,

$$\widetilde{K}(v,v') = [m_0(v,T)/m_0(v,T)]^{1/2} K(v,v'). \tag{28}$$

We impose that the condition (27) remain valid after the discretization in energy. A definition analogous to (28) gives us $\widehat{C}$, symmetric:

$$\widehat{C}' = \widehat{A}\widehat{C}\widehat{A}^{-1},$$

where $(\widehat{A})_{ij} = [m_0(v_i,T)]^{1/2}\delta_{ij}$ is the diagonal transformation matrix. For $n = 2$, $\widehat{A}$ is of the same form; however, the symmetrization can be performed automatically without the assumption of the detailed balance. The condition that $\widehat{C}'$ is symmetric gives $n(n-1)/2$ equations for $n-1$ elements of the diagonal matrix $\widehat{A}$ (with the accuracy to the multiplication factor). If $n = 2$, then these numbers are equal while if $n > 2$ additional constraints for $\widehat{C}$ must appear.

Symmetry of $\widehat{C}$ allows to construct the functional $\mathscr{F}_f(\bar{n})$ the gradient of which is equal to $\widehat{C}(\widehat{I} - \widehat{K}\widehat{C})\bar{n} - \widehat{C}\bar{n}$ and $\mathscr{F}_f(\bar{n})$ is stationary in $\bar{n} = \bar{n}_f$. To ensure the positive definitness of $\widehat{C}(\widehat{I} - \widehat{K}\widehat{C})$ in $L^2_\omega(V)$ one needs the positive definitness of the matrix $\widehat{C}$ in $\mathbb{R}^n$. Note that in the two-group case the latter can be viewed as the requirement that the coupling parameter $(c_{12}\cdot c_{21})/(c_{11}\cdot c_{22})$ is smaller than 1, which is true for most actual situations.

5. EXAMPLE—MILNE PROBLEM IN CYLINDRICAL GEOMETRY

As an example let us take the Milne problem in cylindrical geometry.[14–16] It corresponds to the situation when the infinite long cylindrical black (absolutely absorbing) body of radius R is immersed in the two-group infinite isotropically scattering medium.

For the neutron flux $\bar{n}$ we have the following equation[14–16]

$$(\widehat{I} - \widehat{L}_R\widehat{C})\bar{n}(r) = 0, \quad r \geqslant R, \tag{29}$$

where $\bar{n} = \bar{n}(r)$, r denotes the distance from the axis of the cylinder,

$$(\widehat{L}_R\widehat{C}\bar{n})(r) = \int_R^\infty dr'\, r' \widehat{\mathscr{L}}_R(r,r')\widehat{C}\bar{n}(r'),$$

$$\widehat{\mathscr{L}}_R(r,r') = \frac{4}{\pi}\int_{|r-r'|}^{(r^2-R^2)^{1/2}+(r'^2-R^2)^{1/2}} d\rho \Big/ \{[\rho^2 - (r-r')^2][(r-r')^2 - \rho^2]\}^{-1/2} \int_\rho^\infty dt\, \widehat{K}_0(t)\widehat{\Sigma},$$

$$\widehat{K}_0(t) = \begin{pmatrix} K_0(\sigma t) & 0 \\ 0 & K_0(t) \end{pmatrix};$$

$K_0(t)$ is the modified Bessel function of the second kind and σ is the total cross section of the first group. Here distances are measured in the free path of the second group.

In the limit $R \to 0$ (without the black body) we denote

$$\lim_{R\to 0} \widehat{L}_R = \widehat{L}, \tag{30a}$$

and one can check (the technique is the same as for one group approximation[12,13]) that the equation

$$(\widehat{I} - \widehat{L}\widehat{C})\bar{n}_0 = 0 \tag{30b}$$

has solutions of the form

$$\bar{n}_0(r) = \sum_{i=1}^{\kappa} \alpha_i I_0\left(\frac{r}{\nu_i}\right)\bar{U}(\nu_i), \tag{30c}$$

where $I_0(x)$ is the modified Bessel function and the ν_i are zeros of the characteristic equation[6]:

$$\det\widehat{A}(\nu_i) = 0,$$

where

$$\widehat{A}(\nu_i) = \widehat{I} + \nu_i \int_{-1}^{1} \frac{d\mu}{\mu - \nu_i}\widehat{\Theta}(\mu)\widehat{C},$$

$$\widehat{\Theta}(\mu) = \begin{pmatrix} \theta(\mu) & 0 \\ 0 & 1 \end{pmatrix},$$

with $\theta(\mu)$ being the characteristic function of the interval $(-1/\sigma, 1/\sigma)$, and the characteristic equation is obtained as the solvability condition of the equation for the vectors $\bar{U}(\nu_i)$ which now have the solutions

$$\bar{U}(\nu_i) = \begin{pmatrix} -A_{12}(\nu_i) \\ A_{11}(\nu_i) \end{pmatrix}.$$

In the subcritical case with $k_{\rm BMS} < 1$ there are one or two pairs of real zeros $\{\pm\nu_i\}$. The number of these pairs is denoted by κ. If $\kappa = 2$, the two components of $\bar{U}(\nu_2)$ (we choose $\nu_1 > \nu_2$) are of different signs,[4,9] and the physical solution to (30) is

$$\bar{n}_0(r) = I_0(r/\nu_1)\bar{U}(\nu_1). \tag{31}$$

Now we return to the Milne problem with $R \neq 0$. We look for the unbounded solution to (29) which behaves as (31) for large r (it corresponds to the sources at infinity).

The Placzek lemma[15] allows us to treat the Milne problem as the whole space problem with negative anisotropic cylindrical shell sources situated in $r = R$. If we are interested in the influence of these sources upon the solution at the asymptotically large distances from the cylinder, then these sources in turn can be substituted by the negative spatially distributed sources density of which is of the order $O(e^{-(r-R)})$. Thus the asymptotic part of the correction $\bar{n}^{\rm as}(r) - I_0(r/\nu_1)\bar{U}(r_1)$ due to the presence of the absorber is the same as the asymptotic part $\bar{n}_s^{\rm as}(r)$ of the bounded solution to the equation

$$\bar{n}_s(r) = (\widehat{L}\bar{n}_s)(r) + \bar{s}(r), \tag{32}$$

where $\bar{s}(r)$ represents the latter sources;

$$\bar{n}^{\rm as}(r) = \bar{n}_s^{\rm as}(r) + I_0(r/\nu_1)\bar{U}(\nu_1). \tag{33}$$

$\bar{n}^{\rm as}(r)$ and $\bar{n}_s^{\rm as}(r)$ can be described as follows:

$$[\bar{n}^{\rm as}(r) - \bar{n}(r)]e^r < \infty,$$

$$[\bar{n}_s^{\rm as}(r) - \bar{n}_s(r)]e^r < \infty.$$

Similarly as in the one-group case (Ref. 13, Appendix C) one can obtain from the plane Green matrix[9] the cylindrical shell Green matrix $\widehat{G}(r,r_0)$

$$\widehat{G}(r,r_0) = \sum_{i=1}^{\kappa} \frac{1}{\nu_i N(\nu_i)}\bar{U}(\nu_i)\otimes\bar{U}(\nu_i) \times K_0\left(\frac{r_>}{\nu_i}\right)I_0\left(\frac{r_<}{\nu_i}\right) + O(e^{-|r-r_0|}), \tag{34}$$

where following Ref. 9 the notation

$$\bar{a}\otimes\bar{b}=\begin{pmatrix}a_1b_1 & a_1b_2\\ a_2b_1 & a_2b_2\end{pmatrix}$$

is used, $N(\nu_i)$ is the normalization integral,[6] $r_>=\max(r,r_0)$, $r_<=\min(r,r_0)$. The matrix $\widehat{G}(r,r_0)$ is defined as the solution to the equation.

$$\widehat{G}(r,r_0)=\int_0^\infty dr'\,r'\widehat{\mathscr{L}}(r,r')[\widehat{C}\widehat{G}(r',r_0)+(1/r_0)\delta(r'-r_0)\widehat{I}],$$

where $\widehat{\mathscr{L}}(r,r')=\widehat{\mathscr{L}}_R(r,r')|_{R=0}$. The definition of $\widehat{G}(r,r_0)$ gives for $\bar{n}_s(r)$

$$\bar{n}_s(r)=\int_0^\infty dr_0\,r_0\widehat{G}(r,r_0)\widehat{C}\bar{s}(r_0)+\bar{s}(r).$$

Putting (34) in the above expression and utilizing the $O(e^{-r})$-like behavior of $\bar{s}(r)$ one finally obtains

$$\bar{n}_s^{\rm as}(r)=\sum_{i=1}^{\kappa}\frac{1}{\nu_iN(\nu_i)}\int_0^\infty dr_0\,r_0\times\langle\widehat{C}I_0(r_0/\nu_i)\bar{U}(\nu_i)|\bar{s}(r_0)\rangle K_0(r/\nu_i)\bar{U}(\nu_i). \tag{35}$$

We have established the form of the solution to the Milne problem

$$\bar{n}(r)=I_0\left(\frac{r}{\nu_1}\right)\bar{U}(\nu_1)+\sum_{i=1}^{\kappa}\beta_iK_0\left(\frac{r}{\nu_i}\right)\bar{U}(\nu_i)+O(e^{-r}). \tag{36}$$

Usually the most interesting is the asymptotic part of the flux owing to the importance of the extrapolation distances λ_i,[4,13]

$$\lambda_i\equiv n^{(i)_{\rm as}}(R)\Big/\frac{dn^{(i)_{\rm as}}}{dr^{(i)}}(r)\Big|_{r=R},\quad i=1,2.$$

The remaining part of this section is devoted to the calculation of λ_i. First observe that if we put

$$\bar{n}(r)=\bar{n}_1(r)+\bar{p}_1(r),$$

where

$$\bar{n}_1(r)=I_0(r/\nu_1)\bar{U}(\nu_1),$$

then $\bar{p}_1$, the bounded part of the solution, obeys the inhomogeneous equation

$$\bar{p}_1(r)=(\widehat{L}_R\widehat{C}\bar{p}_1)(r)+\bar{f}_1(r), \tag{37a}$$

with the source term

$$\bar{f}_1(r)=[(\widehat{I}-\widehat{L}_R\widehat{C})\bar{n}_1](r)=O(e^{-r}). \tag{37b}$$

Now, following Marshak (Ref. 15, p. 211), we extend the definition of $\bar{p}_1(r)$ and $\bar{f}_1(r)$ to the region $r<R$

$$\bar{p}_1(r)=\int_R^\infty dr'\,r'\widehat{\mathscr{L}}(r,r')\widehat{C}\bar{p}_1(r'),\quad r<R,$$
$$\bar{f}_1(r)=0,\quad r<R, \tag{38}$$

and Eq. (37) can be rewritten in the form

$$\bar{p}_1(r)=(\widehat{L}\widehat{C}\bar{p}_1)(r)+[(-\widehat{L}_R^{(0)}-\widehat{K}_R)\widehat{C}\bar{p}_1](r)+\bar{f}_1(r), \tag{39}$$

where

$$(\widehat{L}_R^{(0)}\widehat{C}\bar{p}_1)(r)=\int_0^R dr'\,r'\widehat{\mathscr{L}}(r,r')\widehat{C}\bar{p}_1(r'),$$

$$(\widehat{K}_R\widehat{C}\bar{p}_1)(r)=\int_R^\infty dr\,r'\widehat{\mathscr{K}}_R(r,r')\widehat{C}\bar{p}_1(r'),$$

and the matrix $\widehat{\mathscr{K}}_R(r,r')$ is defined as

$$\widehat{\mathscr{K}}(r,r')=\begin{cases}\dfrac{4}{\pi}\displaystyle\int_{(r^2-R^2)^{1/2}+(r'^2-R^2)^{1/2}}^{r+r'}\times\frac{d\rho}{\{[\rho^2-(r-r')^2][(r+r')^2-\rho^2]\}^{1/2}}\times\int_\rho^\infty dt\,\widehat{K}_0(t)\widehat{\Sigma}, & r\geqslant R,\ r'\geqslant R,\\ 0, & r<R\ \text{or}\ r'<R.\end{cases}$$

Similarly, as in the case of $\bar{f}_1(r)$, one can check that

$$[(-\widehat{L}_R^{(0)}-\widehat{K}_R)\widehat{C}\bar{p}_1](r)=O(e^{-r});$$

hence (39) can be viewed as the equation of type (32) with the same $O(e^{-r})$ behavior of the source term

$$[(-\widehat{L}_R^{(0)}-\widehat{K}_R)\widehat{C}\bar{p}_1](r)+\bar{f}_1(r).$$

According to (39) and (35) we obtain formally

$$\bar{p}_1(r)=\sum_{i=1}^{\kappa}\beta_iK_0(r/\nu_i)\bar{U}(\nu_i)+O(e^{-r}), \tag{40}$$

where

$$\beta_i=[1/(\nu_iN(r_i))]\int_0^\infty dr\,r\langle\bar{n}_i(r)|\widehat{C}\times[\bar{f}_1(r)-(\widehat{L}_R^{(0)}\widehat{C}\bar{p}_1)(r)-(\widehat{K}_R\widehat{C}\bar{p}_1)(r)]\rangle, \tag{41}$$

$$\bar{n}_i(r)=I_0(r/\nu_i)\bar{U}(\nu_i). \tag{42}$$

Substituting the explicit definition of $\widehat{L}_R^{(0)}$, changing the order of integration, and using (30) and (38) yields the identity

$$\int_0^\infty dr\,r\langle\bar{n}_i(r)|\widehat{C}(\widehat{L}_R^{(0)}\widehat{C}\bar{p}_1)(r)\rangle=\int_R^\infty dr\,r\langle\widehat{C}\bar{p}_1(r)|(\widehat{L}_R^{(0)}\widehat{C}\bar{n}_i)(r)\rangle.$$

Similarly

$$\int_0^\infty dr\,r\langle\bar{n}_i(r)|\widehat{C}(\widehat{K}_R\widehat{C}\bar{p}_1)(r)\rangle=\int_R^\infty dr\,r\langle\widehat{C}\bar{p}_1(r)|(\widehat{K}_R\widehat{C}\bar{n}_i)(r)\rangle.$$

These results used in (41) give

$$\beta_i=[1/(r_iN(r_i))]\left[\int_R^\infty dr\,r\langle\widehat{C}\bar{n}_i(r)|\bar{f}_1(r)\rangle+\int_R^\infty dr\,r\langle\widehat{C}\bar{p}_1(r)|\bar{f}_i(r)\rangle\right], \tag{43}$$

where according to (37b), (38), and (42) the notation

$$\bar{f}_i(r)=[(\widehat{I}-\widehat{L}_R\widehat{C})\bar{n}_i](r)=-[(\widehat{L}_R^{(0)}+\widehat{K}_R)\widehat{C}\bar{n}_i](r),\quad r\geqslant R,$$
$$\bar{f}_i(r)=0,\quad r<R,$$

is introduced. The source term $\bar{f}_2(r)$ corresponds to the fictitious Milne problem with $\bar{n}_2(r)$ as the unbounded part of the solution $\bar{n}_2(r)+\bar{p}_2(r)$, where the equation for $\bar{p}_2$ is

TABLE I. The values of the extrapolation distances in two cases for $R = 1$. The given results are exact in the sense of the numerical evaluation of the matrix elements of $\hat{I} - \hat{L}_R \hat{C}$.

	Case, references	1 [6]	2 [4]	
Constants characterizing the material	κ	1	2	
	ν_1	1.9360	2.1568	
	ν_2		1.3566	
	k_{BMS}	0.9164	0.9356	
	$C = \det \hat{C}$	0.2580	0.2469	
	Extrapolation distances	$\lambda = \lambda_1 = \lambda_2$	λ_1	λ_2
Last coefficient included	$\tilde{\beta}_1$	0.9554	0.6996	0.9827
	$\tilde{\beta}_2$		0.7281	0.9730
	γ_2	0.9611	0.7311	0.9771
	γ_4	0.9617	0.7318	0.9777

$$\bar{p}_2(r) = (\hat{L}_R \hat{C} \bar{p}_2)(r) + \bar{f}_2(r). \tag{44}$$

The vectors $\bar{p}_1, \bar{p}_2$ and $\bar{f}_1, \bar{f}_2$ belong to the Hilbert space $L^2_r(R,\infty)$ with the scalar product $\langle \mid \rangle_L$:

$$\bar{h},\bar{g} \in L^2_r(R,\infty): \langle \bar{h} | \bar{g} \rangle_L = \int_R^\infty dr\, r \langle \bar{h}(r) | \bar{g}(r) \rangle$$

and the scalar products $[\ |\ ]_L$ and $\{\ |\ \}_L$ are introduced along the way of Sec. 3. Now we observe that

$$\beta_1 = \frac{1}{\nu_1 N(\nu_1)} ([\bar{n}_1 | \bar{f}_1]_L + [\bar{p}_1 | \bar{f}_1]_L)$$

depends on the $\bar{p}_1$ via the minimum of the functional

$$\mathscr{F}_1(\bar{p}) = \tfrac{1}{2}\{\bar{p} | \bar{p}\}_L - [\bar{p} | \bar{f}_1]_L$$

in $L^2_r(R,\infty)$,

$$\beta_1 = \frac{1}{\nu_1 N(\nu_1)} ([\bar{n}_1 | \bar{f}_1]_L - 2\mathscr{F}_1), \tag{45}$$

where

$$\mathscr{F}_1 \equiv \min_{\bar{p} \in L^2_r(R,\infty)} \mathscr{F}_1(\bar{p}).$$

In the expression for β_2

$$\beta_2 = \frac{1}{\nu_2 N(\nu_2)} ([\bar{n}_2 | \bar{f}_1]_L + [\bar{p}_1 | \bar{f}_2]_L);$$

there is the mixed term $[\bar{p}_1 | \bar{f}_2]_L$ From (37a), (44) and the symmetry of $\hat{L}_R \hat{C}$ in the scalar product $[\ |\ ]_L$ it stems:

$$[\bar{p}_1 | \bar{f}_2]_L = [\bar{p}_2 | \bar{f}_1]_L.$$

Let us construct the functional $\mathscr{F}_{(1+2)}(\bar{p})$ corresponding to the equation obtained by adding (37a) and (44);

$$\mathscr{F}_{(1+2)}(\bar{p}) = \tfrac{1}{2}\{\bar{p} | \bar{p}\}_L - [\bar{p} | \bar{f}_1 + \bar{f}_2]_L.$$

TABLE II. The comparison of the results of the extrapolation distance in case 1 for $R = 10$, obtained by the variational method and from the asymptotic expansion (48). In the latter the following notation is used: $\lambda^{(0)} = \rho_0$, $\lambda^{(1)} = \lambda^{(0)} + \rho_1/R$, $\lambda^{(2)} = \lambda^{(1)} + \rho_2/R^2$, $\lambda^{(3)} = \lambda^{(2)} + \rho_3 \log R/R^3$.

	Variational method		Asymptotic expressions, Eq. (48)	
Last coefficient included			$\lambda^{(0)}$	0.7244
	$\tilde{\beta}_1$	0.7490		
			$\lambda^{(1)}$	0.7507
	γ_2	0.7513		
			$\lambda^{(2)}$	0.7518
	γ_4	0.7515		
			$\lambda^{(3)}$	0.7517

The symmetry of the mixed terms gives

$$[\bar{p}_1 | \bar{f}_2]_L = -\mathscr{F}_{(1+2)} + \mathscr{F}_1 + \mathscr{F}_2,$$

where

$$\mathscr{F}_{(1+2)} = \min_{\bar{p} \in L^2_r(R,\infty)} \mathscr{F}_{(1+2)}(\bar{p}),$$

$$\mathscr{F}_2 = \min_{\bar{p} \in L^2_r(R,\infty)} \mathscr{F}_2(\bar{p}),$$

$$\mathscr{F}_2(p) = \tfrac{1}{2}\{\bar{p} | \bar{p}\}_L - [\bar{p} | \bar{f}_2]_L,$$

and the expression for β_2 is

$$\beta_2 = \frac{1}{\nu_2 N(\nu_2)} ([\bar{n}_2 | \bar{f}_1]_L - \mathscr{F}_{(1+2)} + \mathscr{F}_1 + \mathscr{F}_2). \tag{46}$$

The formulas (45) and (46) were utilized from computing the extrapolation distances λ_i in a number of cases.[17] The Ritz method was used with the trial vectors of the form

$$\bar{p}(r) = \sum_{i=1}^{\kappa} \tilde{\beta}_i K_0\left(\frac{r}{\nu_i}\right) \bar{U}(\nu_i) + \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \int_0^1 d\nu\, K_0\left(\frac{r}{\nu}\right) + \begin{pmatrix} \gamma_3 \\ \gamma_4 \end{pmatrix} \int_0^1 d\nu\, \nu K_0\left(\frac{r}{\nu}\right), \tag{47}$$

suggested by the replication property[13,10] of the integral transport equations. The terms in (47) were included successively and minimization was performed with respect to the coefficients $\tilde{\beta}$ and γ present in the terms taken into account. Here in Table I as an example we show the values of λ_i obtained in one case with $\kappa = 1$ [6] and another with $\kappa = 2$, [4] both for $R = 1$. The numerical integration in the matrix elements of the operator $\hat{I} - \hat{L}_R \hat{C}$ with the trial functions and $\bar{n}_i(r)$ were performed using the Kronrod–Gauss method.[25]

For comparison in the Table II we present also the results of case 1 for $R = 10$, obtained by the variational method and from the asymptotic expansion of the form[17]

$$\lambda(R) = \rho_0 + \rho_1/R + \rho_2/R^2 + \rho_3 \log R/R^3 + O(1/R^3). \tag{48}$$

ACKNOWLEDGMENTS

The author whishes to thank Professor R. Żelazny for suggesting this problem and helpful discussions. Thanks are also due to Dr. R. Stankiewicz for reading the manuscript and for many critical remarks.

[1]R. Żelazny and A. Kuszell, Ann. Phys. **10**, 81 (1961).
[2]C.E. Siewert and P.F. Zweifel, Ann. Phys. **36**, 61 (1966).
[3]C.E. Siewert and P.S. Shieh, J. Nucl. Energy **21**, 383 (1967).
[4]W. Baran, Nukleonik **11**, 10 (1968).
[5]P. Silvennoinen and P.F. Zweifel, J. Math. Phys. **13**, 1114 (1972).
[6]C.E. Siewert and Y. Ishiguro, J. Nucl. Energy **26**, 251 (1972).
[7]E.E. Burniston, T.W. Mullikin, and C.E. Siewert, J. Math. Phys. **13**, 1461 (1972).
[8]C.E. Siewert, E.E. Burniston, and J.T. Kriese, J. Nucl. Energy **26**, 469 (1972).
[9]T. Duracz, Atomkernenergie, Bd. 28, Lfg. 3, 161 (1976).
[10]R. Żelazny, T. Duracz, J. Math. Phys. **15**, 1505 (1974).
[11]K.M. CAse and P.F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967).
[12]G. Mitsis, ANL Report 6459, Argonne National Lab., 1963.
[13]O.J. Sheaks, J. Math. Phys. **13**, 203 (1972).
[14]B. Davison, "*Neutron Transport Theory*" (Clarendon, Oxford, 1957).
[15]A. Kavenoky, Nucl. Sci. Eng. **65**, 514 (1978).
[16]B. Davison, Proc. Phys. Soc. A **64**, 881 (1951).
[17]T. Błeński, thesis, Institute of Nuclear Research, Warsaw (1978).
[18]N. Dunford and J. Schwartz, *Linear Operators, Part I* (Interscience, New York, 1958).
[19]M.M. Vainberg, *Variational Method and Method of Monotonic Operators* (Nauka, Moscow, 1972) (in Russian).
[20]P. Lankaster, *Theory of Matrices* (Academic, New York, London, 1969).
[21]S.G. Mikhin, *Variational Methods in Mathematical Physics* (Nauka, Moscow, 1970) (in Russian).
[22]V.S. Vladimirov, "Mathematical Problems in the One Velocity Theory of Particle Transport," Tr. Mat. Akad. Nauk SSSR **61**, 1 (1961) (in Russian).
[23]K. Yosida, *Functional Analysis* (Springer-Verlag, Berlin, Göttingen, Heidelberg, 1965).
[24]G.I. Bell and S. Glasstone, *Nuclear Reactor Theory* (Van Nostrand-Reinhold, New York, 1970).
[25]W. Squire, *Integration for Engineers and Scientists* (Elsevier, New York, London, Amsterdam, 1970).

A homogeneous Hilbert problem for the Kinnersley–Chitre transformations [a)]

Isidore Hauser and Frederick J. Ernst
Department of Physics, Illinois Institute of Technology, Chicago, Illinois 60616

(Received 8 November 1979; accepted for publication 24 December 1979)

A homogeneous Hilbert (Riemann) problem (HHP) is introduced for carrying out the Kinnersley–Chitre transformations of the set V of all axially symmetric stationary vacuum spacetimes, and the spacetimes which are like the axially symmetric stationary ones except that both Killing vectors are spacelike. A proof, which is independent of the Kinnersley–Chitre formalism, establishes that the HHP transforms the potential (for certain closed self-dual 2 forms) $F_0(\mathbf{x}, t)$ of any given member of V into the potential $F(\mathbf{x}, t)$ of another member of V. Two illustrative examples involving the Minkowski space $F_0(\mathbf{x}, t)$ are given. The representation used for the Geroch group K, the singularities and gauge of the potentials, and possible applications of the HHP are discussed.

1. INTRODUCTION

Our principal objective in this paper is to introduce a homogeneous Hilbert (Riemann) problem[1] each of whose solutions gives the result of applying an element u of the Kinnersley–Chitre representation of the Geroch group K.[2-7] The operand of u is a potential F_0 which is determined up to a gauge transformation by a given axially symmetric stationary vacuum metric g_0. The result F of the operation is also a potential from which a new axially symmetric stationary vacuum metric g can be computed.

It is worth mentioning that our HHP (homogeneous Hilbert problem) is also applicable to those vacuum spacetimes which are like the axially symmetric stationary ones except that both Killing vectors are spacelike. We shall let V denote the set of all vacuum spacetimes for which there exist coordinates x^1, x^2, x^3, x^4 such that the line element has the form (signature $= +2$)

$$g_{ij}\delta x^i\delta x^j + g_{ab}\delta x^a\delta x^b \quad (i,j=1,2)\ (a,b=3,4), \tag{1}$$

where g_{ij} and g_{ab} depend at most on x^1, x^2, and where the 2×2 matrix h whose elements are

$$h_{ab} := g_{ab} \quad (a,b=3,4) \tag{2}$$

obeys the condition that $d(\det h)$ is not zero and is not a null 1 form.[8] The axially symmetric stationary members of V are among those which satisfy $\det h < 0$. This paper will explicitly cover the entire set V.

There are eight reasons why we consider the HHP to be an attractive way of effecting the K–C (Kinnersley–Chitre) transformations,[5-7] and we shall now take these up in an appropriate order beginning with some key features of our gauge and our group representation.

(1) First the gauge of the potentials which are transformed by the HHP is elegantly defined in terms of their complex plane singularities. These potentials were originally introduced by K–C[6] (with minor differences of notation and signature) as generating functions for part of their hierarchy of potentials. They are 2×2 complex matrix functions $F(x^1, x^2, t)$ of the nonignorable coordinates x^1, x^2 and a complex variable t. As will be discussed in Sec. 3, we select a gauge so that for fixed $\mathbf{x} = (x^1, x^2)$, $F(\mathbf{x}, t)$ is holomorphic[9] in a neighborhood of $t = 0$, and

$$F(\mathbf{x},t)\begin{pmatrix} t & 0 \\ 0 & 1\end{pmatrix}$$

is holomorphic in a neighborhood of $t = \infty$ (including ∞).[10] *Further restrictions of the gauge involve a minimization of the number of complex plane singularities* and are discussed later in this Introduction. Regardless of the specific gauge, the property of F which is important from the viewpoint of its use in the HHP stems from its holomorphy at $t = 0$; we are referring here to the fact that *if* $\mathbf{x}$ *is restricted to any given compact region of the real plane, then there exists at least one smooth contour L surrounding the origin in the complex plane such that F is holomorphic on $L + L_+$ where L_+ denotes the open set inside L.*

(2) The HHP employs strikingly simple representations K_L of the group K, one for each smooth contour L surrounding the origin in the complex plane and symmetric with respect to the real axis. K_L is the set of all 2×2 complex matrix analytic functions $u(t)$ of a complex variable t such that

$$\det u(t) = 1, \quad u(t)^\dagger \epsilon\, u(t) = \epsilon, \tag{3}$$

$$\epsilon := \begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}, \quad u(t)^\dagger + := \text{h.c. of } u(t^*), \tag{4}$$

and such that $u(t)$ is holomorphic on L, and

$$\begin{pmatrix} t^{-1} & 0 \\ 0 & 1\end{pmatrix} u(t) \begin{pmatrix} t & 0 \\ 0 & 1\end{pmatrix},$$

is holomorphic at $t = \infty$. (We always use the Riemann sphere topology; holomorphy at a point means holomorphy in a neighborhood of the point, and holomorphy in or on a set means holomorphy at every point of the set. The condition at $t = \infty$ derives from the gauge condition satisfied by F at $t = \infty$ and the requirement that FuF^{-1} be holomorphic at $t = \infty$; this is discussed in Sec. 4.) To effect a transformation of any given potential $F_0(x, t)$ for any given $u(t)$ in K_L, we restrict x to a compact region U_c of the real plane such that $F_0(x, t)$ is holomorphic on $L + L_+$ for every x in U_c. Then, for any fixed x in U_c, we apply the HHP with F_0 and u as input data. *The output F potential is automatically also holomorphic on $L + L_+$ for every $\mathbf{x}$ in U_c.*

[a)] Work supported in part by National Science Foundation Grants PHY 75-08750 and PHY 79-08627.

(3) The composition law of the group representation is simply the 2×2 matrix product

$$u_3(t)=u_2(t)u_1(t),$$

where u_1, u_2, and u_3, respectively, transform F_0, into F_1, F_1 into F_2, and F_0 into F_2. This is a nontrivial plus for the representation if we consider the fact that K is an infinite dimensional group.[3] As regards the connection between the representation and the SU (1,1) symmetry[11] of the field equation satisfied by the potential $\mathscr{E}$ of Ernst, observe that Eq. (3) states that $u(t)$ is a member of SU (1,1) for every real value of t.

(4) Another advantage of the group representation is that it is already exponentiated, i.e., we don't have to go through a process of discovering the one parameter group elements corresponding to a given generator everytime we want to carry out a transformation. Also, each $u(t)$ is easily factorized into convenient one parameter group representations. For example, suppose u_3^4 is not identically zero (where u_b^a is in the ath row and b th column). Then u_3^4 has at most a finite number of isolated zeros on L. We can, without loss of generality, subject L to an arbitrarily small deformation so as to avoid these zeros, whereupon we have the following on the new L:

$$u(t)=\begin{pmatrix}1 & \beta\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ u_3^4 & 1\end{pmatrix}\begin{pmatrix}1 & \alpha\\ 0 & 1\end{pmatrix},$$

$$\alpha=(u_3^4)^{-1}(u_4^4-1),\quad \beta=(u_3^4)^{-1}(u_3^3-1).$$

If u_3^4 is identically zero, a like procedure is available with some other element playing the role of a divisor.

(5) The residual gauge transformations are given by the simple mapping

$$F(\mathbf{x},t)\rightarrow F(\mathbf{x},t)v(t),$$

where $v(t)$ is defined like $u(t)$ except that "holomorphy on L" is replaced by holomorphy at $t=0$, and $v(0)=I$. For all choices of the guage, $F(x,t)$ has singularities in the complex plane at the zeros of[12]

$$\lambda(\mathbf{x},t):=[(1-2zt)^2\pm(2t\rho)^2]^{1/2}, \tag{5}$$

where

$$\rho:=|\det h|^{1/2},\quad \pm:=\operatorname{sgn}(-\det h); \tag{6}$$

ρ and z are conjugate harmonic fields in the two dimensional (x^1,x^2) space. The zeros of $\lambda(\mathbf{x},t)$ are generally branch points of $F(\mathbf{x},t)$, and part of the definition of the gauge is a selection of an appropriate branch cut (one which avoids $t=0$). In Sec. 4, we shall show that the gauge can be selected so that the only other t- singularities of $F(\mathbf{x},t)$, if they exist at all, occur at the zeros of $\lambda(\mathbf{x}_0,t)$, where $\mathbf{x}_0$ is used as an initial point in the process of integrating the differential equation which defines $F(\mathbf{x},t)$. For asymptotically flat axially symmetric stationary spacetimes with $\mathbf{x}$ restricted to finite values in a neighborhood of $\rho^2+z^2=0$, we can select a gauge so that the only t singularities are at the zeros of $\lambda(\mathbf{x},t)$; that topic will be covered in a sequel to this paper.

(6) We can now employ the well developed mathematical apparatus of complex analysis to help us obtain new axially symmetric stationary solutions of the Einstein field equations. It should be mentioned that our HHP shares this advantage (and the preceding five ones) with an equivalent linear integral equation of the Cauchy type,[13,14] which was previously discovered by the authors and which will be related to the HHP later in the Introduction.

(7) We feel that the HHP has brought us close to proving some informal conjectures made by various workers in the field. For example, if we are given any two axially symmetric stationary members of V with corresponding potentials F_0 and F, does a u always exist which transforms F_0 into F? Work by Hoenselaers, Kinnersley, and Xanthopoulos[15] strongly indicates that such is the case if the members of V are both asymptotically flat. In Sec. 6, we shall use the HHP to point out that the more general question is equivalent to the problem as to whether there exists a $u(t)$ such that

$$F(\mathbf{x},t)u(t)F_0(\mathbf{x},t)^{-1},$$

is holomorphic in a neighborhood of $t=\infty$, which contains the zeros of $\lambda(\mathbf{x},t)$. The problem is thus reduced to a one in complex analysis, and it appears to us as if its solution is imminent.

(8) The HHP has been extensively used in other fields, especially in particle physics. Moreover, *our* HHP *appears to be a link between the group theoretical*[2–7] *and the recent soliton approaches*[16–19] to exact solution research, though the authors do not yet understand the details of this link.

As regards these last remarks, the recent studies of exact solutions of the Einstein field equations for members of V have taken two principal paths. One is based on the program of using the symmetries of the field equations to generate all or at least a healthy chunk of the space–times from the known ones and, in particular, from Minkowski space. This is the idea initiated by Geroch,[2,3] who discovered K. The idea was then taken up by Kinnersley and Chitre,[4–7] who constructed a useful representation of the generators and set up viable methods for exponentiating them in many cases.

The second approach,[16–19] with which the authors are less familiar, has been based on the application of soliton concepts and techniques. This approach has neglected the group aspect, but it has been clearly successful in relating the exact solution studies of general relativity to other fields of physics.

In our own work, the group theoretical methods of Geroch and K–C have been our starting point. However, our use of complex analysis and our introduction of an HHP seems to have brought us closer to a synthesis of the two approaches.

Now, let us consider some specifics. In a previous paper,[13] we found that the K–C transformations can be effected by solving a linear integral equation of the Cauchy type. This equation can be expressed in the form

$$\int_L ds\,\frac{F(s)u(s)F_0(s)^{-1}}{s(s-t)}=0, \tag{7}$$

subject to the boundary condition

$$F(0)=i\epsilon:=i\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}. \tag{8}$$

The complex variable s lies on L, whereas t is within L. The member $u(s)$ of K_L depends only on s. $F_0(s)$ and $F(s)$ depend on $\mathbf{x}$ as well as on s, but we often suppress the dependence on

$\mathbf{x}$ when this can be done without danger of ambiguity. If h as defined by Eq. (2) is given, F can be computed by solving a pair of linear differential equations subject to gauge conditions; K–C[6,7] derived an expression for the F potentials of the Zipoy–Voorhees metrics in this way.

The F potentials can also be obtained as the solutions of Eqs. (7) and (8) for given u and F_0. In this way, the authors[13] derived the F potential of the Kerr–NUT metric by applying the general B group (a subgroup of K) element u to a Schwarzschild F_0; this was an extension of prior work by K–K[7] who used the same transformation (in their own representation) to derive the Ernst potential of the Kerr–NUT metric. In a similar extension of results due to Hoenselaers, Kinnersley, and Xanthopoulos,[15] the authors[14] derived the general expression[20] for F corresponding to any given F_0 and to a $u(t)$ which has the form

$$u(t) = \begin{pmatrix} 1 & t\beta(t) \\ 0 & 1 \end{pmatrix},$$

where the only singularities of $\beta(t)$ are N simple poles inside L. This result[20] is apparently "isomorphic" to the N-soliton solution of Belinskii and Zakharov,[16] but we have not worked out the details. The F where $\beta(t)$ has two poles at the origin of respective order one and two has also been derived by us as a special case of an electrovac generalization of our work.[14]

All of the F potentials which have been mentioned above can be used as input potentials for further transformations. A definition of the F potential which does not presuppose any knowledge of the K–C hierarchy of potentials[5] will be given in Sec. 3, which will also include a discussion of the gauge conditions.

It is clear that Eqs. (7) and (8) are equivalent[21] to the HHP

$$X_-(x) = X_+(s)G(s), \tag{9}$$

where

$$G(s) := F_0(s)u(s)F_0(s)^{-1}, \quad F(s) := X_+(s)F_0(s), \tag{10}$$

$$X_+(0) = I = \text{unit } 2\times 2 \text{ matrix}, \tag{11}$$

and where, for fixed x,[22]

(1) $X_\pm(s)$ are (respectively, for $\pm$) the boundary values on L of 2×2 matrix functions $X_\pm(t)$ which are continuous in $L + L_\pm$ and holomorphic in $L_\pm$. (L_- is the complement of $L + L_+$; as usual, we include ∞ in L_-.)

(2) The inverse of $X_\pm(t)$ exists (respectively, for $\pm$) at all t in $L + L_\pm$.

In other words, $X_\pm$ is a fundamental solution of a HHP with component indices (0, 0); the boundary conditions at $t = 0$ and $t = \infty$ pin down the solution uniquely.

In connection with Eq. (7) or (9), there remain two existence problems which have their counterparts in the K–C formalism (to which ours is equivalent). First, there is the question as to whether a solution which fits the boundary condition at $t = \infty$ exists for *all choices* of F_0 and u. In other words, are the component indices of the HHP for $G(s)$ always (0, 0)? We shall have to accept this existence as a *working hypothesis*, since we have no proof as yet. Likewise, we shall have to accept the statement that $\partial X_\pm(t)/\partial x^i$ exist, are continuous in $L + L_\pm$, and are holomorphic in $L_\pm$ (respectively, for $\pm$) as a *working hypothesis*. However, in Sec. 6, we shall reduce these existence problems to ones which involve a simple $u(t)$ and appear amenable to analysis.

Our original derivation of Eq. (7) and therefore, of the equivalent HHP was an extremely long and cumbersome one based on the K–C representation of K. A *second objective of this paper is to supply a relatively short elegant derivation of* Eq. (7) which does not presuppose the K–C formalism. Specifically, in Sec. 4, we shall (granting the working hypotheses which were stated above) prove the following theorem:

Theorem: The solution $F(t)$ of Eqs. (9) to (11) is an F potential of a member of V (as defined in Sec. 3). The Ernst potential $\mathscr{E}$ and the metric components h of the members of V are computed from[23]

$$\mathscr{E} = H_{44}, \quad h = -\operatorname{Re}H, \quad H = \dot{F}(0), \tag{12}$$

where $\dot{F}(t) := \partial F(t)/\partial t$; the remaining metric components can be determined from $\mathscr{E}$ or from h by methods which have been given, for example, by Ernst[24] and by Kinnersley.[24]

By proving the above theorem, we shall have obtained a simple derivation of our entire formalism in a relatively few simple strokes. Once again, we stress that the K–C formalism[4–7] is equivalent to ours and that we originally derived[13] ours from theirs. The new derivation given in this paper is a convenient case of hindsight. That does not, however, diminish its importance.

As regards applications of the HHP, we shall give two simple ones in Sec. 5. We shall prove that the Minkowski space F_0 is left unchanged if we apply that element of K which corresponds to

$$u(t) = \exp[\gamma(t)\epsilon], \tag{13}$$

where

$$\gamma(t) = \alpha(t)\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \quad (B\text{-group generator}), \tag{14}$$

and where $\alpha(t)$ is holomorphic on $L + L_-$. Then we shall obtain the general static Weyl metric F from the Minkowski space F_0 by applying that element of K which corresponds to

$$\gamma(t) = \xi(t)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{15}$$

where $\xi(t)$ is holomorphic in $L + L_-$. These results are not sensational, since K–C[6,7] have already obtained them in their representation. However, the brevity of the derivations using the HHP is startling and makes them worthwhile. As regards less trivial applications, these will have to be deferred for a sequel, and the same remark holds for the electrovac generalization of this paper; we are still not close to completing the details on those subjects.

To be able to define the F potential of a given member of V, it is first necessay to introduce the coefficient H of the first degree term in the power series expansion

$$F(t) = \Omega + Ht + \cdots, \quad \Omega := i\epsilon.$$

This 2×2 complex matrix field H is a simple generalization of the potential $\mathscr{E}$ of Ernst. In the next section, we shall define H and derive those equations which govern it and which we shall need in later sections.

2. THE *H* POTENTIAL

The superscripts *, T, and † will, respectively, denote the complex conjugate, transpose, and Hermitian conjugate operations. The wedge symbol $\wedge$ will be omitted in exterior products of forms; thus, $\omega\eta$ and $d\omega$ mean $\omega\wedge\eta$ and $d\wedge\omega$, respectively, for any p- form ω and q- form η. If $p\leqslant q$, the $(q-p)$ form[25] obtained from a maximal contraction of ω with η is denoted by $\omega\ulcorner\eta$; two useful relations are given by

$$\omega\ulcorner\eta=\omega\cdot\eta,\quad \text{if } p=q=1,$$

$$\lambda\ulcorner(\omega\eta)=\omega(\lambda\ulcorner\eta)+(-1)^{q}(\lambda\ulcorner\omega)\eta,$$

if λ is a 1 form.

X will denote the column matrix whose elements are the covectors (1-forms) X_3 and X_4 of the two Killing vectors which characterize the given member of V under consideration in this section. Also, we let W be that self-dual 2-form which is defined by the equation

$$-2\,dX=W+W^*. \tag{16}$$

Since the spacetime is a vacuum,

$$dW=0.$$

Therefore,

$$\mathscr{L}_{\mathbf{x}_a}W=X_a\ulcorner dW-d(X_a\ulcorner W)=-d(X_a\ulcorner W)=0.$$

Therefore, there exists a complex matrix field H such that

$$dH=X\ulcorner W^T. \tag{17}$$

Next, note that a definition of the matrix h which is equivalent to Eq. (2) is given by

$$h:=X\ulcorner X^T. \tag{18}$$

Take the real part of Eq. (17), and use Eqs. (16), (18), and

$$\mathscr{L}_{\mathbf{x}_a}X_b-=X_a\ulcorner dX_b+d(X_a\ulcorner X_b)=0.$$

It follows that

$$d(\mathrm{Re}H)=-dh.$$

We choose the additive constant in ReH so that the integral of the above equation is

$$\mathrm{Re}H=-h. \tag{19}$$

Note that ReH is now symmetric. The additive constant in ImH will remain arbitrary until the details of a more specific problem suggest a good choice.

We next derive a differential equation for H which is equivalent to the self-duality of W. From Eqs. (1) and (2),

$$X=h\,dx,\quad dx:=\begin{pmatrix}dx^3\\dx^4\end{pmatrix}.$$

Therefore, Eq. (16) implies that $W+W^*$ is a linear combination of the 2-forms $dx^i dx^a$ $(i=1,2)$ $(a=3,4)$. Since W is self-dual, it must also be equal to a linear combination of these 2-forms. Therefore, there exists a 2×2 one-form matrix K such that

$$W=K\,dx,\quad X_a\ulcorner K=0.$$

Upon inserting the above into Eq. (17) and using the fact that

$X\ulcorner dx^T=I:=$ unit 2×2 matrix,

we obtain $K=dH^T$; so

$$W=dH^T dx. \tag{20}$$

We now apply the duality operator to the above expression. Let the covectors of the coordinate tetrad $\partial/\partial x^\alpha$ $(\alpha=1,2,3,4)$ be designated by X_α, and let

$$\rho:=|\det h|^{1/2},\quad e^{2\Gamma}:=|g_{11}g_{22}-(g_{12})^2|^{1/2}. \tag{21}$$

Then the duality operation on W is[25]

$$W\ulcorner(X_1X_2X_3X_4e^{-2\Gamma}\rho^{-1})=iW.$$

Upon substituting from Eq. (20) into the above, we get the relation which we are seeking, viz.,

$$-\rho^{-1}h\epsilon{}^*dH=i\,dH, \tag{22}$$

where * is a two *dimensional duality operator* defined by

$$^*dx^i:=dx^i\ulcorner(X_1X_2e^{-2\Gamma}). \tag{23}$$

If we select conjugate harmonic coordinates such that

$$g_{12}=0,\quad g_{11}=\pm g_{22}=\exp^{2\Gamma}, \tag{24}$$

then

$$^*dx^1=\mp dx^2,\quad ^*dx^2=dx^1, \tag{25}$$

where, as will be our usual practice, *the top and bottom signs refer to the cases* $\det h<0$ *and* $\det h>0$, respectively. Observe that Eqs. (25) imply

$$^{**}=\mp1, \tag{26}$$

$$(^*\alpha)\beta=-\alpha(^*\beta)\ \text{for any 1-forms } \alpha \text{ and } \beta. \tag{27}$$

We shall next derive a number of useful relations from Eq. (22). Observe that the definition (21) of ρ is equivalent to the equation

$$h\epsilon h=\mp\rho^2\epsilon. \tag{28}$$

We introduce the field z which is defined by

$$iz\epsilon:=\tfrac12(H-H^T), \tag{29}$$

and which is real according to Eq. (19). By taking the antisymmetric part of the real part of Eq. (22) and by using Eqs. (19), (28), and (29), we obtain

$$^*d\rho=\pm dz. \tag{30}$$

Furthermore, with the aid of Eqs. (19) and (29), we express Eq. (22) in the equivalent form

$$2(z\pm\rho^*)dH=(H+H^\dagger)\Omega dH\quad(\Omega:=i\epsilon). \tag{31}$$

Equation (31) will be called the *self-duality relation*; it will be the key starting point of our work in Sec. 3. To gain a better grasp of this equation and to help us deduce some useful relations from it, we define null coordinates x_A and corresponding fields r_A by

$$X_A:=x^1+Ax^2,\quad A:=[\mathrm{sgn}(\det h)]^{1/2}=\pm1\ \text{or}\ \pm i; \tag{32}$$

$$r_A:=z+A\rho\ \text{if}\ A=\pm i,\quad r_A:=z-A\rho\ \text{if}\ A=\pm1.$$

From Eqs. (25), (26), and (30), observe that

$$^*dx_A=A\,dx_A,\quad ^*dr_A=A\,dr_A,$$

$$\partial r_A/\partial x_B=0,\quad \text{if}\quad A\neq B. \tag{33}$$

From Eqs. (33), it follows that the self-duality relation (31) is equivalent to the pair of equations

$$2r_A \frac{\partial H}{\partial x_A} = (H + H^\dagger)\Omega \frac{\partial H}{\partial x_A} \quad \text{(no sum)}. \tag{34}$$

Multiply Eq. (34) on the left by $(\partial H^T/\partial x_A)\Omega$, and take the antisymmetric part of the result; we get, with the aid of Eqs. (29) and (32),

$$\frac{\partial H^T}{\partial x_A}\Omega \frac{\partial H}{\partial x_A} = 0 \quad \text{(no sum)}.$$

Upon using Eq. (29) to replace H^T in the above equation, we get the important result

$$\left(\frac{\partial H}{\partial x_A}\Omega\right)^2 = \frac{\partial r_A}{\partial x_A}\left(\frac{\partial H}{\partial x_A}\Omega\right) \quad \text{(no sum)}. \tag{35}$$

Next, multiply Eq. (34) on the left by $(\partial H^\dagger/\partial x_{-A})\Omega$ to get a first expression; then get a second expression by subjecting the first one to Hermitian conjugation followed by the script substitution $A \to -A^*$. Take the difference of these two expressions, and one gets

$$\frac{\partial H^\dagger}{\partial x_{-A}}\Omega \frac{\partial H}{\partial x_A} = 0 \quad \text{(no sum)}. \tag{36}$$

The above pair of equations is clearly equivalent to the pair

$$dH^\dagger \Omega dH = dH^\dagger \Omega * dH = 0, \tag{37}$$

which could also have been obtained from Eq. (20) and the fact that the exterior product of any self-dual 2-form with its complex conjugate is zero. That completes all of the relations involving H which we shall need in this paper.

It is time to review the residual arbitrariness in the 2×2 complex potential $H(\mathbf{x})$.

(1) We shall usually select x^1, x^2 so that Eqs. (24) hold. Then the line element of the two dimensional Riemannian space with the metric g_{ij} becomes

$$e^{2\Gamma}[(dx^1)^2 \pm (dx^2)^2]. \tag{38}$$

There remain the *conformal coordinate transformations* which preserve the above form.

(2) The exterior product of the Killing vectors has an arbitrary multiplicative constant, and the antisymmetric part of H has an arbitrary additive constant. Transformations which are equivalent to changes in these constants are

$$\rho \to (\exp b)\rho, \quad z \to z + c, \tag{39}$$

where b, c are any real numbers.

(3) $H - H^\dagger = i\,\mathrm{Im}(H + H^T)$ has an arbitrary additive constant. The corresponding transformation is

$$H - H^\dagger \to H - H^\dagger + 2iB, \tag{40}$$

where B is any 2×2 real symmetric constant.

(4) Finally, X can be subject to the SL(2, R) transformations

$$X \to SX, \quad \det S = 1, \quad dS = 0,$$

which induce the mapping

$$H \to SHS^T. \tag{41}$$

Except in specific cases, there is no obvious advantageous way of using the above transformations when we consider only a *single* member of V. However, the situation is quite different for a *pair* of members.

To deal with that question, we first introduce the notation $V(\mu)$, where $\mu = \pm$, *for that subset of* V *for which*

$$\mathrm{sgn}(-\det h) = \mu. \tag{42}$$

For given μ, let V_4 and V_4' be any two members of $V(\mu)$; primes will be used to distinguish fields in V_4' from fields in V_4. We shall now consider a theorem which is widely known, though we have no reference for it. This theorem is important, because *the transformations of* V *onto* V *which are induced by* K *leave* ρ *and* z *invariant. The theorem shows that this invariance involves no loss in generality in the sense that the invariance does not contradict the conjecture that all of* $V(\mu)$ *can be generated from one of its members by applying* K.

Theorem: For any given V_4 and V_4' in $V(\mu)$, there exist a common coordinate system $\mathbf{x} = (x^1, x^2)$ and choices of the additive constants in $z' - z$ and $\ln\rho' - \ln\rho$ such that, for all $\mathbf{x}$,

$$z'(\mathbf{x}) = z(\mathbf{x}), \quad \rho'(\mathbf{x}) = \rho(\mathbf{x}),$$

and such that the line elements both have the canonical form (38). (Therefore, the two dimensional duality operator * is the same for both spacetimes.)

To prove the above theorem, we first select $\mathbf{x}$ and $\mathbf{x}'$ so that the line elements have the canonical form (38) and so that the coordinate ranges (connected open sets in the real plane) have a point $\mathbf{x}_0$ in common. We then use the transformations (39) to make

$$y'(\mathbf{x}_0) = y(\mathbf{x}_0),$$

where y and y' denote mappings whose domains are U and U', respectively, and whose values are

$$y(\mathbf{x}) := (z(\mathbf{x}), \pm\rho(\mathbf{x})), \quad y'(\mathbf{x}) := (z'(\mathbf{x}), \pm\rho'(\mathbf{x})).$$

We can restrict the domains U and U' so that the inverse mappings y^{-1} and $(y')^{-1}$ exist and have the same domain. Then

$$\sigma := (y')^{-1} \circ y$$

is a coordinate transformation which maps U onto U'. We use this mapping to express all fields in V_4' as functions over the domain U, e.g.,

$$\bar{y}' := y' \circ \sigma, \quad \bar{H} := H \circ \sigma.$$

Now

$$\bar{y}' = y,$$

which is essentially what we had to prove.

As a postscript to the above proof, we can clearly select $x^1 = z$ and $x^2 = \pm\rho$ as our coordinates for both spacetimes. This choice is not always advisable in specific solutions, but it is sometimes useful for general analysis.

3. THE F POTENTIAL

Recall that the self-duality relation (31) was derived from the statement that

$$W_b = -dx^a dH_{ab} = d(dx^a H_{ab})$$

is self-dual. Inspection will reveal that Eq. (31) is in fact, equivalent to the statement that $dx^a H_{ab}$ is a potential for a closed self-dual 2-form, since all of the steps by which the former was derived from the latter are reversible.

We shall now introduce a one parameter family of 2×2 matrix fields $F(t)$ for which $dx^a F_{ab}(t)$ are also potentials for closed self-dual 2-forms. From the operators which act on dH in the two sides of Eq. (31), we construct

$$\Lambda(t) := 1 - 2t(z \pm \rho^*), \tag{43}$$

$$A(t) := I - t(H + H^\dagger)\Omega, \tag{44}$$

whereupon Eq. (31) is expressible in the form

$$t\,dH = A(t)\Gamma(t), \tag{45}$$

$$\Gamma(t) := t\Lambda(t)^{-1}dH. \tag{46}$$

The inverse of $A(t)$ in the above equation is computed from Eqs. (26), (27), (43), and (5), which yield

$$\tilde{A}(t)A(t) = \lambda(t)^2, \quad \tilde{\Lambda}(t) := 1 - 2t(z \mp \rho^*). \tag{47}$$

The integrability condition for the self-duality relation is obtained by taking the exterior derivative of Eq. (45), with the result

$$dA(t)\Gamma(t) + A(t)d\Gamma(t) = 0.$$

However, from Eqs. (37), (44), and (45),

$$dA(t)\Gamma(t) = -t\,dH\Omega\Gamma(t) = -A(t)\Gamma(t)\Omega\Gamma(t).$$

Therefore, the final form for the integrability condition is

$$d\Gamma(t) = \Gamma(t)\Omega\Gamma(t). \tag{48}$$

However, *this is the well known complete integrability condition for the differential equation*

$$dF(t) = \Gamma(t)\Omega F(t), \tag{49}$$

whose solutions are 2×2 nonsingular matrix fields $F(t)$. Except for some gauge conditions which will be discussed later, that is our definition of the F potential.

To convince ourselves that $dx^a F_{ab}(t)$ is the potential for a closed self-dual 2-form, multiply Eq. (49) on the left by $A(t)$ and use Eq. (45) to get

$$A(t)dF(t) = t\,dH\Omega F(t). \tag{50}$$

Then, from Eqs. (43), (44), (46), (49), and (50),

$$2(z \pm \rho^*)dF(t) = (H + H^\dagger)\,\Omega dF(t). \tag{51}$$

This is exactly the same self-duality relation as the one [Eq. (31)] *which is satisfied by dH.*

For any given point $\mathbf{x}_0$ in the domain of H, the general solution of Eq. (49) is given by

$$F(\mathbf{x},t) = \mathscr{F}(\mathbf{x},\mathbf{x}_0,t)F(\mathbf{x}_0,t), \tag{52}$$

where $F(\mathbf{x}_0,t)$ is the value of $F(t)$ at $\mathbf{x} = \mathbf{x}_0$, and $\mathscr{F}(\mathbf{x},\mathbf{x}_0,t)$ is that particular solution of Eq. (49) which satisfies

$$\mathscr{F}(\mathbf{x}_0,\mathbf{x}_0,t) = I, \tag{53}$$

for all t. We shall discuss $F(\mathbf{x}_0, t)$ later, and $\mathscr{F}(\mathbf{x},\mathbf{x}_0,t)$ now. As usual, we shall let $\mathscr{F}(t)$ denote that field whose values are $\mathscr{F}(\mathbf{x},\mathbf{x}_0,t)$.

It is important that we have a clear understanding of the domain and singularities of $\mathscr{F}(t)$. As regards the domain, we shall begin by considering only those charts such that H is a holomorphic function of the coordinates $\mathbf{x} = (x^1, x^2)$ in the range U of each chart. Here U is a *region* (*a connected*[26] *open set*) in the real plane. To say that H is *holomorphic* in U means that H has an extension $[H]$ to a region $[U]$ in

$$C^2 := C \times C \quad (C := \text{complex plane})$$

such that $[H]$ is *holomorphic* in $[U]$, i.e., $[H]$ is single valued and has a Taylor series expansion about every point $\mathbf{x} = (x^1, x^2)$ of $[U]$ such that the series converges to the function in at least one nonempty interval whose center is $\mathbf{x}$.[27] An *interval* in C^n is the Cartesian product of any n open circular disks in C; these intervals are used in the conventional way to define the open sets in C^n.

We are interested in the dependence of $\mathscr{F}(\mathbf{x},\mathbf{x}_0,t)$ on $\mathbf{x}_0$ and t as well as on $\mathbf{x}$. So we consider $[U]\times[U]\times C$, which is a region in C^5. Some of the points in $[U]\times[U]\times C$ are singularities of $\mathscr{F}(t)$. This can be seen by inspecting the self-dual components of Eq. (49), which are given by [see Eqs. (33)]

$$\frac{\partial F(t)}{\partial x_A} = -\frac{1}{2}(r_A - \tau)^{-1}\left(\frac{\partial H}{\partial x_A}\Omega\right)F(t), \quad \tau := (2t)^{-1}. \tag{54}$$

It is clear that the matrix coefficient on the right side of the above equation has singularities only at $(\mathbf{x}, t)$ such that

$$(r_+ - \tau)(r_- - \tau) = [\tau\lambda(t)]^2 = 0,$$

$$r_\pm := r_{\pm i} \text{ or } r_{\pm 1}.$$

Therefore, *we expect all of the singularities of* $\mathscr{F}(t)$ *to be contained in the set of all* $(\mathbf{x},\mathbf{x}_0,t)$ *such that* (r_{0A} is the value of r_A at $\mathbf{x} = \mathbf{x}_0$)

$$(r_+ - \tau)(r_- - \tau)(r_{0+} - \tau)(r_{0-} - \tau) = 0.$$

Let D^5 *denote the set of all* $(\mathbf{x},\mathbf{x}_0,t)$ *in* $[U]\times[U]\times C$ *such that*

$$(r_+ - \tau)(r_- - \tau)(r_{0+} - \tau)(r_{0-} - \tau) \neq 0.$$

Note that $(\mathbf{x}, \mathbf{y}, t)$ is a member of $D^{(5)}$ if and only if both $(\mathbf{x},\mathbf{x},t)$ and $(\mathbf{y}, \mathbf{y}, t)$ are members of $D^{(5)}$.

It is $D^{(5)}$ which is the intended domain[27] of $\mathscr{F}(t)$, and we shall soon state a theorem concerning the existence of a solution of Eq. (49) over this domain. However, the high dimension of $D^{(5)}$ makes it hard to grasp relations in that space. Therefore we introduce the auxiliary spaces

$D^{(3)}$: = set of all $(\mathbf{x}, t)$ such that $\mathbf{x}$ is in $[U]$, t is in C, and $(r_+ - \tau)(r_- - \tau) = [\tau\lambda(\mathbf{x}, t)]^2 \neq 0$;

$D_t^{(2)}$: = set of all $\mathbf{x}$ such that $(\mathbf{x}, t)$ is in $D^{(3)}$ (t = any point in C);

$D_{\mathbf{x}}^{(1)}$: = set of all t such that $(\mathbf{x}, t)$ is in $D^{(3)}$ ($\mathbf{x}$ = any point in $[U]$).

Thus, $D_t^{(2)}$ and $D_{\mathbf{x}}^{(1)}$ are subsets of $[U]$ and C, respectively. Obvserve that $D_{\mathbf{x}}^{(1)}$ is simply C minus the pair of complex numbers

$$t = (2r_\pm)^{-1}.$$

$D_t^{(2)}$ is $[U]$ minus the pair of surfaces whose equations are

$$r_\pm = \tau.$$

We may regard C^2 as a four dimensional real manifold with the real and imaginary parts of x^1, x^2 serving as coordinates; then the excluded surfaces are two dimensional. The following additional statements can be verified with ease:

(1) $D_{\mathbf{x}}^{(1)}$, $D_t^{(2)}$, $D^{(3)}$ and $D^{(5)}$ are, respectively, regions in C, C^2, C^3 and C^5.

(2) $D_t^{(2)}$ is the set of all $\mathbf{x}$ in C^2 such that $(\mathbf{x}, \mathbf{x}, t)$ is in $D^{(5)}$.

(3) $D^{(5)}$ is the set of all $(\mathbf{x}, \mathbf{y}, t)$ in C^5 such that $\mathbf{x}$ and $\mathbf{y}$ are both in D_t.

A disadvantage of the sets $D_{\mathbf{x}}^{(1)}$ and $D_t^{(2)}$ is that they are multiply connected. We need simply connected versions of these domains for our applications. In principle, the simply connected domains may be constructed by first selecting an appropriate (five dimensional) hypersurface $K^{(3)}$ in the six dimensional real manifold $D^{(3)}$; then we may define

$$S^{(3)}:=D^{(3)}-K^{(3)},$$

$$S_t^{(2)}:=\text{set of all } \mathbf{x} \text{ in } D_t^{(2)} \text{ such that } (\mathbf{x},t) \text{ is in } S^{(3)},$$

$$S_{\mathbf{x}}^{(1)}:=\text{set of all } t \text{ in } D_{\mathbf{x}}^{(1)} \text{ such that } (\mathbf{x},t) \text{ is in } S^{(3)}.$$

The idea is to select $K^{(3)}$ so that the following conditions are fulfilled:

(1) $S_t^{(2)}$ is a simply connected region in C^2.

(2) $S_{\mathbf{x}}^{(1)}$ is a simple connected region in C (with the Riemann sphere topology). Specifically,

$$K_{\mathbf{x}}:=D_{\mathbf{x}}^{(1)}-S_{\mathbf{x}}^{(1)}$$

is an arc which joins $(2r_+)^{-1}$ to $(2r_-)^{-1}$.

Actually, it is easier to select the complex plane cuts $K_{\mathbf{x}}$ first, and then to find $K^{(3)}$ from the condition

$$K^{(3)}=\text{set of all } (\mathbf{x},t) \text{ in } D^{(3)} \text{ such that } t \text{ is in } K_{\mathbf{x}}.$$

In practice, almost all applications which we have in mind involving only a limited set of values of $\mathbf{x}$ and do not require any global knowledge of a specific $K^{(3)}$. The concept is required mostly for the development of the general theory.

Now, let $S^{(5)}$ denote the set of all $(\mathbf{x},\mathbf{y},t)$ in C^5 such that both $\mathbf{x}$ and $\mathbf{y}$ are in $S_t^{(2)}$. In the Appendix, we prove the following basic theorem[28] which holds for any choice of the simply connected subregions:

Theorem: (1) There exists exactly one (multiple valued) solution $\mathscr{F}(t)$ of Eq. (49) such that its domain is $D^{(5)}$ and such that the restriction of $\mathscr{F}(t)$ to any given $S^{(5)}$ has a unique branch which satisfies

$$\mathscr{F}(\mathbf{x},\mathbf{x},t)=I, \tag{55}$$

whenever $(\mathbf{x},\mathbf{x},t)$ is in $S^{(5)}$. (2) For any given $\mathbf{x},\mathbf{y},\mathbf{z}$ in $D^{(2)}$ and for any given value of $\mathscr{F}(\mathbf{x},\mathbf{y},t)$, there exist values of $\mathscr{F}(\mathbf{x},\mathbf{z},t)$, $\mathscr{F}(\mathbf{z},\mathbf{y},t)$, and $\mathscr{F}(\mathbf{y},\mathbf{x},t)$ such that

$$\mathscr{F}(\mathbf{x},\mathbf{y},t)=\mathscr{F}(\mathbf{x},\mathbf{z},t)\mathscr{F}(\mathbf{z},\mathbf{y},t), \tag{56}$$

$$[\mathscr{F}(\mathbf{x},\mathbf{y},t)]^{-1}=\mathscr{F}(\mathbf{y},\mathbf{x},t). \tag{57}$$

The branch of $\mathscr{F}(t)$ for which Eq. (55) holds satisfies Eqs. (56) and (57) for all $\mathbf{x},\mathbf{y},\mathbf{z}$ in $S_t^{(2)}$. (3) $\mathscr{F}(t)$ is analytic in $D^{(5)}$.

The key point in the above theorem is *the absence of any singularities in* $\mathscr{F}(\mathbf{x},\mathbf{x}_0,t)$ *over the domain* $D^{(5)}$, from which we have excluded (by definition) the zeros of

$$\tau\lambda(\mathbf{x},t)\cdot\tau\lambda(\mathbf{x}_0,t).$$

Now, we investigate the character of the singularities at these zeros.

Upon restricting U, we can select $x^1=z, x^2=\pm\rho$, and this is what we choose to do. In the actual analysis of the singularities, we use null coordinates

$$x_A=r_A,$$

as defined by Eq. (32). We shall let $A=\pm$ to cover both cases $A=\pm i$ and $A=\pm 1$.

Note tht Eq. (54) is, in effect, a linear homogeneous ordinary differential equation, with r_{-A}, r_{0+}, r_{0-}, and t playing the roles of parameters. $r_A=\tau$ is a regular singular point of this equation.[29] Inspection of Eq. (35) shows that *the eigenvalues of the coefficient matrix*

$$-\tfrac{1}{2}\frac{\partial H}{\partial r_A}\Omega,$$

on the right side of Eq. (54) *are*

$$p=0,\ -\tfrac{1}{2}.$$

Therefore, if $(\mathbf{x},\mathbf{x}_0)$ is in $[U]\times[U]$, and $(r_{-A}-\tau)(r_{0+}-\tau)(r_0-\tau)\neq 0$, we have

$$\mathscr{F}(\mathbf{x},\mathbf{x}_0,t)=\mathscr{F}_0(\mathbf{x},\mathbf{x}_0,t)+(r_A-\tau)^{-1/2}\mathscr{F}_1(\mathbf{x},\mathbf{x}_0,t), \tag{58}$$

where $\mathscr{F}_0$ and $\mathscr{F}_1$ are regular functions of r_A in a neighborhood of $r_A=\tau$. In view of the analyticity of $\mathscr{F}$ as a function of $(\mathbf{x},\mathbf{x}_0,t)$ in $D^{(5)}$, we can also say that for fixed $(\mathbf{x},\mathbf{x}_0)$ in $[U]\times[U]$, $\mathscr{F}_0(t)$ and $\mathscr{F}_1(t)$ are regular functions of t in a neighborhood of $\tau=r_A$.

In summary, there are branch points at the zeros of $\tau\lambda(\mathbf{x},t)$ *and* $\tau\lambda(\mathbf{x}_0,t)$, *and there are no other singularities over the domain* $[U]\times[U]\times C$. In view of Eqs. (55), (57), and (58), the branch points at $r_A=\tau$ will be of index $-\frac{1}{2}$, while those at $r_{0A}=\tau$ will be of index $\frac{1}{2}$. We have verified that this is the case for the Zipoy–Voorhees and the Kerr–NUT $\mathscr{F}(t)$.[30]

We now return to the general soultion of Eq. (49) as given by Eq. (52), and we discuss the problem of selecting a "suitable $F(\mathbf{x}_0,t)$" with the aid of criteria which are independent of the particular member of V under discussion and which are independent of the choice of $\mathbf{x}_0$. Such criteria are to be found in the relations

$$dF(0)=d[\dot F(0)-H]=0, \tag{59}$$

$$d[\lambda(t)\det F(t)]=0, \tag{60}$$

$$d[F(t)^\dagger\Omega A(t)F(t)]=0, \tag{61}$$

where $\dot F(t)=\partial F(t)/\partial t$ and where it is to be understood that

$$F(\mathbf{x},\mathbf{x}_0,t)^\dagger:=\text{h.c. of } F(\mathbf{x}^*,\mathbf{x}_0^*,t^*).$$

Equations (59) follow from Eq. (49) and the definitions (43) and (44) of $\Lambda(t)$ and $A(t)$. Equation (60) is derived from Eq. (49) by using the relations

$$F(t)^T\Omega F(t)=\Omega[\det F(t)], \tag{62}$$

$$2tA(t)^{-1}dz=-\lambda(t)^{-1}d\lambda(t), \tag{63}$$

and Eq. (29). [Note that Eq. (62) holds if $F(t)$ is replaced by *any* 2×2 matrix.] Equation (61) is deduced from Eq. (50) and the facts that Ω and $\Omega A(t)$ are Hermitian.

We shall select $F(\mathbf{x}_0,t)$ so that the integrals of Eqs. (59) to (61) are simply

$$F(0)=\Omega,\quad \dot F(0)=H, \tag{64}$$

$$\lambda(t)\det F(t)=-1, \tag{65}$$

$$F(t)^\dagger\Omega A(t)F(t)=\Omega. \tag{66}$$

Observe that Eqs. (65) and (66) are consistent with Eqs. (64) and the facts that

$$\Omega^2=I,\quad A(0)=I,\quad \lambda(0):=1.$$

Equations (64)–(66) are also consistent with the gauge conditions of K–C.[6,7]

We shall now prove that there exists at least one choice of $F(\mathbf{x}_0, t)$ and of the additive imaginary symmetric constant in $H - H^T$ such that Eqs. (64)–(66) are true and such that $F(\mathbf{x}_0, t)$ has no singularities except for the branch points of $\lambda(\mathbf{x}_0, t)$. We start by selecting the additive constant in $H - H^T$ so that $H(\mathbf{x}_0)$ is Hermitian at a particular point $\mathbf{x}_0$ which will remain fixed throughout the remainder of this proof [see Eq. (40)]. Then, we let $A(t)^{-1/2}$ denote that branch of the square root of $A(t)$ which is holomorphic at $t = 0$ and has the value $A(0)^{-1/2} = I$. To compute this square root, we use Eqs. (5), (19), (28), (29), and (62) and the fact that $\mathrm{tr}(h\epsilon) = 0$ to show that

$$\mathrm{tr}A(t) = 2(1 - 2zt), \quad \det A(t) = \lambda(t)^2. \tag{67}$$

This enables us to construct the minimal polynomial of $A(t)$, which gives us

$$A(t)^{-1/2} = \frac{[1 - 4tz + \lambda(t)]I + t(H + H^\dagger)\Omega}{\lambda(t)[2(1 - 2tz + \lambda(t))]^{1/2}}. \tag{68}$$

We claim that a choice which satisfies our criteria is the following assignment at any *one* point $\mathbf{x}_0$ such that $\rho_0 \neq 0$ and $z_0^2 \pm \rho_0^2 \neq 0$:

$$F(\mathbf{x}_0, t) = A(\mathbf{x}_0, t)^{-1/2}\Omega w(\mathbf{x}_0, t), \quad \rho \neq 0, \quad z_0^2 \pm \rho_0^2 \neq 0, \quad \lambda(\mathbf{x}_0, 0) = 1 \tag{69}$$

(does not supply $F(t)$ at $\mathbf{x} \neq \mathbf{x}_0$),

$$w(t) := \begin{pmatrix} 2^{1/2}[1 - 2tz + \lambda(t)]^{-1/2} & 0 \\ 0 & 2^{-1/2}[1 - 2tz + \lambda(t)]^{1/2} \end{pmatrix}.$$

That Eqs. (64) to (66) are satisfied with this choice can be proven from the relations

$$\dot\lambda(0) = -2z,$$

$$\frac{\partial}{\partial t}[A(t)^{-1/2}] = \tfrac{1}{2}(H + H^\dagger)\Omega, \quad \text{when } t = 0,$$

$$\Omega[A(t)^{-1/2}]^\dagger = A(t)^{-1/2}\Omega,$$

$$w(t)^\dagger \Omega w(t) = \Omega, \quad \det w(t) = 1,$$

$$w(0) = I, \ \dot w(0) = 0,$$

which can be derived from Eqs. (5), (69), and (44) or (68). *As regards singularities in the t-plane,* observe that $1 - 2tz + \lambda(t)$ can vanish only if $\rho = 0$ and $z = \tau$, and observe that $\tau\lambda(t) \to (z^2 \pm \rho^2)^{1/2}$ as $t \to \infty$. Therefore, it can be seen that our choice for $F(\mathbf{x}_0, t)$ *has no singularities except for a branch point of index* $-(1/2)$ *at the zeros of* $\lambda(\mathbf{x}_0, t)$ and that

$$F(\mathbf{x}_0, t)\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}$$

is *holomorphic in a neighborhood of* $t = \infty$ *and has an inverse there.*

We now summarize the key points which have been made so far concerning $F(t)$.

(1) $F(t)$ is defined, for given H, as any solution of Eq. (49) which is (for given $\mathbf{x}$) holomorphic in a neighborhood of $t = 0$ and which satisfies the gauge conditions (64) to (66).

(2) We can further specialize the gauge of $F(t)$ so that its only singularities are at the zeros of $\tau\lambda(\mathbf{x}, t)$ and, possibly, at the zeros of $\tau\lambda(\mathbf{x}_0, t)$ for one other point $\mathbf{x}_0$ in the domain $[U]$ of H. These singularities are branch points of index $-1/2$ (except when there are confluences).

(3) In this specialization of the gauge,

$$F(t)\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix},$$

is, with the exception of the points at which $z^2 \pm \rho^2 = 0$,[22] holomorphic in a neighborhood of $t = \infty$ and has an inverse there.

Actually, the gauge conditions (2) *and* (3) *given above are independent of one another if one restricts* (2) *to finite t.* All of the F potentials used so far by the authors and by K–C satisfy condition (3). However, they do not all satisfy condition (2). For example, the Zipoy–Voorhees $F(t)$ used both by the authors[13] and by K–C[7] in the past has poles or branch points at $t = \pm\frac{1}{2}$ in addition to those at the zeros of $\tau\lambda(\mathbf{x}, t)$. (The $\mathbf{x}_0$ in this case is, in effect, "at infinity," which is outside of the domain of H and which will be covered in a sequel to this paper.) The singularities at $t = \pm\frac{1}{2}$ can be removed by multiplying the given $F^{zv}(t)$ with

$$\begin{pmatrix} \left(\dfrac{1-2t}{1+2t}\right)^{\delta/2} & 0 \\ 0 & \left(\dfrac{1+2t}{1-2t}\right)^{\delta/2} \end{pmatrix} \quad (\delta = \text{real parameter}),$$

on the right, where we use a cut whose section in the t-plane intersects the real axis in the open interval between $t = \frac{1}{2}$ and $t = 0$ or between $t = -\frac{1}{2}$ and $t = 0$.

It is essential to point out that we are not advocating the use of our specialized gauge under all circumstances or by everyone. *However, it is important to know that it exists.* If we are given $F(\mathbf{x}, t)$ in a different gauge, then the $F(\mathbf{x}, t)$ in the specialized (sp) gauge can be constructed from

$$F_{\mathrm{sp}}(\mathbf{x}, t) = \mathscr{F}(\mathbf{x}, \mathbf{x}_0, t)F_{\mathrm{sp}}(\mathbf{x}_0, t),$$

$$\mathscr{F}(\mathbf{x}, \mathbf{x}_0, t) := F(\mathbf{x}, t)F(\mathbf{x}_0, t)^{-1},$$

where $F_{\mathrm{sp}}(\mathbf{x}_0, t)$ can be chosen to be the expression (69).

4. THE HOMOGENEOUS HILBERT PROBLEM

In this section, $F_0(t)$ will denote the F potential of any given member V_{40} of $V(u)$. *On the other hand, $F(t)$ will denote the result of a specific transformation of $F_0(t)$ which is induced by a member $u(t)$ of K_L, and $F(t)$ is not to be construed as being the F potential of any member of $V(\mu)$ until we have proven that such is the case.*

We select an arbitrary compact region $[U_c]$ in the domain $[U]$ of

$$H_0 = \dot F_0(0),$$

(where we are letting $H_0 := [H_0]$ to avoid the cumbersome brackets). It will be understood that $\mathbf{x}$ is restricted to $[U_c]$ throughout the remainder of this section.

Then there exists at least one smooth contour L about the origin such that $F_0(t)$ is holomorphic on $L + L_+$. Choose any L is this category, and let $u(t)$ be any member of K_L.

Suppose $u(t)$ is holomorphic on L_+. If $F_0(t)$ is subject to the transformation

$$F_0(t) \to F(t) := [u(0)^{-1}]^T F_0(t) u(t)^{-1},$$

then it can be seen from Eqs. (3), (49), and (64)–(66) that $F(t)$ satisfies all of the defining conditions for an F potential of V_{40}. The only possible point of difficulty in seeing this arises from the fact that

$$H_0 \to \dot{F}(0) := [u(0)^{-1}]^T H_0 u(0)^{-1} - \Omega \dot{u}(0) u(0)^{-1}.$$

However, it is easily proven that $\Omega \dot{u}(0) u(0)^{-1}$ is imaginary and symmetric; so the above is a trivial gauge transformation. [See Eqs. (40) and (41)].

If $u(t)$ is not holomorphic in L_+, then the above transformation of $F_0(t)$ by $u(t)$ may no longer define a simple gauge transformation of the F potential. However, the theory[1] of the HHP informs us that there exist $F(t)$ and $X_-(t)$ such that

$$F(s) = X_-(s) F_0(s) u(s)^{-1}, \quad F(0) := F_0(0),$$

for all s in L, where the following conditions hold:

(1) $F(t)$ is continuous and has an inverse for all t in $L + L_+$, and $F(t)$ is holomorphic in L_+.

(2) For all finite t in $L + L_-$, $X_-(t)$ is continuous and has an inverse. For all finite t in L_-, $X_-(t)$ is holomorphic.

(3) The rows of $X_-(t)$ have finite degrees $-m$ and $-n$ in a neighborhood of $t = \infty$ such that

$$M_- := \lim_{t \to \infty} Z(t) X_-(t)$$

has an inverse, where

$$Z(t) := \begin{pmatrix} t^m & 0 \\ 0 & t^n \end{pmatrix}.$$

It is clear that this HHP defines a generalization of the previously considered transformation when $u(t)$ was holomorphic on $L + L_+$, and $X_-(s)$ was $[u(0)^{-1}]^T$. We still have to justify the use of the notation $F(t)$ in this generalization.

We shall prefer to put the above HHP equations in the more standard form of Eqs. (9)–(11), viz.,

$$X_-(s) = X_+(s) G(s), \quad X_+(0) = I,$$

$$G(t) := F_0(t) u(t) F_0(t)^{-1}, \quad X_+(t) := F(t) F_0(t)^{-1}.$$

Since, in our particular HHP, $G(t)$ satisfies the strong condition that it is holomorphic on L (as opposed to just satisfying a Hölder condition on L), *there is the correspondingly stronger conclusion that $X_+(t)$ and $X_-(t)$ are each holomorphic on L.* We introduce, as is customary, the *sectionally* holomorphic function

$$X(t) := X_+(t), \quad \text{if } t \text{ is in } L + L_+,$$

$$X(t) := X_-(t), \quad \text{if } t \text{ is in } L + L_-.$$

Then $X(t)$ [or the ordered pair of functions $X_\pm(t)$, if we prefer] is called a *fundamental solution*[1] of the HHP for $G(s)$.

The integers n, m, whose negatives are the degrees at infinity of the rows of $X(t)$, are called the component indices of the HHP for $G(s)$. Since, in our particular HHP,

$$\det G(s) = 1,$$

there is a theorem which tells us that

$$m + n = 0.$$

In our proof that the solution $F(t)$ of the HHP is an F potential of some member V_4 of $V(\mu)$, we shall have to rely on the premise $m = 0$. The question of whether the component indices are 0, 0 is important for us since, otherwise, we would have to place constraints of an (as yet) unknown character on our choice of the members $u(t)$ of K_L. *We shall assume as a working hypothesis that $m = 0$.* In the discussion of Sec. 6, we shall reduce the problem of determining the component indices for $u(t)$ and $F_0(t)$ to that of proving the existence of a solution of our HHP for arbitrary $F_0(t)$ and for an especially simple kind of $u(t)$, viz.,

$$u_1(t) := \begin{pmatrix} 1 & t\alpha(t) \\ 0 & 1 \end{pmatrix},$$

where $\alpha(t)$ is holomorphic on L and at $t = \infty$. [The solution of the HHP corresponding to arbitrary $F_0(t)$ and to any $u_1(t)$ for which $\alpha(t)$ has n simple poles in L_+ has already been found; so the question is answered insofar as that case is concerned.]

Like remarks apply to a *second working hypothesis* which we shall make, viz., that $dX_\pm(t)$ have the same domains of continuity and holomorphy as (respectively, for $\pm$) $X_\pm(t)$.

We shall now take up the main goal of this section. We first prove that $X_+(t)$ is holomorphic at $t = \infty$ and that it has an inverse at that point. This follows from the extension

$$X_-(t) = X_+(t) G(t)$$

of the HHP into the complex plane. *From the gauge condition for $F_0(t)$ at $t = \infty$ and from the analogous condition satisfied by $u(t)$ at $t = \infty$* [Sec. 1, after Eqs. (3)], *$G(t)$ is holomorphic at $t = \infty$, and $G(\infty)$ has an inverse.* Therefore, since $X_-(t)$ is also holomorphic and has an inverse at $t = \infty$, it follows that

$$X_+(t) \text{ is holomorphic at } t = \infty, \qquad X_+(\infty)^{-1} \text{ exists.} \tag{70}$$

We next define

$$*, \rho, z, \Lambda(t) := \text{same as those defined for } V_{40},$$

$$H := H_0 + \dot{X}_+(0)\Omega, \tag{71}$$

$$A(t) := I - t(H + H^\dagger)\Omega, \tag{72}$$

$$\Gamma(t) := t\Lambda(t)^{-1} dH. \tag{73}$$

It is essential to grasp that the above *are definitions*; we still have to prove that they are what the notations suggest.

We shall next prove that, for all t,

$$\det X(t) = 1, \tag{74}$$

$$X(t)^\dagger \Omega A(t) X(t) = \Omega A_0(t), \tag{75}$$

$$\det A(t) = \det A_0(t) = \lambda(t)^2, \tag{76}$$

$$dX(t) = \Gamma(t)\Omega X(t) - X(t)\Gamma_0(t)\Omega, \tag{77}$$

$$\Omega A(t)dX(t) + t[X(t)^\dagger]^{-1}\Omega dH_0\Omega = t\Omega dH\Omega X. \tag{78}$$

To prove Eq. (74), use Eqs. (3), (9), (10), and (for F_0) (65) to prove

$$\det X_+(s) = \det X_-(s).$$

Therefore, $\det X(t)$ is an entire function of t. Since $X(t)$ is regular at $t = \infty$,

$$\det X(t) = \text{constant}.$$

The value of this constant is unity, from Eq. (11).

To prove Eq. (75), use Eqs. (3), (9), (10), and (for F_0, A_0) (66) to get

$$[X_+(s)^\dagger]^{-1}\Omega A_0(s)[X_+(a)]^{-1} = \text{same with } - \text{ replacing } +.$$

Therefore,

$$[X(t)^\dagger]^{-1}\Omega A_0(t)[X(t)]^{-1},$$

is an entire function of t. Since $A_0(t)$ is linear in t, so is this entire function. By using Eqs. (11), (71), and (for H_0) (44), we obtain the coefficients in this linear function of t, which turns out to be $\Omega A(t)$ as defined by Eq. (72). That proves Eq. (75).

Equation (76) follows simply from Eqs. (74), (75), and (for A_0) (67).

To prove Eq. (77), operate on Eq. (9) with $A(s)d$, and use Eqs. (10) and (for F_0) (49) to get

$$A(s)dX_+(s)X_+(s)^{-1} + sX_+(s)dH_0\Omega X_+(s)^{-1}$$
$$= \text{same with } - \text{ replacing } +. \tag{79}$$

Therefore,

$$A(t)dX(t)X(t)^{-1} + tX(t)dH_0\Omega X(t)^{-1}$$

is a linear function of t. [Note Eq. (74)]. By manipulations similar to the proof of Eq. (75) and by using Eqs. (11), (71), (43), and (73), we get Eq. (77).

To prove Eq. (78), first multiply Eq. (79) through by $A(s)^{-1}$ and introduce $\Gamma_0(s)$ as defined (for H_0) by Eq. (46). Then use Eqs. (72), (75), and (for H_0) (45) to get

$$\Omega A(s)dX_+(s)X_+(s)^{-1} + s[X_+(s)^\dagger]^{-1}\Omega dH_0\Omega[X_+(s)]^{-1}$$
$$= \text{same with } - \text{ replacing } +.$$

The rest of the proof of Eq. (78) uses Eq. (71) and closely resembles the proof of Eq. (77).

We consider by defining $F(t)$ as in Eq. (10), whereupon we deduce

$$dF(t) = \Gamma(t)\Omega F(t), \tag{80a}$$

$$A(t)dF(t) = t\,dH\Omega F(t), \tag{80b}$$

$$F(0) = \Omega, \quad \dot{F}(0) = H, \tag{80c,d}$$

$$-\lambda(t)\det F(t) = 1, \tag{80e}$$

$$F(t)^\dagger\Omega A(t)F(t) = \Omega. \tag{80f}$$

The above are, respectivley, derived from Eqs. (49) and (77), Eqs. (50) and (78), Eqs. (64) and (11), Eqs. (64) and (71), Eqs. (65) and (74), and Eqs. (66) and (75). [In each pair of this list of equations, the first one is for H_0 or F_0, i.e., the subscript "0" is to be inserted at the appropriate places in Eqs. (49), (50), (64), (65), and (66).]

From Eqs. (80a) to (80e), *we see that $F(t)$ satisfies all of the defining equations for the F potential corresponding to a given H. Moreover, from* Eq. (70), *we see that $F(t)$ satisfies our gauge condition at $t = \infty$*, i.e., this condition is preserved by the transformations induced by K_L.

It remains to prove that H is an H potential for some spacetime. To show this, we now prove that

$$2(z \pm \rho^*)dH = (H + H^\dagger)\Omega dH, \tag{81a}$$

$$\tfrac{1}{2}(H - H^T) = i\epsilon z, \tag{81b}$$

$$h := -\operatorname{Re}H \text{ is symmetric}, \tag{81c}$$

$$\rho^2 := \mp \det h. \tag{81d}$$

To derive Eq. (81a), note that Eqs. (73), (80a), and (80b) imply

$$A(t)dF(t) = A(t)dF(t).$$

So, from Eqs. (72) and (43)

$$2(z \pm \rho^*)dF(t) = (H + H^\dagger)\Omega dF(t).$$

Take the t-derivative of the above equation; then set $t = 0$, and use Eq. (80d) to get Eq. (81a).

To derive Eq. (81b), note that Eq. (80e) is equivalent to

$$-\lambda(t)F(t)^T\epsilon F(t) = \epsilon.$$

Take the t-derivative of the above, and set $t = 0$ to get Eq. (81b).

Equation (81c) is a trivial implication of Eq. (81b). Equation (81d) is deduced from Eq. (76) by using the relations

$$A(t)^T\epsilon A(t) = \epsilon \det A(t),$$

and Eq. (72), as well as the relation

$$H + H^\dagger = -2h + 2i\epsilon z,$$

which follows from the definition of h in Eq. (81c) and from Eqs. (81b) and (81c).

Equations (81a) to (81d) show that H fulfills the definition of an H potential for a $V(\mu)$ spacetime V_4 whose metric components g_{ab} are

$$g_{ab} := h_{ab}.$$

Equation (81a) is equivalent to the statement that $dx^a dH_{ab}$ is a closed self-dual 2-form, which is equivalent to the statement that g_{ab} satisfies the vacuum field equations. We refer the reader to Sec. 2 for other details.

5. TWO SIMPLE HHP APPLICATIONS

We shall consider two transformations involving a Minkowski space $F_0(t)$ corresponding to a *rotation about an axis and a time translation* as the pair of isometries characterizing the metrical form in Eq. (1). This $F_0(t)$ has been computed by K–C[6] and is given by[13]

$$F_0(t) = \begin{pmatrix} -\dfrac{\lambda - 1 + 2tz}{2t\lambda} & i\left(\dfrac{\lambda + 1 - 2tz}{2\lambda}\right) \\ -\dfrac{i}{\lambda} & \dfrac{t}{\lambda} \end{pmatrix}, \tag{82}$$

where $\lambda = \lambda(t)$, and z and ρ are the conventional cylindrical coordinates. For given z and ρ, the branch points are at

$$t = (2r_{\pm})^{-1} = \tfrac{1}{2}(z \pm i\rho)^{-1} \quad (\rho > 0).$$

We select a branch cut and a contour L so that $F_0(t)$ is holomorphic on $L + L_+$. For example, if z and ρ are real, we can choose a circular arc with center at $t = 0$ as our cut, and we can choose a circle with center at $t = 0$ and radius $< |2r_{\pm}|^{-1}$ as our L. Any other choices which satisfy the requirement that $F_0(t)$ be holomorphic on $L + L_+$ will do as well and may even be useful at times.

Consider any $u(t)$ which is expressible in the exponential form

$$u(t) = \exp[\gamma(t)\epsilon],$$

$$\gamma(t) = \begin{pmatrix} t\alpha(t) & \xi(t) \\ \xi(t) & t^{-1}\beta(t) \end{pmatrix}, \tag{83}$$

where $\alpha(t)$, $\beta(t)$, and $\xi(t)$ are holomorphic on L and at $t = \infty$. Though we shall not consider the transformations of $F_0(t)$ corresponding to arbitrary α, β, and γ, it is expedient not to specialize until later. The kernel $G(s)$ of the HHP is given by

$$G(s) = \exp M(s), \quad M(s) := F_0(s)\gamma(s)\epsilon F_0(s)^{-1}. \tag{84}$$

From Eqs. (82) and (83), we readily compute (supressing s in some places)

$$M(s) = \lambda^{-1} \begin{pmatrix} -\frac{i}{2}(\alpha - \beta + 2i\xi)(1 - 2sz) & (-\alpha + \beta + 2i\xi)\rho^2 s \\ & + (-\alpha + \beta)(2s)^{-1}(1 - 2sz)^2 \\ -(\alpha - \beta + 2i\xi)s & \frac{i}{2}(\alpha - \beta + 2i\xi)(1 - 2sz) \end{pmatrix} + \begin{pmatrix} \frac{i}{2}(\alpha + \beta) & (\alpha + \beta)(2s)^{-1}(1 - 2sz) \\ 0 & \frac{i}{2}(\alpha + \beta) \end{pmatrix}. \tag{85}$$

Now we are ready to discuss the special cases.

First consider the *B group* of K–C[7] which is defined by

$$\alpha = \beta \text{ is holomorphic in } L_-, \qquad \xi = 0 \quad (B \text{ group}). \tag{86}$$

Upon inserting (86) into (85), *we obtain an* $M(s)$ *which is holomorphic in* L_- (including ∞). Hence, from Eq. (84), the solution of Eqs. (9) to (11) which satisfies the condition at $t = \infty$ is given by

$$X_+(s) = I, \quad X_-(s) = \exp M(s).$$

So, $F(s) = F_0(s)$; *we have proven that the B group leaves the Minkowski space* $F_0(t)$ *of* Eq. (82) *invariant.*

As our second example, consider

$$\alpha = \beta = 0. \tag{87}$$

Equation (85) becomes, as can be seen with the aid of Eqs. (83) and (84):

$$M(s) = \lambda(s)^{-1}\xi(s)P(s) = \xi(s)F_0(s)\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}F_0(s)^{-1}, \tag{88}$$

where

$$P(s) := \begin{pmatrix} 1 - 2sz & 2is\rho^2 \\ -2is & 2sz - 1 \end{pmatrix}.$$

Observe that $P(s)$ is holomorphic in L_+ and at finite values of s in L_-. This suggests that we seek a solution of the form

$$X_+(s) = e^{-\phi_+(0)P(0)}e^{\phi_+(s)P(s)}, \tag{89}$$

where $\phi_+(t)$ is a complex valued function of ρ, z, and t which is (for fixed ρ, z) holomorphic in L_+. Note that the expression (89) has been constructed so that it automatically satisfies Eq. (11).

Upon inserting (89) into Eq. (9) and using Eqs. (84) and (88), we get

$$X_-(s) = e^{-\phi_+(0)P(0)}e^{[\phi_+(s) + \lambda(s)^{-1}\xi(s)]P(s)}, \tag{90}$$

So we see that we have the solution of the HHP which satisfies the boundary conditions at $t = 0$ and at $t = \infty$ if we find $\phi_+(t)$ and $\phi_-(t)$ such that

$$\phi_+(s) - \phi_-(s) = -\lambda(s)^{-1}\xi(s), \tag{91}$$

where $\phi_-(t)$ is continuous on $L + L_-$, is holomorphic in L_-, and satisfies

$$\phi_-(\infty) = 0. \tag{92}$$

Equation (92) is necessary, because $P(t)$ is of degree 1 in t.

Equations (91) and (92) constitute a standard problem in analysis with the solution

$$\phi(t) = -\frac{1}{2\pi i}\int_L ds \frac{\xi(s)}{\lambda(s)(s - t)}, \tag{93}$$

$$\phi(t) = \phi_{\pm}(t) \text{ if } t \text{ is in } L_{\pm} \text{ (respectively)},$$

where it should be remembered that the branch points of $\lambda(t)$ are in L_-. From Eqs. (88) and (89), we have (recalling that $F_0(0) = i\epsilon$)

$$F(s) = \begin{pmatrix} e^{-\psi(0)} & 0 \\ 0 & e^{\psi(0)} \end{pmatrix} F_0(s) \begin{pmatrix} e^{-\psi(s)} & 0 \\ 0 & e^{\psi(s)} \end{pmatrix}, \tag{94}$$

where

$$\psi(t) := \lambda(t)\phi_+(t). \tag{95}$$

In summary, the solution is given by Eqs. (82), (94), (95), and (93).

The expression (94) was originally derived[6] by K–C via a different route and is the $F(t)$ for the general static axially symmetric stationary vacuum spacetime (the Weyl metric). Note that

$$\lambda(s) = 2s[(2s)^{-2} - 2(2s)^{-1}r\cos\theta + r^2]^{1/2},$$

where

$$z = r\cos\theta, \quad \rho = r\sin\theta.$$

Therefore, $\phi(t)$ as given by Eq. (93) is a solution of Laplace's equation (in three dimensional Euclidean space). On account of the reality condition

$$[\lambda(s)^{-1}\xi(s)]^* = \lambda(s^*)^{-1}\xi(s^*),$$

$\psi(0) = \phi_+(0)$ is real; $\psi(0)$ is the general axially symmetric solution of Laplace's equation in a neighborhood of $r = 0$ which is free of singularities.

6. PERSPECTIVES

There remain two central themes concerning the exact solutions for members of V. One of these is the problem of finding the solutions of the HHP or of the equivalent integral equation corresponding to any given $u(t)$ and $F_0(t)$. The second is the question of whether or not all of V can be generated from Minkowski space by the group K.

In a certain sense, the first of these problems can be reduced to a succession of relatively simple ones due to the fact that the general $u(t)$ can be expressed as a product involving factors of the form

$$u^{(1)}(t) = \begin{pmatrix} 1 & t\alpha(t) \\ 0 & 1 \end{pmatrix}, \quad u^{(2)}(t) = \begin{pmatrix} 1 & 0 \\ -t^{-1}\beta(t) & 1 \end{pmatrix},$$

where $\alpha(t)$ and $\beta(t)$ are holomorphic on L and at $t = \infty$.

The point is that the transformations involving a $u^{(1)}(t)$ or a $u^{(2)}(t)$ seem to be amenable to analysis for a *general* $F_0(t)$. For example, as we already mentioned in Sec. 1, the solution, corresponding to an arbitrary $F_0(t)$ and to a $u^{(1)}(t)$ [or $u^{(2)}(t)$] for which $\alpha(t)$ [or $\beta(t)$] has n simple poles in L_+ has been found.[14,20] At least, this is true to the extent that a certain known $n \times n$ matrix which appears in the solution can be inverted; the problem has thus been solved in the sense that it has been reduced to a standard algebraic problem which can be machine computed for reasonably small n.

An important extension of the result for n simple poles is the corresponding "solution" for an $\alpha(t)$ or a $\beta(t)$ which has a smooth (or even more arbitrary) line distribution of singularities in L_+. It is plausible that the same form of solution holds, but the inverse of a known $n \times n$ matrix is replaced by the inverse of a known Fredholm operator. The problem is thus reduced to one about which much has been written.

As a more formal application of the factorization into $u^{(1)}$ and $u^{(2)}$ elements, there is a good chance of proving (or disproving) that the Fredholm operator corresponding to a general $\alpha(t)$ or $\beta(t)$ is invertible for every $F_0(t)$. That would be equivalent to proving that the component indices of the HHP for every $u(t)$ and every $F_0(t)$ are (0, 0). A like remark holds for the question whether $dX_\pm(t)$ has the same domains of continuity and of holomorphy as $X_\pm(t)$.

There may also be the possibility of taking advantage of the special properties of a particular type of $F_0(t)$ to facilitate the inversion of the Fredholm operator. The case of the static Weyl $F_0(t)$ is especially important because there is the distinct possibility based on work of K–C[6,7] and of Hoenselaers, Kinnersley, and Xanthopoulos[15] that every axially symmetric stationary vacuum (at least the asymptotically flat ones) can be generated from the general static Weyl $F_0(t)$ by the $u^{(1)}(t)$ transformations alone.

Finally, there is the second central theme. Can every member of V be generated from Minkowski space by the group K? With the aid of the HHP, this question is reduced to a reasonable one in analysis (or in the topic of linear differential equations), viz., given any $F_0(t)$ and $F(t)$ in the same $V(\mu)$, does a $u(t)$ always exist such that

$$F(t)u(t)F_0(t)^{-1}$$

is holomorphic in a neighborhood of $t = \infty$ which contains the singularities of both $F_0(t)$ and $F(t)$? [In our special gauge, that would simply be the zeros of $\tau\lambda(\mathbf{x}, t)$ and $\tau\lambda(\mathbf{x}_0, t)$.] This is a problem which should interest everyone.

APPENDIX: PROOF OF THEOREM IN SEC. 3

A_1, A_2, and T shall denote any open circular disks in the complex plane C. We let

$$A^{(2)} := A_1 \times A_2, \quad A^{(5)} := A^{(2)} \times A^{(2)} \times T.$$

$A^{(2)}$ will be called an (open) *interval* in C^2, and $A^{(5)}$ will be called a *symmetric interval* in C^5. We shall lead up to our proof of the theorem in Sec. 3 by proving five lemmas. The first lemma is a restricted version of the theorem.

Lemma (1): Let $A^{(5)}$ denote any symmetric interval in $D^{(5)}$. (1) There exists exactly one solution $\mathscr{F}(t)$ of Eq. (49) such that its domain is $A^{(5)}$, and such that

$$\mathscr{F}(\mathbf{x},\mathbf{x},t) = I. \tag{A1}$$

for every $\mathbf{x}$ in $A^{(2)}$ and t in T. (2) For all $\mathbf{x}, \mathbf{y}, \mathbf{x}_1$ in $A^{(2)}$ and t in T, the inverse of $\mathscr{F}(\mathbf{x}, \mathbf{y}, t)$ exists, and

$$\mathscr{F}(\mathbf{x},\mathbf{y},t) = \mathscr{F}(\mathbf{x},\mathbf{x}_1,t)\mathscr{F}(\mathbf{x}_1,\mathbf{y},t), \tag{A2}$$

$$\mathscr{F}(\mathbf{x},\mathbf{y},t)^{-1} = \mathscr{F}(\mathbf{y},\mathbf{x},t). \tag{A3}$$

(3) $\mathscr{F}(t)$ is holomorphic in $A^{(5)}$.

Proof: Introduce components $\Gamma_i(t)$ of the 1 form $\Gamma(t)$ by

$$\Gamma(t) = dx^i\Gamma_i(t).$$

Consider the ordinary differential equation

$$\frac{\partial f(x^1,\mathbf{y},t)}{\partial x^1} = \Gamma_1(x^1,y^2,t)\Omega f(x^1,\mathbf{y},t), \tag{A4}$$

subject to the initial condition

$$f(y^1,\mathbf{y},t) = I, \tag{A5}$$

in the domain $A_1 \times A^{(2)} \times T$. According to a standard theorem,[31] the solution of (A4) and (A5) exists, is unique, has an inverse, and is holomorphic in $A_1 \times A^{(2)} \times T$. Next, consider the ordinary differential equation

$$\frac{\partial \mathscr{F}(\mathbf{x},\mathbf{y},t)}{\partial x^2} = \Gamma_2(\mathbf{x},t)\Omega\mathscr{F}(\mathbf{x},\mathbf{y},t), \tag{A6}$$

subject to the initial condition

$$\mathscr{F}(x^1, y^2, \mathbf{y}, t) = f(x^1, \mathbf{y}, t) \tag{A7}$$

in the domain $A^{(5)}$. According to the same standard theorem,[31] the solution of (A6) and (A7) exists, is unique, has an inverse, and is holomorphic in $A^{(5)}$.

Now, the integrability condition (48) has the component form

$$\frac{\partial \Gamma_2(\mathbf{x},t)}{\partial x^1} - \frac{\partial \Gamma_1(\mathbf{x},t)}{\partial x^2}$$

$$= \Gamma_1(\mathbf{x},t)\Omega\Gamma_2(\mathbf{x},t) - \Gamma_2(\mathbf{x},t)\Omega\Gamma_1(x,t).$$

After differentiating (A6) with respect to x^1, using the integrability condition to replace $\partial\Gamma_2/\partial x^1$, and regrouping terms, we get

$$\frac{\partial g(\mathbf{x},\mathbf{y},t)}{\partial \mathbf{x}^2} = \Gamma_2(\mathbf{x},t)\Omega g(\mathbf{x},t), \tag{A8}$$

where

$$g(\mathbf{x},\mathbf{y},t) := \frac{\partial \mathscr{F}(\mathbf{x},\mathbf{y},t)}{\partial x^1} - \Gamma_1(\mathbf{x},t)\Omega\mathscr{F}(\mathbf{x},\mathbf{y},t). \tag{A9}$$

However, note that $g(\mathbf{x},\mathbf{y},t)$ and $\mathscr{F}(\mathbf{x},\mathbf{y},t)$ are solutions of the same ordinary linear homogeneous differential equation (A8) and (A6), respectively. Therefore, since $\mathscr{F}(\mathbf{x},\mathbf{y},t)$ has an inverse (so that its columns are linearly independent and constitute a fundamental pair of solutions), there exists a 2×2 matrix $k(x^1,\mathbf{y},t)$ such that

$$g(\mathbf{x},\mathbf{y},t) = \mathscr{F}(\mathbf{x},\mathbf{y},t)k(x^1,\mathbf{y},t) \tag{A10}$$

throughout $A^{(5)}$. Now $x^2 = y^2$ in (A10), whereupon (A7) implies

$$g(x^1,y^2,\mathbf{y},t) = f(x^1,\mathbf{y},t)k(x^1,\mathbf{y},t). \tag{A11}$$

However, from (A4) and (A7), if we set $x^2 = y^2$ in (A9), we get

$$g(x^1,y^2,\mathbf{y},t) = 0.$$

So, from (A11), $k(x^1,\mathbf{y},t) = 0$, whereupon (A10) implies

$$g(\mathbf{x},\mathbf{y},t) = 0. \tag{A12}$$

Also, upon replacing x^1 by y^1 in (A7), Eq. (A5) yields

$$\mathscr{F}(\mathbf{y},\mathbf{y},t) = I. \tag{A13}$$

Equations (A9), (A12), (A6), and (A13) establish the main part of the lemma.

Next, we prove (A2) and (A3). Since $\mathscr{F}(\mathbf{x},\mathbf{x}_1,t)$ is a fundamental solution of Eq. (49), there exists a 2×2 matrix $M(\mathbf{x}_1,\mathbf{y},t)$ such that

$$\mathscr{F}(\mathbf{x},\mathbf{y},t) = \mathscr{F}(\mathbf{x},\mathbf{x}_1,t)M(\mathbf{x}_1,\mathbf{y},t).$$

Set $\mathbf{x} = \mathbf{x}_1$ above, and use (A1) to get (A2). To get (A3), set $\mathbf{x} = \mathbf{y}$ in (A2).

We want to extend the solutions in symmetric intervals to $D^{(5)}$. This process has some of the features of analytic continuation in the complex plane, and Lemma (1) has already supplied us with the analog of the power series. The next lemma supplies key results which make the continuation possible.

Lemma (2): Suppose

$$A^{(5)} = A^{(2)}\times A^{(2)}\times T, \quad A^{(5)\prime} = A^{(2)\prime}\times A^{(2)\prime}\times T'$$

are any overlapping symmetric intervals in $D^{(5)}$, and $\mathscr{F}(t)$ and $\mathscr{F}'(t)$ are the solutions of Eqs. (49) and (A1) in the respective intervals. (1) In the region of overlap,

$$\mathscr{F}'(\mathbf{x},\mathbf{y},t) = \mathscr{F}(\mathbf{x},\mathbf{y},t). \tag{A14}$$

(2) For all $\mathbf{y}$ in $A^{(2)}$, $\mathbf{y}'$ in $A^{(2)\prime}$, $\mathbf{x}$ in $A^{(2)}\cap A^{(2)\prime}$, and t in $T\cap T'$,

$$\mathscr{F}'(y',x,t)\mathscr{F}(x,y,t) \tag{A15}$$

is independent of $\mathbf{x}$.

Proof: For all t in $T\cap T'$, $\mathbf{x}$ in $A^{(2)}\cap A^{(2)\prime}$, $\mathbf{y}$ in $A^{(2)}$, $\mathbf{y}'$ in $A^{(2)\prime}$, Eq. (49) yields $[d = (\partial/\partial x^i)dx^i]$

$$d[\mathscr{F}'(\mathbf{x},\mathbf{y}',t)^{-1}\cdot\mathscr{F}(\mathbf{x},\mathbf{y},t)] = 0.$$

So $M(\mathbf{y},\mathbf{y}',t)$ exists such that

$$\mathscr{F}'(\mathbf{x},\mathbf{y}',t)^{-1}\cdot\mathscr{F}(\mathbf{x},\mathbf{y},t) = M(\mathbf{y},\mathbf{y}',t). \tag{A16}$$

Use (A3) and (A16) to get (A15). Set $\mathbf{x} = \mathbf{y} = \mathbf{y}'$ in (A16) to prove $M(\mathbf{y},\mathbf{y},t) = I$; then set $\mathbf{y}' = \mathbf{y}$ in (A16) to get (A14).

The significance of (A14) is that it tells us that *the value of* $\mathscr{F}(\mathbf{x},\mathbf{y},t)$ *is independent of the specific symmetric interval in* $D^{(5)}$ *which covers* $(\mathbf{x},\mathbf{y},t)$. As regards (A15), it suggests how $\mathscr{F}(\mathbf{x},\mathbf{y},t)$ can be defined for arbitrary $(\mathbf{x},\mathbf{y},t)$ in $D^{(5)}$ by using products like that in (A15).

To help us formalize the idea, we need some more definitions. For any point $(\mathbf{x},\mathbf{y},t)$ in $D^{(5)}$, we let $L(\mathbf{x},\mathbf{y},t)$ denote the set of all simple[32] oriented smooth lines in $D_t^{(2)}$ which join $\mathbf{y}$ to $\mathbf{x}$; $\mathbf{y}$ is the initial and $\mathbf{x}$ is the final point. $\mathbf{l}$ will denote any line in $L(\mathbf{x},\mathbf{y},t)$.

Since $\mathbf{l}$ is compact, there exists at least one finite sequence of intervals $A_i^{(2)}$ in $D_t^{(2)}$ such that

$$0\leqslant i\leqslant n,\ \mathbf{y} \text{ is in } A_0^{(2)},\ \mathbf{x} \text{ is in } A_n^{(2)},$$

$$\bigcup_{i=0}^{n} A_i^{(2)} \text{ contains } \mathbf{l},$$

$$\mathbf{l}\cap A_{i-1}^{(2)}\cap A_i^{(2)} \text{ is not void } (1\leqslant i\leqslant n).$$

The set of all finite sequences of intervals in $D_t^{(2)}$ which satisfy the above conditions will be denoted by $K(\mathbf{l})$.

From the above definition, if $(A_0^{(2)},\cdots A_n^{(2)})$ is any given member of $K(\mathbf{l})$, there exists at least one sequence of points $\mathbf{x}$ in $D_t^{(2)}$ such that

$$\begin{gathered} 0\leqslant j\leqslant n+1, \quad \mathbf{x}_0 = \mathbf{y}, \quad \mathbf{x}_{n+1} = \mathbf{x}, \\ \mathbf{x}_i \text{ lies in } A_{i-1}^{(2)}\cap A_i^{(2)}\ (1\leqslant i\leqslant n). \end{gathered} \tag{A17}$$

Moreover, these points may always be selected so that they lie on $\mathbf{l}$. Note that $\mathbf{x}_{i+1}$ and $\mathbf{x}_i$ are both in $A_i^{(2)}$. Define

$$A_i^{(5)} := A_i^{(2)}\times A_i^{(2)}\times T_i,$$

where T_i is any covering of t in C such that $A_i^{(5)}$ is in $D^{(5)}$. [Clearly, an $A_i^{(5)}$ exists since (x_i,y_i,t) lies in $D^{(5)}$ for all x_i, y_i in $A_i^{(2)}$.] Define

$$\mathscr{F}(\mathbf{x},\mathbf{y},t,\mathbf{l}) := \prod_{i=0}^{n} \mathscr{F}_i(\mathbf{x}_{i+1},\mathbf{x}_i,t), \tag{A18}$$

where $\mathscr{F}_i(t)$ denotes the solution of Eqs. (49) and (A1) in the symmetric interval $A_i^{(5)}$. *In the next two lemmas, we shall establish that the value of the above expression* (A18) *is uniquely determined by* $\mathbf{x}$, $\mathbf{y}$, t, *and* $\mathbf{l}$.

Lemma (3): For any given $\mathbf{l}$ in $L(\mathbf{x},\mathbf{y},t)$ and for any given $(A_0^{(2)},\ldots,A_n^{(2)})$ in $K(\mathbf{l})$, the expression (A18) has a value which is independent of the choice of the points $\mathbf{x}_1,\ldots,\mathbf{x}_n$ which satisfy (A17).

Proof: $A_{i-1}^{(2)}$ and $A_i^{(2)}$ overlap. Also, T_{i-1} and T_i overlap, because they both contain t. Therefore, $A_{i-1}^{(5)}$ and $A_i^{(5)}$ overlap $(1\leqslant i\leqslant n)$. The conclusion then follows from the second part [Eq. (A15)] of Lemma (2).

Q.E.D.

Lemma (4): For any given $\mathbf{l}$ in $L(\mathbf{x},\mathbf{y},t)$, the expression (A18) has a value which is independent of the choice of the member $(A_0^{(2)},\ldots,A_n^{(2)})$ of $K(\mathbf{l})$.

Proof: Since the value of (A18) is independent of the choice of $\mathbf{x}_1,\ldots,\mathbf{x}_n$, let us select these points so that all lie on $\mathbf{l}$.

Then consider that, by definition of a simple[32] oriented smooth line, there exist C^∞ functions $z^1(\sigma)$ and $z^2(\sigma)$ of a real variable σ defined over an open interval (c, d) such that **l** is the set of all points

$$z(s) := (z^1(\sigma), z^2(\sigma)), \quad \text{for } a \leqslant \sigma \leqslant b,$$

where

$$c < a \leqslant b < d,$$

$$\mathbf{z}(a) = \mathbf{y}, \quad \mathbf{z}(b) = \mathbf{x},$$

$$\left| \frac{dz^1(\sigma)}{ds} \right| + \left| \frac{dz^2(\sigma)}{ds} \right| \neq 0, \tag{A19}$$

$\mathbf{z}(\sigma) \neq \mathbf{z}(\sigma')$ if $\sigma \neq \sigma'$, except possibly when $\sigma = a$ and $\sigma' = b$.

Consider the ordinary differential equation

$$\frac{df(\sigma,\sigma_0,t)}{d\sigma} = \frac{dz^i(\sigma)}{d\sigma} \Gamma_i(\mathbf{z}(\sigma),t)\Omega f(\sigma,\sigma_0,t),$$

$$c < \sigma_0 < d, \quad f(\sigma_0,\sigma_0,t) = I.$$

We have the standard results

$$f(\sigma,\sigma',t) f(\sigma',\sigma'',t) = f(\sigma,\sigma'',t),$$
$$f(\sigma_{i+1},\sigma_i,t) = \mathscr{F}_i(\mathbf{x}_{i+1},\mathbf{x}_i,t), \tag{A20}$$

where σ_j is that value of σ for which

$$\mathbf{x}_j = \mathbf{z}(\sigma_j) \quad (0 \leqslant j \leqslant n+1).$$

From Eqs. (A18), (A19), and (A20),

$$\mathscr{F}(x,y,t,l) = f(b,a,t). \qquad \text{Q.E.D.}$$

The final stages of our proof require that we say something about the dependence of $\mathscr{F}(x,y,t,l)$ on l. Suppose l and l' are any members of $L(x,y,t)$. We shall write

$$l \sim l'$$

whenever there exists at least one finite sequence of members l_k of $L(x,y,t)$, and there exists at least one family of intervals $A^{(2)}_{ki}$ in $D^{(2)}_t$ such that

$$0 \leqslant k \leqslant N, \quad l_0 = l, \quad l_N = l',$$

$$0 \leqslant i \leqslant n_k,$$

$(A^{(2)}_{k0},\dots,A^{(2)}_{kn_k})$ is a member both of $K(l_{k-1})$ and of $K(l_k)$ $(1 \leqslant k \leqslant N)$. (A21)

[This is one way of formalizing the concept of being able to transform **l** into **l**′ continuously without leaving $L(\mathbf{x},\mathbf{y},t)$.] It is easy to show that $\sim$ is an equivalence relation.

Lemma (5): If $\mathbf{l} \sim \mathbf{l}'$,

$$\mathscr{F}(\mathbf{x},\mathbf{y},t,\mathbf{l}) = \mathscr{F}(\mathbf{x},\mathbf{y},t,\mathbf{l}').$$

Proof: From (A21), the same expression of the form (A18) can be used both for $\mathbf{l}_{k-1}$ and for $\mathbf{l}_k$. Therefore,

$$\mathscr{F}(\mathbf{x},\mathbf{y},t,\mathbf{l}_{k-1}) = \mathscr{F}(\mathbf{x},\mathbf{y},t,\mathbf{l}_k), \quad 0 \leqslant k \leqslant N$$

which gives us our conclusion, since $\mathbf{l}_0 = \mathbf{l}$, and $\mathbf{l}_N = \mathbf{l}'$.

We shall say only a few more words about the final stages of the proof. The above lemma enables us to replace the notation $\mathscr{F}(\mathbf{x},\mathbf{y},t,\mathbf{l})$ by the notation $\mathscr{F}(\mathbf{x},\mathbf{y},t,v)$, where v is a label for the various equivalence ($\sim$) classes of lines in $L(\mathbf{x},\mathbf{y},t)$. Now, suppose $S^{(2)}_t$ is any simply connected subregion of $D^{(2)}_t$, and $S^{(5)}$ is the set of all $(\mathbf{x},\mathbf{y},t)$ in $D^{(5)}$ such that **x** and **y** are both in $S^{(2)}_t$. Suppose, furthermore, that we restrict $L(\mathbf{x},\mathbf{y},t)$ so that its members all lie in $S^{(2)}_t$. Then, all lines in this restricted $L(\mathbf{x},\mathbf{y},t)$ are equivalent ($\sim$), and $\mathscr{F}$ (so restricted) is single valued and is that branch of the function which satisfies (A1).

We feel that the remainder of the proof is sufficiently straightforward so that we can stop here. It is perhaps worth noting that the same proof is applicable to any completely integrable $n \times n$ linear homogeneous differential equation like Eq. (49), for which the one form $\Gamma\Omega$ is a holomorphic function of (x^1, x^2, t); the extension to any finite number of coordinates and parameters involves no difficulties.

[1]More specifically, this the homogeneous Hilbert problem for nonsingular square matrix functions of a complex variable. The theory is given, for example, in Chap. 18, N.I. Muskhelishvili, *Singular Integral Equations* by (Noordhoff, Groningen, 1953). In particular, Eq. (127.15) in this reference corresponds to our Eq. (9). We differ in only one way from the conventions of Muskhelishvili, viz., our $X_\pm$ are the transposes of their $X^{\pm}$, and our G is their $(G')^{-1}$; therefore, where we refer to rows, he would refer to columns.

[2]R. Geroch, J. Math. Phys. **12**, 918 (1971). Though this paper is largely on space–times with one Killing vector, Sec. 2 presents the essential idea for the group K.

[3]R. Geroch, J. Math. Phys. **13**, 394 (1972).

[4]W. Kinnersley, J. Math. Phys. **18**, 1529 (1977). this paper clarified the field equation symmetries which are responsible for K. It also enlarged K to form a group K' which covers the stationary axially symmetric electrovacs.

[5]W. Kinnersley and D. Chitre, J. Math. Phys. **18**, 1538 (1977).

[6]W. Kinnersley and D. Chitre, J. Math. Phys. **19**, 1926 (1978).

[7]W. Kinnersley and D. Chitre, J. Math. Phys. **19**, 2037 (1978). This paper contains a new five parameter vacuum solution which has the $\delta = 2$ Tomimatsu–Sato solution as a special case.

[8]Actually, the general theory in this paper is applicable to the cases for which $d(\det h)$ is identically zero or is identically a null one form. However, some of the specific statements concerning our choice of gauge and the character of the singularities of the F potential (as discussed in Sec. 3) have to be modified or revised in these special cases.

[10]In all statements concerning t dependence, the Riemann sphere topology is used.

[11]W. Kinnersley, J. Math. Phys. **14**, 651 (1973).

[12]More precisely, the singularities occur at the zeros of $(2t)^{-1}\lambda(\mathbf{x}, t)$, as will become clear in Sec. 4.

[13]I. Hauser and F.J. Ernst, Phys. Rev. D **20**, 362 (1979).

[14]I. Hauser and F.J. Ernst, Phys. Rev. D **20**, 1783 (1979). This paper contains the extension of the integral equation to electrovacs. In the extension, 3×3 matrices are employed, and $u(t)$ is a member of SU (2,1).

[15]C. Hoenselaers, W. Kinnersley, and B. Xanthopoulos, Phys. Rev. Lett. **42**, 481 (1979); J. Math. Phys. **20**, 2530 (1979).

[16]V.A. Belinskii and V.E. Zakharov, Sov. Phys. JETP **75**, 1953 (1978).

[17]D. Maison, Phys. Rev. Lett. **41**, 521 (1978); J. Math. Phys. **20**, 871 (1979).

[18]B.K. Harrison, Phys. Rev. Lett. **41**, 1197 (1978).

[19]G. Neugebauer, J. Phys. A **12**, L67 (1979).

[20]Our n-pole solution in Ref. 14 is actually for an electrovac. It is the vacuum specialization of our results which correspond to those in Refs. 15 and 16.

[21]This follows, for example, from Eq. (125.9) in Ref. 1 and from the statement immediately following the equation in the book.

[22]Values of **x** corresponding to $z^2 + \rho^2 = 0$ [$\pm := -(\sin(\det h))$] are excluded from the domains of both Eqs. (7) and (8) and of the HHP defined by Eqs. (9) (11) and the subsequent conditions. The reason is that $F(t)$ has a singularity at $t = \infty$ for such values of **x**; this will become clear in Sec. 4.

[23]In the K–C papers, the Ernst potential is the upper left element of their H, whereas $\mathscr{E}$ is the lower right element in our H.

[24]One can use Eq. (21d), with $S_{ii} = 0$ and $f^{-1}P^{-2} = \exp(2\Gamma)$ of F.J. Ernst, J. Math. Phys. **15**, 1409 (1974). Alternatively, one can use Eq. (2.12) of Ref. 4. These are simply examples.

[25]Our notations and conventions concerning differential forms, Grassmann products, duality operations, etc., are described in an Appendix to the paper by I. Hauser and F.J. Ernst, J. Math. Phys. **19**, 1316 (1978).

[26]The term "connected" is used here in the sense of "arcwise connected." Specifically, a region is called *connected* if any two of its points lie on at least one simple smooth arc which is contained by the region.

[27]In this extension to complex $\mathbf{x}$, $[H(\mathbf{x})]^*$ denotes the complex conjugate of $[H(\mathbf{x}^*)]$, and $[H(\mathbf{x})]^\dagger$ denotes the Hermitian conjugate of $[H(\mathbf{x}^*)]$. The duality oprator * may be defined by Eqs. (24) and (25), and Eqs. (26), (30), (31), etc., are correct, as they stand, for the extension. *The distinction between $V(+)$ and $V(-)$ is unnecessary for $[H(\mathbf{x})]$; in fact, the $V(-)$ expressions and relations can be obtained from the $V(+)$ ones by the substitutions* $x^1 \to x^1$, $x^2 \to -ix^2$, $z \to z$, $\rho \to i\rho$, and $* \to i*$. In our equations, however, we shall continue to maintain a formal distinction between the two "cases" as an aid to those who would prefer to think in terms of real $\mathbf{x}$ or who would prefer to maintain the distinction.

[28]We have not yet been able to find a reference for this theorem and would welcome one. The related theorems which we have found are strictly local in their existence claims or say nothing about analyticity or about parameter dependence.

[29]See, for example, Chap. XIX, E. L. Ince, *Ordinary Differential Equations* (Dover, New Yor, 1956).

[30]For this purpose, we used Eqs. (3.1), (3.3), and (3.6) of Ref. 13.

[31]See, for example, Solomon Lefschetz, II, *Differential Equations: Geometric Theory*, 2nd ed. (Dover, New York, 1977), Chap. II, theorem (10.3).

[32]However, we do permit the line to return to $\mathbf{y}$; i.e., we do permit $\mathbf{x} = \mathbf{y}$. Therefore, the term "simple" does not have its conventional meaning. A definition is given in Eq. (A19).

Symmetric vectors and algebraic classification

Elhanan Leibowitz
Department of Mathematics and Department of Physics, Ben Gurion University of the Negev, Beer-Sheva, Israel

(Received 5 January 1979; accepted for publication 25 May 1979)

The concept of symmetric vector field in Riemannian manifolds, which arises in the study of relativistic cosmological models, is analyzed. Symmetric vectors are tied up with the algebraic properties of the manifold curvature. A procedure for generating a congruence of symmetric fields out of a given pair is outlined. The case of a three-dimensional manifold of constant curvature ("isotropic universe") is studied in detail, with all its symmetric vector fields being explicitly constructed.

I. INTRODUCTION

The concept of locally symmetric vector field in general Riemannian manifold has recently been introduced by Walker.[1] He was motivated by his earlier investigations[2] on possible laws of orientation of galaxies. The distribution of galaxies in the three-dimensional curved space is assumed, in the standard cosmological model of general relativity, to be isotropic (and homogeneous). Most of the galaxies, however, reveal in their structure one or more axis of symmetry, and the question arises as to whether the orientation of the galaxies ought to be random, or there are other laws of orientation which do not violate cosmological principles. The galactic axes of symmetry are represented by unit vector fields, and thus one is led to consider the symmetry of such fields. A unit vector field is defined by Walker[1] to be symmetric if it exhibits rotational symmetry. To be precise, let M be an n-Riemannian manifold with metric tensor g. A unit vector field $V \in T(M)$ is said to have symmetry about a point $p \in M$ if it is invariant under all linear transformations of normal coordinates centered at p which leave $V(p)$ and $g(p)$ invariant. The condition of symmetry is expressed as the vanishing of a certain function Φ of the normal coordinates ($\Phi = 0$ identically at p). The vector field is said to have first (second) order local symmetry about p if a weaker condition is satisfied by Φ, namely, its first (and second) order partial derivatives vanish at p. Finally, the vector field is defined to have first (second) order local symmetry in M if it has first (second) local symmetry about every $p \in M$. After some manipulations, Walker arrives at a covariant formulation of the condition of local symmetry. It turns out that the case of three-dimensional Riemannian manifolds (which is of fundamental importance for the underlying cosmological considerations) is special. A necessary and sufficient condition for the unit vector field V to have first order local symmetry is that in local chart $(x^1,\ldots,x^n)$ it satisfies

$$\nabla_\nu V_\mu = \alpha(g_{\mu\nu} - V_\mu V_\nu), \qquad \text{for } n > 3$$

$$\nabla_\nu V_\mu = \alpha(g_{\mu\nu} - V_\mu V_\nu) + \beta e_{\mu\nu\lambda} V^\lambda, \quad \text{for } n = 3$$

where $\mu,\nu = 1,\ldots,n$; ∇_ν denotes covariant derivative with respect to the metric g; $e_{\mu\nu\lambda}$ is the three-dimensional alternating tensor; and α and β being some scalars. The vector field V has second order local symmetry iff in addition to these conditions it satifies

$$\alpha_\mu = \alpha_\nu V^\nu V_\mu, \qquad \text{for } n > 3$$

$$\beta = 0 \text{ and } \alpha_\mu = \alpha_\nu V^\nu V_\mu,$$

$$\text{or } \alpha = 0 \text{ and } \beta = \text{constant}, \qquad \text{for } n = 3$$

whereby $\alpha_\mu = \alpha_{,\mu}$ denotes the partial derivative of α with respect to x^μ. The existence of a locally symmetric vector field restricts the manifold, and it is easy to establish a canonical form for the metric tensor of a manifold admitting such field.

In the next section the implications of the existence of one or more locally symmetric vector field will be studied. In Sec. III the same investigation will be carried out for the three-dimensional case, and in Sec. IV we focus our attention on a three-manifold of constant curvature. Section V is devoted to concluding remarks.

II. $n>3$

In this section we analyze the case $n>3$. Let V_μ be a first order locally symmetric vector field

$$\nabla_\nu V_\mu = \alpha(g_{\mu\nu} - V_\mu V_\nu), \quad V_\mu V^\mu = 1. \tag{2.1}$$

The integrability conditions of these equations are[1,3]

$$R_{\mu o \lambda\tau} = (g_{\mu\lambda} - V_\mu V_\lambda)\alpha_\tau - (g_{\mu\tau} - V_\mu V_\tau)\alpha_\lambda + \alpha^2(g_{\mu\lambda} V_\tau - g_{\mu\tau} V_\lambda), \tag{2.2}$$

where the subscript o denotes contraction with the vector V_μ, e.g., $R_{\mu o\lambda\tau} = R_{\mu\nu\lambda\tau} V^\nu$. From Eq. (2.2) we have:

$$R_{\mu o \nu o} = (\alpha^2 + \alpha_o)(g_{\mu\nu} - V_\mu V_\nu), \tag{2.3}$$

$$R_{\mu o} = (n-2)\alpha_\mu + [(n-1)\alpha^2 + \alpha_o] V_\mu. \tag{2.4}$$

Theorem 2.1: A first order locally symmetric vector field is a Ricci principal direction if and only if it has second order local symmetry. In that case the associated eigenvalue is $(n-1)(\alpha^2 + \alpha_o)$.

Proof: Let V_μ be a first order locally symmetric vector field which is a Ricci principal direction. Hence (2.4) is satisfied, and in addition

$$R_{\mu o} = p V_\mu,$$

where p is the eigenvalue. The latter is substituted in (2.4), to yield

$$(n-2)\alpha_\mu = [p - (n-1)\alpha^2 - \alpha_o] V_\mu.$$

This implies that $\alpha = \alpha(V)$, i.e., V_μ has second order local symmetry. Contraction with V^μ gives $p = (n-1)(\alpha^2 + \alpha_o)$. The converse is proved similarly. □

 0022-2488/80/051141-08$01.00

Now suppose that a manifold admits two first order locally symmetric vector fields V_μ and $\bar{V}_\mu$

$$\nabla_\nu V_\mu = \alpha(g_{\mu\nu} - V_\mu V_\nu),$$
$$\nabla_\nu \bar{V}_\mu = \bar{\alpha}(g_{\mu\nu} - \bar{V}_\mu \bar{V}_\nu).$$

(Note that we do not consider V_μ and $-V_\mu$ as distinct.) Thus in addition to (2.2)–(2.4), we have the integrability conditions:

$$R_{\mu\bar{o}\lambda\tau} = (g_{\mu\lambda} - \bar{V}_\mu \bar{V}_\lambda)\bar{\alpha}_\tau - (g_{\mu\tau} - \bar{V}_\mu \bar{V}_\tau)\bar{\alpha}_\lambda + \bar{\alpha}^2(g_{\mu\lambda}\bar{V}_\tau - g_{\mu\tau}\bar{V}_\lambda), \quad (2.5)$$

$$R_{\mu\bar{o}\nu\bar{o}} = (\bar{\alpha}^2 + \bar{\alpha}_{\bar{o}})(g_{\mu\nu} - \bar{V}_\mu \bar{V}_\nu), \quad (2.6)$$

$$R_{\mu\bar{o}} = (n-2)\bar{\alpha}_\mu + [(n-1)\bar{\alpha}^2 + \bar{\alpha}_{\bar{o}}]\bar{V}_\mu. \quad (2.7)$$

Contracting (2.3) with $\bar{V}^\mu \bar{V}^\nu$ and (2.6) with $V^\mu V^\nu$ we get two expressions for $R_{o\bar{o}o\bar{o}}$, from which follows:

$$\bar{\alpha}^2 + \bar{\alpha}_{\bar{o}} = \alpha^2 + \alpha_o. \quad (2.8)$$

Likewise, from (2.4) and (2.7) [in view of (2.8)] follows:

$$\bar{\alpha}_o - \alpha_{\bar{o}} = \phi(\bar{\alpha}_{\bar{o}} - \alpha_o), \quad (2.9)$$

with

$$\phi = g^{\mu\nu} V_\mu \bar{V}_\nu.$$

Contract now Eq. (2.2) with $\bar{V}^\mu \bar{V}^\tau$, to obtain

$$R_{\mu\bar{o}o\bar{o}} = (1-\phi^2)\alpha_\mu + (\alpha^2 + \phi\alpha_{\bar{o}})V_\mu - (\phi\alpha^2 + \alpha_{\bar{o}})\bar{V}_\mu.$$

On the other hand, contracting (2.6) with V^ν and using (2.8) lead to

$$R_{\mu\bar{o}o\bar{o}} = (\alpha^2 + \alpha_o)(V_\mu - \phi\bar{V}_\mu).$$

From the last two equations we can deduce the last identity needed for the next theorem, viz.:

$$(1-\phi^2)\alpha_\mu = (\alpha_o - \phi\alpha_{\bar{o}})V_\mu + (\alpha_{\bar{o}} - \phi\alpha_o)\bar{V}_\mu. \quad (2.10)$$

Theorem 2.2: Let V_μ and $\bar{V}_\mu$ be two first order locally symmetric vector fields. Then there exist two scalars A and B such that

$$R_{\mu\nu}V^\nu = AV_\mu + B\bar{V}_\mu,$$
$$R_{\mu\nu}\bar{V}^\nu = BV_\mu + A\bar{V}_\mu, \quad (2.11)$$

and consequently the vectors $(1/\sqrt{2})(V_\mu + \bar{V}_\mu)$ and $(1/\sqrt{2})(V_\mu - \bar{V}_\mu)$ are Ricci principal directions.

Proof: we have

$$\nabla_\nu V_\mu = \alpha(g_{\mu\nu} - V_\mu V_\nu),$$
$$\nabla_\nu \bar{V}_\mu = \bar{\alpha}(g_{\mu\nu} - \bar{V}_\mu \bar{V}_\nu).$$

As was shown above, this entails that α satisfies (2.10), and $\bar{\alpha}$ satisfies an analogous equation. Substituting (2.10) into (2.4) one obtains:

$$(1-\phi^2)R_{\mu o} = \{(n-2)(\alpha_o - \phi\alpha_{\bar{o}}) + (1-\phi^2)[(n-1)\alpha^2 + \alpha_o]\}V_\mu + \{(n-2)(\alpha_{\bar{o}} - \phi\alpha_o)\}\bar{V}_\mu. \quad \text{(i)}$$

Interchanging the roles of V_μ and $\bar{V}_\mu$ we can similarly write:

$$(1-\phi^2)R_{\mu\bar{o}} = \{(n-2)(\bar{\alpha}_o - \phi\bar{\alpha}_{\bar{o}})\}V_\mu + \{(n-2)(\bar{\alpha}_{\bar{o}} - \phi\bar{\alpha}_o) + (1-\phi^2)[(n-1)\bar{\alpha}^2 + \bar{\alpha}_{\bar{o}}]\}\bar{V}_\mu.$$

Straightforward manipulation of the last equation, with the aid of (2.8) and (2.9), yields:

$$(1-\phi^2)R_{\mu\bar{o}} = \{(n-2)(\alpha_{\bar{o}} - \phi\alpha_o)\}V_\mu + \{(n-2)(\alpha_o - \phi\alpha_{\bar{o}}) + (1-\phi^2)[(n-1)\alpha^2 + \alpha_o]\}\bar{V}_\mu. \quad \text{(ii)}$$

The two expression (i) and (ii) provide the representation stated in the theorem. □

Corollaries:

(i) If a manifold admits two first order locally symmetric vectors, they lie in a subspace of the tangent space at each point spanned by two Ricci principal directions.

(ii) If a manifold admits two first order locally symmetric vector fields, one of them being of the second order, then both vectors are of the second order, both are Ricci principal directions corresponding to a common constant eigenvalue.

(iii) If a manifold admits three first order locally symmetric vector fields, then all of them are of the second order, and consequently are Ricci principal directions corresponding to a common constant eigenvalue.

(iv) If an n-manifold admits $n-1$ linearly independent locally symmetric vector fields, the manifold is of constant curvature.

Remark: some of these corollaries have been derived through a different approach by Gauchman.[4]

Proof:

(i) This proposition follows directly from the theorem.

(ii) Suppose V_μ in the last theorem has second order local symmetry. By theorem (2.1) V_μ is a Ricci principal direction, and from (2.11a) $B = 0$, which by (2.11b) implies that $\bar{V}_\mu$ is a Ricci principal direction with the same eigenvalue, and is of the second order. By theorem (2.1) the common eigenvector is

$$(n-1)(\alpha^2 + \alpha_o) = (n-1)(\bar{\alpha}^2 + \bar{\alpha}_{\bar{o}}), \quad (*)$$

and it is being left to demonstrate its constancy. Since V_μ and $\bar{V}_\mu$ have second order symmetry, $\alpha = \alpha(V)$, $\bar{\alpha} = \bar{\alpha}(\bar{V})$, where $V_\mu = V_{,\mu}$ and $\bar{V}_\mu = \bar{V}_{,\mu}$. Hence the left-hand side of (*) is a function of V, the right-hand side is a function of $\bar{V}$, and as V and $\bar{V}$ are functionally independent (otherwise $\bar{V}_\mu = \pm V_\mu$), both sides must be constant.

(iii) Let U, V and W be three first order locally symmetric vector fields. By the theorem, the six vectors $(1/\sqrt{2})(U \pm V)$, $(1/\sqrt{2})(U \pm W)$, $(1/\sqrt{2})(V \pm W)$ are Ricci principal directions. At least one of the two vectors $(1/\sqrt{2})(U \pm W)$ is not orthogonal to $(1/\sqrt{2})(U + V)$. Without loss of generality we can assume this vector to be $(1/\sqrt{2})(U + W)$. Thus the two vectors $(1/\sqrt{2})(U + V)$ and $(1/\sqrt{2})(U + W)$ constitute a pair of nonorthogonal eigenvectors of the Ricci tensor, and hence they correspond to the same eigenvalue, say p. Since the inner products $(U + V, U + W)$ and $(U + V, V + W)$ are equal, the vectors $(1/\sqrt{2})(U + V)$ and $(1/\sqrt{2})(V + W)$ are not orthogonal, and hence $(1/\sqrt{2})(V + W)$ corresponds also to the same eigenvalue p. Consequently any linear combination of the three vectors $(1/\sqrt{2})(U + V)$, $(1/\sqrt{2})(V + W)$ and $(1/\sqrt{2})(V + W)$, in particular the vector U, is a Ricci principal direction. The proposition is thereby proved in virtue of corollary (ii).

(iv) Let $V_\mu^{(1)},\ldots,V_\mu^{(n-1)}$ be $n-1$ linearly independent locally symmetric vector fields. By (iii) all of them have second order local symmetry, and hence their scalars $\alpha^{(k)}$, $k=1,\dots,n-1$, satisfy

$$\alpha^{(k)}_{\mu}=(\alpha^{(k)}_{\nu}V^{(k)\nu})V^{(k)}_{\mu},$$
$$(\alpha^{(k)})^2+\alpha^{(k)}_{\mu}V^{(k)\mu}=(\alpha^{(1)})^2+\alpha^{(1)}_{\mu}V^{(1)\mu}\equiv -p.$$

The integrability conditions (2.2) for these fields assume the form:

$$T_{\mu\nu\lambda\tau}V^{(k)\nu}=0,\quad k=1,\dots,n-1$$

with

$$T_{\mu\nu\lambda\tau}=R_{\mu\nu\lambda\tau}+p(g_{\mu\lambda}g_{\nu\tau}-g_{\mu\tau}g_{\nu\lambda}).$$

Denote

$$V^{(n)}_{\mu}=e^{\mu\lambda_1\cdots\lambda_{n-1}}V^{(1)}_{\lambda_1}V^{(2)}_{\lambda_2}\dots V^{(n-1)}_{\lambda_{n-1}},$$

Where $\exp(\mu\,\lambda_1,\dots,\lambda_{n-1}$ is the alternating tensor. Since antisymmetrization over $(n+1)$ indices annihilates, we have identically

$$T_{\rho\sigma\mu}{}^{[\nu}e^{\mu\lambda_1\cdots\lambda_{n-1}]}\dots V^{(n-1)}_{\lambda_{n-1}}=0$$

(square bracket symbols are used for antisymmetrization), from which follows that the vector field $V^{(n)}_{\mu}$ too satisfies

$$T_{\mu\nu\lambda\tau}V^{(n)\nu}=0.$$

Thus the ennuple of independent vectors $V^{(1)}_{\mu},\dots,V^{(n)}_{\mu}$ satisfy

$$T_{\mu\nu\lambda\tau}V^{(a)\nu}=0,\quad a=1,\dots,n,$$

which implies

$$T_{\mu\nu\lambda\tau}=0,$$

namely, the manifold is of constant curvature p.

The last corollary deals with linearly independent symmetric vectors. The next theorem shows that there is no upper bound to the number of linearly dependent symmetric vectors which a manifold can admit.

Theorem 2.3: Let a manifold admit two second order locally symmetric vector fields $V^{(a)}_{\mu}, a=1,2$:

$$\nabla_{\nu}V^{(a)}_{\mu}=\alpha^{(a)}(g_{\mu\nu}-V^{(a)}_{\mu}V^{(a)}_{\nu}),$$
$$\alpha^{(1)}=\alpha^{(1)}(V^{(1)}),$$
$$\alpha^{(2)}=\alpha^{(2)}(V^{(2)}),$$

where

$$V^{(a)}_{\mu}=V^{(a)}_{,\mu}.$$

Denote

$$\psi^{(a)}(V^{(a)})=\exp\int\alpha^{(a)}(V^{(a)})dV^{(a)},$$
$$\Delta^2=[\psi^{(1)}]^2+[\psi^{(2)}]^2+2\psi^{(1)}\psi^{(2)}V^{(1)}_{\mu}V^{(2)\mu},$$
$$p^{(a)}=\psi^{(a)}/\Delta.$$

Then the unit vector field

$$V_{\mu}=p^{(1)}V^{(1)}_{\mu}+p^{(2)}V^{(2)}_{\mu}$$

too is a second order locally symmetric vector field, i.e., it satisfies the equation

$$\nabla_{\nu}V_{\mu}=\alpha(g_{\mu\nu}-V_{\mu}V_{\nu}),$$

with

$$\alpha=p^{(1)}\alpha^{(1)}+p^{(2)}\alpha^{(2)}.$$

Proof: From the definition of $\psi^{(a)}$,

$$\psi^{(a)}{}_{,\mu}=\alpha^{(a)}\psi^{(a)}V^{(a)}_{\mu}.$$

With the aid of this relation we differentiate Δ^2:

$$\Delta\Delta_{,\mu}=\Delta^2(p^{(1)}\alpha^{(1)}+p^{(2)}\alpha^{(2)})(p^{(1)}V^{(1)}_{\mu}+p^{(2)}V^{(2)}_{\mu}),$$

i.e.,

$$\frac{\Delta_{,\mu}}{\Delta}=\alpha V_{\mu}.$$

Therefore

$$p^{(1)}{}_{,\mu}=p^{(1)}[\alpha^{(1)}V^{(1)}_{\mu}-\alpha V_{\mu}],$$

$$p^{(2)}{}_{,\mu}=p^{(2)}[\alpha^{(2)}V^{(2)}_{\mu}-\alpha V_{\mu}].$$

When these two relations and the symmetry equations for $V^{(1)}_{\mu}$ and $V^{(2)}_{\mu}$ are substituted into

$$\nabla_{\nu}V_{\mu}=p^{(1)}\nabla_{\nu}V^{(1)}_{\mu}+p^{(2)}\nabla_{\nu}V^{(2)}_{\mu}+p^{(1)}{}_{,\nu}V^{(1)}_{\mu}+p^{(2)}{}_{,\nu}V^{(2)}_{\mu},$$

one obtains finally

$$\nabla_{\nu}V_{\mu}=\alpha(g_{\mu\nu}-V_{\mu}V_{\nu}).$$

The vector field V_{μ} has then local symmetry, necessarily of the second order according to corollary (ii). □

From this theorem we can conclude that the existence of two second order locally symmetric vector fields implies the existence of an infinite number (in fact, a two-parametric congruence) of second order locally symmetric vector fields. Indeed, the indefinite integrals in the definition of $\psi^{(a)}$ contain arbitrary constants. Let $\psi^{(1)}$ be one particular solution of the equation

$$\frac{d\psi^{(1)}}{dV^{(1)}}=\alpha^{(1)}\psi^{(1)},$$

and similarly for $\psi^{(2)}$. Then for arbitrary real numbers $A^{(1)}$ and $A^{(2)}$ we put

$$\Delta^2(A^{(1)},A^{(2)})=[A^{(1)}\psi^{(1)}]^2+[A^{(2)}\psi^{(2)}]^2+2A^{(1)}A^{(2)}\psi^{(1)}\psi^{(2)}V^{(1)}_{\mu}V^{(2)\mu},$$

$$p^{(a)}(A^{(1)},A^{(2)})=\frac{A^{(a)}\psi^{(a)}}{\Delta(A^{(1)},A^{(2)})},\quad a=1,2.$$

A direct calculation verifies that the set of vector fields:

$$V_{\mu}(A^{(1)},A^{(2)})=p^{(1)}(A^{(1)},A^{(2)})V^{(1)}_{\mu}+p^{(2)}(A^{(1)},A^{(2)})V^{(2)}_{\mu}$$

is a two-parameter congruence of second order locally symmetric vector fields.

The statement of theorem 2.3 gives rise to the question whether the existence of two first order locally symmetric vector fields is sufficient to guarantee the existence of a third locally symmetric vector field. The question can be answered in the negative, and it is enough to demonstrate it via an example.

Example: Let R_{n-2} be an $(n-2)$-Euclidean space, and let M_2 be a two-dimensional Riemannian manifold with metric (in local chart xy): $(1+x^2)(dx^2+dy^2)$. Let $M_n=M_2X(x^2)R_{n-2}$, the meaning of this notation being that local coordinates $(u^1,\dots,u^n)=(x,y,u^3,\dots,u^n)$ exist, such that M_n has the metric

TABLE I. Classification of Riemannian manifolds into five categories. The table is exhaustive—every manifold belongs to one of the five classes, and all the five classes are realized.

Class No.		1	2	3	4	5
No. of locally symmetric vector field of	first order	0	1	1	2	∞
	second order	0	0	1	0	∞

$$ds^2 = (1 + x^2)(dx^2 + dy^2) + x^2[(du^3)^2 + (du^4)^2 + \cdots + (du^n)^2].$$

It is readily verified that the two unit vector fields:

$$V^{(1)}_\mu = x\delta^1_\mu + \delta^2_\mu,$$
$$V^{(2)}_\mu = x\delta^1_\mu - \delta^2_\mu,$$

have first order local symmetry, with scalars

$$\alpha^{(1)} = \alpha^{(2)} = \frac{1}{1+x^2}.$$

Since the vectors are derived from the scalars

$$V^{(1)} = \frac{1}{2}x^2 + y,$$
$$V^{(2)} = \frac{1}{2}x^2 - y,$$

the scalar $\alpha^{(1)}$ is manifestly not a function of $V^{(1)}$ alone, and hence $V^{(1)}_\mu$ is not of the second order. Thus by corollary (iii), the manifold does not admit any locally symmetric vectors apart from $V^{(1)}_\mu$ and $V^{(2)}_\mu$. [It is evident that R_{n-2} in the example can be replaced by an arbitrary $(n-2)$ dimensional Riemannian manifold].

The various possibilities with regard to the number of locally symmetric vector fields admitted by a manifold can be summarized in Table I, which classifies all Riemannian manifolds into five categories according to the number of locally symmetric vector fields of first and second order admitted by them.

III. $n = 3$

In this section the case $n = 3$ will be analyzed. Let V_μ be a first order locally symmetric vector field in a three-dimensional Riemannian manifold;

$$\nabla_\nu V_\mu = \alpha(g_{\mu\nu} - V_\mu V_\nu) + \beta e_{\mu\nu o}. \tag{3.1}$$

The integrability conditions of these equations are[1]

$$R_{\mu o\lambda\tau} = (g_{\mu\lambda} - V_\mu V_\lambda)\alpha_\tau - (g_{\mu\tau} - V_\mu V_\tau)\alpha_\lambda + (\alpha^2 - \beta^2)(g_{\mu\lambda}V_\tau - g_{\mu\tau}V_\lambda) + 2\alpha\beta(e_{\mu\lambda\tau} - V_\mu e_{\lambda\tau o}) + (e_{\mu\lambda o}\beta_\tau - e_{\mu\tau o}\beta_\lambda).$$

Contracting with V^τ and antisymmetrizing on (μ,λ) one finds

$$\beta_o + 2\alpha\beta = 0. \tag{3.2}$$

From this we can conclude that

$$2\alpha\beta e_{\mu\lambda\tau} + e_{\mu\lambda o}\beta_\tau - e_{\mu\tau o}\beta_\lambda = -\beta_\mu e_{\lambda\tau o},$$

in virtue of which the expression above for $R_{\mu o\lambda\tau}$ becomes:

$$R_{\mu o\lambda\tau} = (g_{\mu\lambda} - V_\mu V_\lambda)\alpha_\tau - (g_{\mu\tau} - V_\mu V_\tau)\alpha_\lambda + (\alpha^2 - \beta^2)(g_{\mu\lambda}V_\tau - g_{\mu\tau}V_\lambda) - (\beta_\mu + 2\alpha\beta V_\mu)e_{\lambda\tau o}. \tag{3.3}$$

From Eq. (3.3) we have:

$$R_{\mu o\nu o} = (\alpha^2 - \beta^2 + \alpha_o)(g_{\mu\nu} - V_\mu V_\nu), \tag{3.4}$$

$$R_{\mu o} = \alpha_\mu + (2\alpha^2 - 2\beta^2 + \alpha_o)V_\mu + e_{\mu\nu o}\beta^\nu. \tag{3.5}$$

It follows from (3.4) that if a manifold admits two first order locally symmetric vector fields V_μ and $\bar{V}_\mu$, then

$$\bar{\alpha}^2 - \bar{\beta}^2 + \bar{\alpha}_{\bar{o}} = \alpha^2 - \beta^2 + \alpha_o.$$

In (3.4) one can employ the identical vanishing of the three-dimensional Weyl conformal tensor, to express $R_{\mu o\nu o}$ in terms of the Ricci tensor and scalar. One finds then:

$$R_{\mu\nu} + (\alpha^2 - \beta^2 + \alpha_o - \tfrac{1}{2}R)g_{\mu\nu} + (\alpha^2 - \beta^2 + \alpha_o + \tfrac{1}{2}R)V_\mu V_\nu - R_{o\mu}V_\nu - R_{o\nu}V_\mu = 0. \tag{3.6}$$

Theorem 3.1: Let a three-manifold admit a first order locally symmetric vector field V_μ, and let a, b, c be the Ricci eigenvalues with $a \leqslant c \leqslant b$. Then V_μ must be of the form

$$V_\mu = pA_\mu + qB_\mu,$$
$$p^2 = (c-a)/(b-a), \quad q^2 = (b-c)/(b-a),$$

where A_μ and B_μ are two orthonormal Ricci principal directions corresponding to the eigenvalues a and b respectively.

Proof: Let C_μ be a unit vector orthogonal to the two vectors V_μ and $R_{o\mu}$. Contracting (3.6) with C^ν one obtains:

$$R_{\mu\nu}C^\nu + (\alpha^2 - \beta^2 + \alpha_o - \tfrac{1}{2}R)C_\mu = 0,$$

i.e., C_μ is a Ricci principal direction with the eigenvalue

$$c = \tfrac{1}{2}R - \alpha^2 + \beta^2 - \alpha_o$$

(so far, c is not necessarily the middle eigenvalue). Consequently, the vectors V_μ and $R_{o\mu}$ lie in the subspace of the tangent space at each point spanned by two orthonormal Ricci principal directions orthogonal to C_μ, with eigenvalues a and b, $a \leqslant b$, viz.:

$$V_\mu = pA_\mu + qB_\mu,$$

for some scalars p and q, $p^2 + q^2 = 1$, and

$$R_{o\mu} = paA_\mu + qbB_\mu.$$

Equation (3.6) reduces now to

$$p^2b + q^2a - c = 0,$$

implying

$$p^2 = (c-a)/(b-a), q^2 = (b-c)/(b-a),$$

and in particular $a \leqslant c \leqslant b$. □

This theorem furnishes a simple derivation of a result due to Gauchman[4], namely

Corollary: Let a three-manifold admit a first order locally symmetric vector field V_μ, and let a, b, c be the Ricci eigenvalues.

(i) If the manifold is "degenerate", i.e., $a = c \neq b$, then V_μ is the Ricci principal direction corresponding to b.

(ii) If the manifold is "nondegenerate", i.e., $a < c < b$, then V_μ is one of the four vectors

$$\pm\left(\sqrt{\frac{c-a}{b-a}}A_\mu \pm \sqrt{\frac{b-c}{b-a}}B_\mu\right),$$

where A_μ and B_μ are the Ricci principal directions corresponding to a and b.
(If $a = c = b$, then the manifold is of constant curvature).

In particular, it follows that a manifold which is not of constant curvature can admit at most two first order locally symmetric vector fields (recall that we do not distinguish between V_μ and $-V_\mu$), and in view of the next theorem, it can admit at most one second order locally symmetric vector field. The case of constant curvature will be dealt with in the next sec. IV.

We are now in the position to investigate the order of the symmetry. We shall call a Riemannian three manifold axial if two of its Ricci eigenvalues coincide, while the third one is constant. In this terminology, a three-dimensional version of theorem (2.1) can be formulated.

Theorem 3.2: A first order locally symmetric vector field is a Ricci principal direction if and only if it has second order local symmetry or the manifold is axial.

Proof: One direction is obvious, if V_μ has second order local symmetry, then either $\alpha = 0$ and $\beta = \text{const}$, or $\beta = 0$ and $\alpha_\mu = \alpha_o V_\mu$. In both cases (3.5) entails that V_μ is a Ricci principal direction. Likewise, if V_μ is a first order locally symmetric vector field in an axial three-manifold, then by proposition (i) of the last corollary it is a Ricci principal direction.

Conversely, let V_μ be a first order locally symmetric vector field which is a Ricci principal direction. Then, by the foregoing discussion, the manifold is degenerate, two of its Ricci eigenvalues are equal (to a, say), and V_μ is a Ricci principal direction corresponding to an eigenvalue b,

$$R_{o\mu} = bV_\mu,$$
$$R = 2a + b,$$
$$b = 2(\alpha^2 - \beta^2 + \alpha_o). \tag{3.7}$$

Equations (3.5) and (3.6) assume the form

$$\alpha_\mu - \alpha_o V_\mu + e_{\mu\nu o}\beta^\nu = 0, \tag{3.8}$$
$$R_{\mu\nu} - ag_{\mu\nu} + (a-b)V_\mu V_\nu = 0. \tag{3.9}$$

Taking the covariant divergence of (3.9) and employing the Bianchi identity, we obtain

$$b_\mu = 2[b_o - a_o + 2\alpha(b-a)]V_\mu.$$

Contraction with V^μ and substitution back yields

$$b_\mu - b_0 V_\mu = 0. \tag{3.10}$$

The solutions of Eq. (3.10) fall into two groups:

(i) $b_o \neq 0$,

(ii) $b_0 \neq 0$.

In case (i) V_μ by (3.10) is proportional to a gradient, and as V_μ is a unit vector, it must be a gradient. In order to show this we take the derivative of (3.10) with respect to x^ν, and antisymmetrize over μ and ν

$$b_o(V_{\mu,\nu} - V_{\nu,\mu}) + b_{o\nu}V_\mu - b_{o\mu}V_\nu = 0.$$

Contract with V^ν

$$b_{o\mu} = b_{oo}V_\mu,$$

and substitute back

$$b_o(V_{\mu,\nu} - V_{\nu,\mu}) = 0.$$

Thus V_μ is a gradient, $\beta = 0$, $\alpha_\mu = \alpha_o V_\mu$ by (3.8), and V_μ has local symmetry of the second order. In case (ii), by (3.10),

$$b_\mu = 0,$$

and the manifold is axial. □

It is to be noted that in the case of a first order locally symmetric vector field admitted by an axial three manifold, the associated scalars α and β are both harmonic, viz.,

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\alpha = 0, \quad g^{\mu\nu}\nabla_\mu\nabla_\nu\beta = 0.$$

This follows directly from (3.8) and (3.2), when the constancy of the Ricci eigenvalue b is being invoked.

IV. CONSTANT CURVATURE

We now turn to the case which is the most pertinent to the underlying cosmological considerations, namely, to the case of a three-diminsional Riemannian manifold of constant curvature. In this section we enumerate and analyze all the second order locally symmetric vector fields admitted by such a manifold. These fields fall into two categories, viz., locally symmetric vector field V_μ of type 1:

$$\nabla_\nu V_\mu = \beta e_{\mu\nu o}, \quad \beta = \text{constant} \neq 0, \tag{4.1}$$

and locally symmetric vector field V_μ of type 2:

$$\nabla_\nu V_\mu = \alpha(g_{\mu\nu} - V_\mu V_\nu), \quad \alpha_\mu = \alpha_o V_\mu. \tag{4.2}$$

We first address the question of the number of locally symmetric fields of type 1 which are admitted by the manifold. Substitution of $\alpha = 0$, $\beta = $ constant in the integrability conditions (3.3) gives:

$$R_{\mu o\lambda\tau} = -\beta^2(g_{\mu\lambda}V_\tau - g_{\mu\tau}V_\lambda).$$

On the other hand, if the manifold is of constant curvature K, then

$$R_{\mu o\lambda\tau} = -K(g_{\mu\lambda}V_\tau - g_{\mu\tau}V_\lambda).$$

We conclude, therefore:

Theorem 4.1: If a manifold of constant curvature K admits a locally symmetric vector field of type 1 [Eq. (4.1)], then $K = \beta^2$. In particular, a manifold of negative constant curvature does not admit locally symmetric fields of type 1.

Suppose now that a manifold of constant curvature $K = \beta^2$ admits two locally symmetric vector fields of type 1. Then according to the last theorem we can write without loss of generality:

$$\begin{aligned} \nabla_\nu V_\mu &= \beta e_{\mu\nu\lambda}V^\lambda, \\ \nabla_\nu \bar{V}_\mu &= \bar{\beta} e_{\mu\nu\lambda}\bar{V}^\lambda, \end{aligned} \tag{4.3}$$

where $\bar{\beta} = \beta$ or $\bar{\beta} = -\beta$. In the first case, it follows immediately that the inner product $g^{\mu\nu}V_\mu \bar{V}_\nu$ is constant. If the second possibility is realized, i.e.,

$$\nabla_\nu V_\mu = \beta e_{\mu\nu\lambda}V^\lambda,$$
$$\nabla_\nu \bar{V}_\mu = -\beta e_{\mu\nu\lambda}\bar{V}^\lambda,$$

we deduce from this that the two vector fields commute:

$$[V,\bar{V}]_\mu = V^\nu\nabla_\nu\bar{V}_\mu - \bar{V}^\nu\nabla_\nu V_\mu = 0.$$

Furthermore, consider the unit vector field U orthogonal to both V and $\bar{V}$

$$U_\mu = (1-\phi^2)^{-1/2} e_{\mu\nu\lambda} V^\nu \bar{V}^\lambda,$$

where

$$\phi = g_{\mu\nu} V^\mu \bar{V}^\nu.$$

As a consequence of the symmetry conditions for V and $\bar{V}$, this vector field commutes with both V and $\bar{V}$:

$$[U,V]_\mu = [U,\bar{V}]_\mu = 0.$$

Thus the three unit vectors $V_\mu, \bar{V}_\mu$ and U_μ form a triad of commuting linearly independent vectors. We can choose, therefore, a local coordinate chart (xyz), such that

$$V = \frac{\partial}{\partial x}, \quad \bar{V} = \frac{\partial}{\partial y}, \quad U = \frac{\partial}{\partial z}.$$

In view of the relations:

$$g^{\mu\nu} V_\mu V_\nu = g^{\mu\nu} \bar{V}_\mu \bar{V}_\nu = g^{\mu\nu} U_\mu U_\nu = 1,$$

$$g^{\mu\nu} U_\mu V_\nu = g^{\mu\nu} U_\mu \bar{V}_\nu = 0, \quad g^{\mu\nu} V_\mu \bar{V}_\nu = \phi,$$

the metric tensor in this coordinate system assumes the form

$$ds^2 = dx^2 + 2\phi dx dy + dy^2 + dz^2.$$

Thus if the manifold admits two locally symmetric vector fields of type 1 with β and $-\beta$, then the metric can be reduced to this form. It is easily verified that ϕ can indeed be chosen such that $0 < \phi^2 < 1$, the manifold is of constant curvature $K = \beta^2$, and the two vectors $\partial/\partial x$ and $\partial/\partial y$ have local symmetry of type 1. It proves later more convenient to change to new coordinates

$$x \to x + y,$$

$$y \to x - y,$$

in terms of which all the solutions of the symmetry condition (4.1) can be calculated. The final result is formulated in the next theorem.

Theorem 4.2: The most general locally symmetric vector field of type 1 admitted by a manifold of constant curvature β^2, $\beta > 0$, is of either of the two forms:

$$V_\mu = aA_\mu + bB_\mu + cC_\mu, \quad a^2 + b^2 + c^2 = 1,$$
$$\bar{V}_\mu = \bar{a}\bar{A}_\mu + \bar{b}\bar{B}_\mu + \bar{c}\bar{C}_\mu, \quad \bar{a}^2 + \bar{b}^2 + \bar{c}^2 = 1, \tag{4.4}$$

where $\{A,B,C\}$ and $\{\bar{A},\bar{B},\bar{C}\}$ are two fixed triads of orthonormal vector fields, and $a,b,c,\bar{a},\bar{b},\bar{c}$ are arbitrary constants. A local coordinate chart (xyz) exists, in terms of which the metric tensor is

$$ds^2 = \cos^2 z dx^2 + \sin^2 z dy^2 + (1/\beta^2) dz^2, \tag{4.5}$$

and the two triads are given by:

$$A = \cos\theta\Big(-tgz\frac{\partial}{\partial x} + \cot z\frac{\partial}{\partial y}\Big) + \beta\sin\theta\frac{\partial}{\partial z},$$
$$B = \sin\theta\Big(tgz\frac{\partial}{\partial x} - \cot z\frac{\partial}{\partial y}\Big) + \beta\cos\theta\frac{\partial}{\partial z},$$
$$C = \frac{\partial}{\partial x} + \frac{\partial}{\partial y},$$
$$\bar{A} = \cos\bar{\theta}\Big(tgz\frac{\partial}{\partial x} + \cot z\frac{\partial}{\partial y}\Big) - \beta\sin\bar{\theta}\frac{\partial}{\partial z},$$
$$\bar{B} = \sin\bar{\theta}\Big(tgz\frac{\partial}{\partial x} + \cot z\frac{\partial}{\partial y}\Big) + \beta\cos\bar{\theta}\frac{\partial}{\partial z},$$
$$\bar{C} = \frac{\partial}{\partial x} - \frac{\partial}{\partial y}, \tag{4.6}$$

with

$$\theta = \beta(x+y), \quad \bar{\theta} = \beta(x-y).$$

All the vectors V_μ satisfy

$$\nabla_\nu V_\mu = \beta e_{\mu\nu\lambda} V^\lambda,$$

and all the vectors $\bar{V}_\mu$ satisfy

$$\nabla_\nu \bar{V}_\mu = -\beta e_{\mu\nu\lambda} \bar{V}^\lambda.$$

Proof: The line element (4.5) is indeed positive definite, and direct calculation reveals that it satisfies

$$R_{\mu\nu} = -2\beta^2 g_{\mu\nu},$$

i.e., it describes a manifold of constant curvature. It follows from the isometry of all manifolds of the same dimension and the same curvature that for a three-dimensional manifold of constant curvature β^2 a local coordinate chart (xyz) exists, in terms of which the metric tensor is given according to (4.5). Employing the sign convention

$$e_{123} = g^{1/2} = +\beta\cos z\sin z,$$

it is easily verified that the two triads $\{A,B,C\}$ and $\{\bar{A},\bar{B},\bar{C}\}$ have, indeed, local symmetry of type 1 with constants $+\beta$ and $-\beta$, respectively.

Conversely, suppose that U_μ is a locally symmetric vector field of type 1, with associated constant $+\beta$. Since the inner product of vector fields of type 1 having the same associated constant are constant, it follows that $\langle U,A\rangle$, $\langle U,B\rangle$, and $\langle U,C\rangle$ are constants. Hence, U_μ has the first form of (4.4). Similarly, if the associated constant of U_μ is $-\beta$, then U_μ has the second form of (4.4). □

We turn now to the case of locally symmetric vector fields of type 2. Again we have to distinguish between manifolds of positive and negative constant curvature. For a manifold of positive constant curvature β^2, it is convenient to use the coordinates (xyz) of theorem 4.2. The metric tensor is given by (4.5), and it is possible to find solutions for the symmetry Eq. (4.2) in a direct maner. By a tedious calculation one finds the following solution:

$$V^{(1)}_\mu = -(1-\cos^2\beta x\cos^2 z)^{-1/2}\Big(\sin\beta x\cos z\frac{\partial}{\partial x} + \frac{1}{\beta}\cos\beta x\sin z\frac{\partial}{\partial z}\Big),$$

with the associated scalar:

$$\alpha^{(1)} = -\beta\cos\beta x\cos z(1-\cos^2\beta x\cos^2 z)^{-1/2}.$$

Three more solutions of a similar character are obtained, and it facilitates the writing to express all the four solutions $V^{(a)}_\mu$, $a = 1,2,3,4$ in a uniform way, viz.:

$$V^{(a)}_\mu = (1/\beta)[1-(\xi^{(a)})^2]^{-1/2}\xi^{(a)}{}_{,\mu}$$

(no summation over a), where the four scalars $\xi^{(a)}$ are given by:

$$\begin{aligned}\xi^{(1)} &= \cos \beta x \cos z,\\ \xi^{(2)} &= \sin \beta x \cos z,\\ \xi^{(3)} &= \cos \beta y \sin z,\\ \xi^{(4)} &= \sin \beta y \sin z.\end{aligned} \tag{4.7}$$

These vector fields satisfy:

$$\nabla_\nu V_\mu^{(a)} = \alpha^{(a)}(g_{\mu\nu} - V_\mu^{(a)} V_\nu^{(a)}),$$

with

$$\alpha^{(a)} = -\beta \xi^{(a)} [1 - (\xi^{(a)})^2]^{-1/2}.$$

Any three of the four vectors are linearly independent, and none of the four is a linear combination with constant coefficients of the remaining three. Furthermore, the conditions for local symmetry of type 2 in three dimensions are analogous to the conditions for second order local symmetry in higher dimensions. Theorem 2.3, therefore, together with the remarks following its proof, can be consulted to conclude that there exists an infinite number of locally symmetric vector fields of type 2. It can be shown (the details will be published elsewhere) that a process of generating more vectors out of these four along the line of theorem 2.3 does, in fact, exhaust the set of locally symmetric vector fields of type 2 admitted by the manifold. The precise meaning of this statement is contained in the following theorem, which gives also a summary of the foregoing discussion.

Theorem 4.3: A three-dimensional Riemannian manifold of positive constant curvature $\beta^2, \beta > 0$, admits two two-parameter congruences of locally symmetric vector fields of type 1, and a four-parameter congruence of locally symmetric vector fields of type 2. A local chart of coordinates $(x^1, x^2, x^3) = (x, y, z)$ exists, in terms of which the metric tensor is given by (4.5), the two congruences of locally symmetric vector fields of type 1 are given by (4.4) and (4.6), and the locally symmetric vector fields of type 2 are given by

$$V_\mu(C) = N(C)\,\xi(C)_{,\mu},$$

where $C = (C_{(1)}, C_{(2)}, C_{(3)}, C_{(4)})$ are four arbitrary real numbers, $\xi(C)$ is a scalar defined in terms of the scalars (4.7) as

$$\xi(C) = \sum_{a=1}^{4} C_{(a)} \xi^{(a)},$$

and $N(C)$ is a normalizing factor (to make $V_\mu(C)$ a unit vector). The manifold admits no other vector fields with second order local symmetry.

The situation in the case of negative constant curvature can by analyzed, *mutatis mutandis*, by the same method, and we quote only the final result.

Theorem 4.4: A three-dimensional Riemannian manifold of negative constant curvature $-\beta^2, \beta > 0$, admits no locally symmetric vector fields of type 1, and admits a four-parameter congruence of locally symmetric vector fields of type 2. A local chart of coordinates $(x^1, x^2, x^3) = (x, y, z)$ exists, in terms of which the metric tensor is given by:

$$ds^2 = \cosh^2 z\, dx^2 + \sinh^2 z\, dy^2 + (1/\beta^2) dz^2,$$

and the locally symmetric vector fields of type 2 are given by

$$V_\mu(C) = N(C)\,\eta(C)_{,\mu},$$

where

$$\eta(C) = \sum_{a=1}^{4} C_{(a)} \eta^{(a)},$$

and $\eta^{(a)}$ being the four scalars:

$$\begin{aligned}\eta^{(1)} &= \cosh \beta x \cosh z,\\ \eta^{(2)} &= \sinh \beta x \cosh z,\\ \eta^{(3)} &= \cos \beta y \sinh z,\\ \eta^{(4)} &= \sin \beta y \sinh z\end{aligned}$$

(the rest of the notation is the same as in the last theorem). The manifold admits no other vector fields with second order local symmetry.

For completeness we add the case of locally flat manifold, which was excluded in the preceding theorems. The following result is immediately obtained.

Theorem 4.5: The most general second order locally symmetric vector field admitted by a locally flat three-manifold with metric tensor

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

is of type 2, and is given by:

$$V = N\left[(A + Dx)\frac{\partial}{\partial x} + (B + Dy)\frac{\partial}{\partial y} + (C + Dz)\frac{\partial}{\partial z}\right],$$

where A, B, C and D are arbitrarily real constants, and N is a normalizing factor.

Returning now to the cosmological implications, we see that an isotropic universe allows for the various types of systematic distributions of orientations throughout space, according to the sign of the curvature and the list of vector fields constructed above. The existence of these distinguished but symmetric directions should in principle be detectable. Angles formed between axes of galaxies (and other celestial sources of radiation) and the line of sight are observable, and deviations from pure random distributions of orientation could be measured. Of course, to implement such an undertaking, a sufficient amount of data from distant objects ought to be available, so as to eliminate statistically the impact of local fluctuations.

V. CONCLUSIONS

The relation between locally symmetric vector fields and Ricci principal directions has been analyzed. It has been found that a locally symmetric vector field is a Ricci principal direction if it has second order symmetry, but not necessarily so if it has only first order symmetry. Furthermore, it has been shown that if two first order locally symmetric vectors are admitted by a manifold, they must lie in a Ricci two-space at each point. As a consequence a bound has been set for the number of linearly independent locally symmetric vector fields admitted by a manifold. As for linearly dependent locally symmetric vector fields, a procedure has been established for the generation of a two-parameter congruence of such vectors out of a given pair having second order local symmetry. The case of a three-dimensional Riemannian manifold of constant curvature, which attains special importance in relativistic cosmological models, has been studied in details. A coordinate system adapted to the symmetric vectors was found, and the list of all second order locally symmetric vector fields has been explicitly constructed. Classification of these vector fields according to their type distinguishes between the cases of positive and negative constant curvature. These results, coupled with further astronomical observations of the orientation of matter and radiation fields throughout space, may shed some light on the question of the sign of the curvature of the universe.

ACKNOWLEDGMENT

I wish to thank Professor H. Gauchman for helpful discussions.

[1] A.G. Walker, "Note on locally symmetric vector fields in a Riemannian space," in *Topics in Differential Geometry*, edited by H. Rund and W.F. Forbes (Academic, New York, 1976).

[2] A.G. Walker, Mon. Not. R. Astron. Soc. **100**, 622 (1940).

[3] The sign convention throughout the paper is:

$$(\nabla_\tau \nabla_\lambda - \nabla_\lambda \nabla_\tau)\xi^\mu = R^{\mu}_{\ \nu\lambda\tau}\,\xi^\nu,$$
$$R_{\mu\nu} = R^{\lambda}_{\ \mu\lambda\nu},$$
$$R = g^{\mu\nu}R_{\mu\nu}.$$

[4] H. Gauchman, "On locally symmetric vector fields on Riemannian manifolds," Preprint MATH-204, 1978.

Coordinate free relativity

Paul Carlson
Department of Physics, University of Windsor, Windsor, Ontario N9B 3P4, Canada

(Received 15 February 1979; accepted for publication 8 June 1979)

A technique is presented using orthonormal tetrads which enables efficient algebraic manipulation of Einstein's equations by computer and ease of physical interpretation. The results are applied to the spherically symmetric case and Birkhoff's theorem is proved in the formalism. Several exact solutions such as Tolman's and Schwarzschild's are derived, and isotropic expansion of Pefect fluids considered.

1. INTRODUCTION

The use of tetrads has grown rapidly in the last 20 years in general relativity, and some of the more recent examples are summarized in Refs. 1–7. Some earlier work is quoted in these sources, particularly Ref. 5, and has been of value in furthering our understanding of solutions to Einstein's equations. It is the orthonormal tetrad with metric components $\eta_{ab} = \text{diag}_{ab}\ (1,1,1,-1)$ defined in a local neighborhood of space–time that is of interest here as contrasted to the more common null tetrad. Brans[1] has examined such tetrads for Petrov type I (general) case and looked at the integrability conditions by computer finding certain surprising degeneracies. Ciubotariu[2] has set up an observer field in terms of tetrads for the Schwarzchild metric, and Koppel[3] has examined the gauge of tetrad potentials. Moller[4] claims the tetrad formalism may be of value in avoiding certain coordinate singularities and Asgekar and Date[5] consider using tetrads for investigating a charged fluid. Mitskievic and Nesterov[6] examine the Bel–Robinson superenergy tensor and Hoenselaers[7] uses a triad system in space–times with one Killing vector. Here we present an examination of the spherically symmetric space–time using orthonormal tetrads minimizing the role played by coordinates. The same format has been applied to unidirectional space–times Ref. 8, p. 172, and the stationary infinite cylinder as well as the finite stationary axisymmetric fluid,[9] where the equations can be placed in a complex notation which bears resemblance to some of the equations of Ernst.[10]

2. THE BASIC THEORY

We let the indices $a, b, c, d, \cdots = 1, 2, 3, 4$ refer to the orthonormal tetrad, and $i, j, k, l, \cdots = 1, 2, 3, 4$ refer to some coordinate system, and let $v^i_{(a)}$ be the transformation coefficients. Suppose, denotes frame differentiation and * coordinate differentiation. Let $\mathbf{v}_{(a)}$ denote the ath vector field defined locally on some neighborhood U in space–time which makes up the orthonormal tetrad, $v^i_{(a)}$ being its coordinate components. Then $v^i_{(a)}\, v^j_{(b)}\, g_{ij} = \eta_{ab}$ where $\eta_{ab} = \text{diag}_{ab}\ (1,1,1,-1)$ is the frame component representation of the metric tensor. Since the $\mathbf{v}_{(a)}$ form a basis for the tangent space to space–time evaluated at each $x \in U$ we may write $[\mathbf{v}_{(a)},\mathbf{v}_{(b)}] = T_b{}^c{}_a \mathbf{v}_c$ where the scalars $T_b{}^c{}_a$ are called the *Ricci rotation coefficients*. This Lie bracket relation is equivalent to $v^i_{(a),b} - v^i_{(b),a} = T_a{}^c{}_b v^i_c$ which is just the integrability condition for the coordinate function x^i since $x^i_{,c} = v^i_{(c)}$. In general, the integrability condition for a scalar function φ is $\varphi_{,ab} - \varphi_{,ba} = T_a{}^c{}_b \varphi_{,c}$ or equivalently $\varphi_{*ji} = \varphi_{*ij}$. Of course, $\varphi_{,a} = v^i_{(a)}\varphi_{*i}$ for any scalar function φ.

The Ricci rotation coefficients must satisfy the Jacobi identity $T_{[b}{}^a{}_{c,d]} + T_e{}^a{}_{[d}\, T_b{}^e{}_{c]} = 0$. Since the $T_b{}^a{}_c$ are scalars, they must likewise be integrable, and we refer to the equations $T_b{}^a{}_{c,de} - T_b{}^a{}_{c,ed} = T_d{}^f{}_e T_b{}^a{}_{c,f}$ as the integrability conditions for the Ricci rotation coefficients.

The symmetric metric connection $\{{}_j{}^i{}_k\} = \{{}_k{}^i{}_j\}$ referred to as the Christoffel symbols also is represented in frame components as $\{{}_b{}^a{}_c\}$, although this time $\{{}_b{}^a{}_c\} \neq \{{}_c{}^a{}_b\}$. They are given by $\{{}_b{}^a{}_c\} = \frac{1}{2}(T_b{}^a{}_c - T^a{}_{bc} - T^a{}_{cb})$ where the indices are raised and lowered on the scalars $T_b{}^a{}_c$ using the metric η_{ab} or η^{ab} in the usual way. Frame covariant differentiation is then analogously done as $A^a{}_{b;c} = A^a{}_{b,c} + A^d{}_b\{{}_d{}^a{}_c\} - A^a{}_d\{{}_b{}^d{}_c\}$, etc.

The Riemann tensor can then be expressed in terms of linear first derivative terms in $T_b{}^a{}_c$ and quadratic single contractions of $T_b{}^a{}_c$ with itself. This is given in Ref. 8, p. 142. The Ricci tensor, of interest here, is given by $R_{ab} = T_{(a,b)} + T^c{}_{(ab),c} + \frac{1}{2}B_{ab} - \frac{1}{4}E_{ab} + A_{(ab)} + \frac{1}{2}G_{ab}$ where $T_a = T_a{}^b{}_b$, $B_{ab} = T^c{}_{da} T_c{}^d{}_b$, $E_{ab} = T_{ca}{}^d T^c{}_{bd}$, $A_{ab} = T_c T^c{}_{ab}$, $G_{ab} = T_c{}^d{}_a T_d{}^c{}_b$, and the Einstein summation conventions is followed even though these are only scalars. We can easily see $B_{ab} = B_{ba}$, $E_{ab} = E_{ba}$, $G_{ab} = G_{ba}$.

The tetrad vectors $\mathbf{v}_{(a)}$ are also denoted alternately as $\mathbf{v}_{(1)} = \mathbf{r}$, $\mathbf{v}_{(2)} = \mathbf{s}$, $\mathbf{v}_{(3)} = \mathbf{t}$, $\mathbf{v}_{(4)} = \mathbf{u}$. Their components in frame components are $r^a = \delta^a{}_1$, $s^a = \delta^a{}_2$, $t^a = \delta^a{}_3 = t_a$, $u^a = \delta^a{}_4 = -u_a$. It should be mentioned that the Bianchi and Cyclic identities for the Riemann tensor, the Ricci identities, and all local differential geometry properties are consequences of the Jacobi identity and integrability conditions for the Ricci coefficients. Thus we may write out the field equations, impose integrability and in principle, if not in practice, solve the problem or show no solution exists.

3. ENERGY–MOMENTUM TENSOR

We choose $\mathbf{u}$ to be the timelike normalized flow for the material medium. Then following the notation of Ehlers[11] we write $u_{a;b} = \theta_{ab} + \omega_{ab} - \dot{u}_a u_b$. In frame components, assuming A, B, etc. $= 1, 2, 3$, cover the spatial indices only, we have the acceleration $\dot{u}_a = T_a{}^4{}_4$, the deformation rate along the flow $\theta_{AB} = T_{4(AB)}$, and the rotation between mate-

 0022-2488/80/051149-06$01.00

rially nonrotating and spatially nonrotating frames along the flow $\omega_{AB} = \frac{1}{2} T_A{}^4{}_B$.

The general energy-momentum tensor is given by

$$T^{ab} = -(\rho c^2 + \epsilon) u^a u^b - \lambda u^a v^b - \lambda v^a u^b + \widehat{T}^{ab},$$

where $\widehat{T}^{ab} = \widehat{T}^{ba}$, $\widehat{T}^{ab} u_b = 0$, $\rho c^2 + \epsilon > 0$, $v^a v_b = 1$, $v^a u_a = 0$. We interpret ρ as mass density, ϵ as thermal and potential energy density, $\lambda v^a = q^a$ as heat flux vector of magnitude λ and direction v^a, $\lambda \geqslant 0$. We assume the triad **r**, **s**, **t** is chosen so as to diagonalize the stress tensor $\widehat{T}^{ab}$, in this case $\widehat{T}^{ab} = \sigma_1 r^a r^b + \sigma_2 s^a s^b + \sigma_3 t^a t^b$. We call $\sigma_1, \sigma_2, \sigma_3$ the *principal* stresses, and with the spacelike axes oriented along the principal directions of stress we refer to the tetrad as an *adapted frame component system*. The energy domination condition of Hawking,[12] that $0 > T^{ab} \tilde{u}_a \tilde{u}_b$ for all timelike $\tilde{u}^a$ means $\rho c^2 + \epsilon \geqslant 2\lambda + \sigma_{\max}$ where $\sigma_{\max} = \max\{|\sigma_1|, |\sigma_2|, |\sigma_3|\}$. This guarantees that $T^{ab} \tilde{u}_b$ is not spacelike. It is to be noted that our condition that $T^{ab} \tilde{u}_a \tilde{u}_b$ be *negative* is simply a consequence of the choice $\tilde{u}_a \tilde{u}^a < 0$, so T^a_b maps future pointing timelike vectors to future pointing timelike vectors indicating a *positive* energy and mass density, provided $T^a_b \tilde{u}^b$ is timelike. For any material continuum we expect a timelike eigenvector to exist which may or may not be parallel to u^a. For a perfect fluid, $\lambda = 0$ and $\sigma_1 = \sigma_2 = \sigma_3 = -P$, $P \geqslant 0$, $P < \rho c^2 + \epsilon$ will hold.

If $\kappa = 8\pi G / c^4$ we find that Einstein's equations give us

$$\sigma_1 = -\frac{1}{2\kappa}(R_{33} + R_{22} - R_{11} - R_{44}),$$

$$\sigma_2 = -\frac{1}{2\kappa}(R_{33} + R_{11} - R_{22} - R_{44}),$$

$$\sigma_3 = -\frac{1}{2\kappa}(R_{11} + R_{22} - R_{33} - R_{44}),$$

$$\rho c^2 + \epsilon = -\frac{1}{2\kappa}(R_{11} + R_{22} + R_{33} + R_{44}),$$

$$\lambda \alpha^i = \frac{1}{\kappa} R_{i4}, \quad (\alpha^1)^2 + (\alpha^2)^2 + (\alpha^3)^2 = 1,$$

where $v^i = \alpha^1 r^i + \alpha^2 s^i + \alpha^3 t^i$ is the heat flux direction. As well, $R_{AB} = 0$, $A \neq B$ which is the condition of adapted frame components.

The conservation equations $T^{ab}{}_{;b} = 0$ split up into energy and momentum equations in the tetrad notation. They are

$$T_4{}^I{}_J \sigma_I = \epsilon T_4{}^I{}_I + \epsilon_{,4} + (\lambda \alpha^I)_{,I} + \lambda \alpha^I [T_I{}^b{}_b + T_I{}^4{}_4],$$

$$\begin{aligned}(\rho c^2 + \epsilon) T_A{}^4{}_4 = \sigma_{A,A} &+ \sigma_A T_A{}^b{}_b - \sigma_I T_A{}^I{}_I \\ &- [\lambda (T_A{}^4{}_B + T_4{}^B{}_A) \alpha^B \\ &+ (\lambda_{,4} + \lambda T_4{}^I{}_I) \alpha^A \\ &+ \lambda \alpha^A{}_{,4}].\end{aligned}$$

A few comments are in order here. The first equation is a scalar equation and the second holds for $A = 1, 2, 3$, while all other indices have sums, lower case from 1 to 4, upper case from 1 to 3. In the first equation, ϵ can be replaced by $\rho c^2 + \epsilon$ since $(\rho u^a)_{;a} = 0$. Of course, $T^{ab}{}_{;b} = 0$ is a consequence of the contracted Bianchi identity and, therefore, the equations above will be equivalent to integrability conditions for certain of the Ricci coefficients. We will see this in examples to follow. More details are given in Ref. 8, p. 159.

4. SPHERICALLY SYMMETRIC SPACE–TIME

We shall illustrate this by examining the metric

$$\begin{aligned}ds^2 = &-A(r,t) dt^2 + B(r,t) dr^2 + C(r,t) \\ &\times (d\theta^2 + \sin^2\theta d\varphi^2).\end{aligned}$$

Since this is diagonal, we simply rescale the coordinates to obtain the desired frame. Thus

$$v^4_{(4)} = A^{-\frac{1}{2}}, \quad v^1_{(1)} = B^{-\frac{1}{2}}, \quad v^2_{(2)} = C^{-\frac{1}{2}},$$

$$v^3_{(3)} = C^{-\frac{1}{2}} \csc\theta,$$

and

$$r_{,1} = v^1_{(1)}, \quad \theta_{,2} = v^2_{(2)}, \quad t_{,4} = v^4_{(4)}.$$

Using

$$v^i_{(a),b} - v^i_{(b),a} = T_a{}^c{}_b v^i{}_{(c)},$$

we obtain

$$\begin{aligned}T_b{}^a{}_c = g t^a t_{[b} r_{c]} &+ g s^a s_{[b} r_{c]} + d s^a s_{[b} u_{c]} \\ &+ d t^a t_{[b} u_{c]} + k u^a u_{[b} r_{c]} + y r^a r_{[b} u_{c]} + h t^a t_{[b} s_{c]},\end{aligned} \tag{4.1}$$

where

$$g = 2T_3{}^3{}_1 = 2T_2{}^2{}_1 = -\frac{1}{C\sqrt{B}} \frac{\partial C}{\partial r},$$

$$d = 2T_4{}^2{}_2 = 2T_4{}^3{}_3 = \frac{1}{C\sqrt{A}} \frac{\partial C}{\partial t},$$

$$k = 2T_1{}^4{}_4 = \frac{1}{A\sqrt{B}} \frac{\partial A}{\partial r},$$

$$y = 2T_4{}^1{}_1 = \frac{1}{B\sqrt{A}} \frac{\partial B}{\partial t},$$

$$h = 2T_3{}^3{}_2 = -2\cot\theta / \sqrt{C}.$$

In (4.1) we see that h has one-, two-, and four-derivatives, while all other functions g, d, k, y have only one- and four-derivatives. We write $g_{,1}$ as g_1, $d_{,4}$ as d_4, $h_{,2}$ as h_2, etc.

We now start with (4.1) and the above nonzero derivatives, completely ignoring the coordinates and the functions A, B, and C. The starting point now is the Ricci coefficients (4.1) and we simply ignore the line element ds^2. The kinematics and physics can be completely described without coordinates and without vector or tensor fields, simply with scalars and the comma derivative. All physical quantities are in canonical (tetrad) form, not to be interpreted through obscure coordinates, and describe the matter in its infinitesimal rest frame.

5. THE TETRAD SOLUTION

We apply the general method by imposing Jacobi identity and integrability conditions to (4.1) and working out the principal stresses, heat flux, and mass and energy density. In order to fully determine the tetrad it may be necessary also to impose some form of gauge condition. The most convenient one the author has found so far is the differential gauge condition $T_b{}^a{}_{c,a} = 0$. One can show (Ref. 8, p. 149) that the formula for the exterior derivative $df_a|_b$ of a one form f_a in tetrad components is

$$df_a|_b = f_{a,b} - f_{b,a} - f_c T_a{}^c{}_b. \tag{5.1}$$

If $T_a = T_a{}^b{}_b$ are considered to be the tetrad components in this particular frame component system of some 1-form, then the contracted Jacobi identity can be used to show that $dT_a|_b = T_a{}^c{}_{b,c}$ so the gauge condition is equivalent to the 1-form T_a being closed, i.e., $dT_a|_b = 0$. Since a closed form is locally exact and the tetrad fields are only defined on neighborhoods, this means that a potential $\widehat{\Omega}$ exists with $\widehat{\Omega}_{,a} = \widehat{\Omega}_a = T_a$, i.e., the integrability condition is satisfied for $\widehat{\Omega}$, which we call the *basic gauge potential*. For the spherically symmetric case this is closely related to, but not identical with the gauge potential of Sec. 7.

In Ref. 9 Carlson shows that for perfect stationary axisymmetric fluids, the differential gauge condition can always be imposed, and although not unique, has a special unique integrated gauge determined by a relationship among Weyl tensor components in the case of Kerr and Wahlquist[13] solutions. Here it leads to the usual prolate spheroidal coordinates from the tetrad frame. Alternate gauges are considered as well, and a geodesic radial gauge is used to form the central leading terms that exhibit Newtonian behavior. The differential gauge condition, however, gives the simplest expression for known exact solutions in its unique integrated form.

Likewise we see in Sec. 7 that for spherical symmetry the condition $T_b{}^a{}_{c,a} = 0$ is equivalent to the flow being canonically oriented. This is not so for all solutions[14] and the orientations defined need not be unique, but are a restricted class. Thus this gauge condition is not always satisfied for adapted frame components, and others may be considered.[3] The conformally flat metric with $T_b{}^a{}_c = \frac{2}{3}\widehat{\Omega}_{,[b}\delta^a_{c]}$ does satisfy this gauge. Now we examine the tetrad formalism.

The calculations are straightforward but lengthy, however they are much more suited to algebraic computer solution than are the usual coordinate forms, where derivatives are required to second order and large messy denominators appear. The tetrad solutions in coordinate free relativity have been put on an algebraic computer program by the author using REDUCE 2[15] and are very much superior in speed and compactness than corresponding coordinate solutions.

In the spherically symmetric case, using (4.1) we have the integrability conditions

$$\varphi_{41} - \varphi_{14} = y\varphi_1/2 - k\varphi_4/2$$

for all functions with one- and four-derivatives only, i.e., g, d, k, and y. For h we have

$$h_{41} - h_{14} = \frac{y}{2}h_1 - \frac{k}{2}h_4, \tag{5.2}$$

$$h_{42} - h_{24} = \frac{d}{2}h_2, \tag{5.3}$$

$$h_{12} - h_{21} = -\frac{g}{2}h_2. \tag{5.4}$$

The Jacobi identity gives us

$$g_4 + d_1 = -kd/2 - gy/2, \tag{5.5}$$

$$h_1 = gh/2, \quad h_4 = -dh/2. \tag{5.6}$$

Now (5.6) and (5.5) imply (5.2) and using (5.3) and (5.4) we obtain

$$h_{21} = gh_2, \quad h_{24} = -dh_2. \tag{5.7}$$

Working out the principal stresses, etc., we obtain

$$\begin{aligned}
\sigma_1 &= -\frac{1}{2\kappa}\left(-2d_4 - h_2 + \frac{h^2}{2} + \frac{g^2}{2} - \frac{3d^2}{2} - gk\right),\\
\sigma_2 = \sigma_3 &= -\frac{1}{2\kappa}\Big(k_1 - g_1 - y_4 - d_4\\
&\qquad + \frac{g^2}{2} + \frac{k^2}{2} - \frac{y^2}{2} - \frac{d^2}{2} - \frac{yd}{2} - \frac{kg}{2}\Big),\\
\rho c^2 + \epsilon &= -\frac{1}{2\kappa}\Big(-2g_1 - h_2 + \frac{h^2}{2} + \frac{3g^2}{2}\\
&\qquad - \frac{d^2}{2} - dy\Big),\\
\lambda &= -\frac{1}{2\kappa}\left(g_4 - d_1 + gd - \frac{yg}{2} + \frac{kd}{2}\right), \quad \alpha^i = r^i,\\
&= -\frac{1}{2\kappa}(-2d_1 + gd - gy) = -\frac{1}{2\kappa}(2g_4 + kd + gd).
\end{aligned} \tag{5.8}$$

The conditions $(\sigma_1)_{,2} = 0$ and $(\rho c^2 + \epsilon)_{,2} = 0$ of spherical symmetry give us

$$h_{22} = hh_2 \tag{5.9}$$

and this completes the solution for h, guaranteeing h and h_2 are integrable.

The conservation equations of energy and momentum are

$$\begin{aligned}
d\sigma_2 + \frac{y\sigma_1}{2} &= (\rho c^2 + \epsilon)\left(d + \frac{y}{2}\right) + (\rho c^2 + \epsilon)_{,4}\\
&\quad + \lambda_{,1} + \lambda(k - g),\\
(\rho c^2 + \epsilon)\frac{k}{2} &= \sigma_{1,1} + \sigma_1\left(\frac{k}{2} - g\right) + g\sigma_2\\
&\quad - \lambda_{,4} - \lambda(y + d),
\end{aligned} \tag{5.10}$$

and these give, respectively, the integrability conditions for g and d. Of course, kinematically $\omega_{ab} = 0$, $\dot{u}^a = (k/2)r^a$ and $\theta_{ab} = (d/2)(s_a s_b + t_a t_b) + (y/2)r_a r_b$. Thus if $y = d$ the deformation is isotropic or shear free. In the coordinate formulation (Sec. 4) this means B/C is independent of t.

If we put $y = d$ in (5.8) we see, among other things, that $\lambda = d_1/\kappa$. Thus, in an isotropic expansion of a spherically symmetric space–time the heat flux, which is radially directed, is proportional to the radial derivative of the expansion rate. Also, the frame rotation tensor $\Lambda_{AB} = T^4_{[AB]}$ is zero whether or not $y = d$. Λ_{AB} measures the rotation rate between the tetrad frame and the materially nonrotating frame.

6. BIRKHOFF'S THEOREM

In this section we look for vacuum solutions, putting the quantities in (5.8) to zero. We are free to orient the flow vector as we please, and choose it so as to make the scalar invariant $\alpha = C_{4141}$ have zero 4-derivative. For a spherically symmetric space–time, the Weyl tensor is pure electric and type D (or zero) and in the vacuum case $\alpha = -h_2/2$

$+ h^2/2 + g^2/4 - d^2/4$. Since $\alpha \neq 0$ for vacuum, and $\alpha_{,4} = -(3d/2)\alpha$, the orientation condition is $d = 0$. Hence $\lambda = 0 \rightarrow g_4 = 0$ and (5.5) implies $gy = 0$. But since $g \neq 0$ (from g_1 equation) then $y = 0$.

The vacuum Eqs. (5.8) then become

$$g_1 = -\frac{h_2}{2} + \frac{h^2}{4} + \frac{3g^2}{4}, \tag{6.1}$$

$$0 = d_4 = -\frac{h_2}{2} + \frac{h^2}{4} + \frac{g^2}{4} - \frac{gk}{2}, \tag{6.2}$$

$$k_1 = -\frac{3h_2}{2} + \frac{3h^2}{4} + \frac{3g^2}{4} - \frac{k^2}{2} - \frac{gk}{2}. \tag{6.3}$$

Now g is trivially integrable from (5.10) and (6.2) implies $k_4 = 0$ which together with (6.3) implies k is integrable. Thus the solution is time independent and without rotation and hence is static.

Even if the flow vector is not oriented so that we have this obvious time independence, and even if $y \neq d$, in the vacuum case we have a number of interesting relations. As well as α, the quantity $h_2 - h^2/2 = 2/C(r,t)$ (see Sec. 4) is also an invariant and $\alpha = -(m/2^{1/2})(h_2 - h^2/2)^{3/2}$ where m is a constant, identical with the Schwarzschild mass. Furthermore $\alpha = dy/2 + gk/2$

and so

$$d_4 = -\frac{d}{2}(d-y), \quad d_1 = \frac{g}{2}(d-y),$$

$$g_4 = -\frac{d}{2}(k+g), \quad g_1 = \frac{g}{2}(k+g),$$

and $k_1 - y_4 = dy + gk - k^2/2 + y^2/2$. These relations are useful for placing the Schwarzschild solution in other than the canonical tetrad frame for matching on the boundary to various interior solutions.

7. THE SCHWARZSCHILD SOLUTION

In this section we shall solve for the above vacuum solution in the unique canonically oriented isotropic frame, which makes all four-derivatives zero. We say the flow vector in a spherically symmetric space-time is *canonically oriented* if the gauge condition $y_1 = k_4$ is satisfied. This means there exists a gauge potential Ω with $\Omega_1 = k, \Omega_4 = y$, since Ω is integrable. In terms of the coordinate functions, A, B, and C, this means that $A/B = F(t)G(r)$, a factorization, since $\Omega = \ln A + f(t) = \ln B + g(r)$. We can then readjust coordinates $r \rightarrow f_1(r)$, $t \rightarrow f_2(t)$ so as to make $A = B$. Thus the flow is oriented so as to make the conformal flatness of the two-space part of the metric in Sec. 4 obvious. It is a restricted class of orientations, not unique.

If we impose this condition, plus $y = d$ in the vacuum case, we obtain $y = d = 0$ and all four-derivatives zero as in the first part of Sec. 6. The equations we obtain are

$$h_2 = \frac{h^2}{2} + \frac{g^2}{2} - gk, \quad h_1 = \frac{gh}{2}, \tag{7.1}$$

$$g_1 = \frac{g^2}{2} + \frac{gk}{2}, \quad k_1 = -\frac{k^2}{2} + gk. \tag{7.2}$$

Because $T_a{}^1{}_b = 0$ (since $y = 0$) we see that the one-form r_a has zero exterior derivative, i.e., $dr_{a|b} = 0$ from (5.1). Thus there exists a local metric radial coordinate x with $x_{,a} = r_a$ and $\varphi_1 = d\varphi/dx$ for any φ. The functions g and k depend only on x so the pair (7.2) can be solved.

To solve (7.2) we change variables to separate putting $U = kg$ and $V = g^2/k$, so $dU/dx = \frac{3}{2}kg^2$ and $dV/dx = \frac{3}{2}g^2$. Hence $dU/dV = k = U^{2/3}V^{-1/3}$ so integration gives $2U^{1/3} = V^{2/3} + c$ for constant $c = -(2/m)^{2/3}$, where m is Schwarzschild mass. We take

$$\frac{dU}{dx} = \frac{3}{2}U^{4/3}\left(2U^{1/3} + \left(\frac{2}{m}\right)^{2/3}\right)^{1/2},$$

and substituting $U = -4m/r^3$ we obtain

$$\frac{dr}{dx} = \left(1 - \frac{2m}{r}\right)^{1/2},$$

and

$$x = m\cosh^{-1}(r/m - 1) + (r^2 - 2mr)^{1/2}.$$

We call r the Schwarzschild coordinate. Now

$$V = (2/m)(1 - 2m/r)^{3/2},$$

and so

$$k = (2m/r^2)(1 - 2m/r)^{-1/2},$$

and

$$g = -(2/r)(1 - 2m/r)^{1/2}.$$

Now

$$h_2 - h^2/2 = g^2/2 - gk = 2/r^2,$$

so

$$h = -2(\cot\theta/r),$$

where $\theta_{,b} = (1/r)s_b$. Going back to A, B, C in Sec. 4 we find

$$ds^2 = -(1 - 2m/r)dt^2 + (1 - 2m/r)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2).$$

The beauty of this is that it is not necessary to ever go back so far as to work out the metric in coordinates, and physical behavior can be inferred by simple equations such as (7.2) with their solutions.

8. THE TOLMAN SOLUTION

It is possible to show in this tetrad formalism that for spherically symmetric dust, the flow is canonically oriented (i.e., $y_1 = k_4$) if and only if the deformation is isotropic $y = d$. Under these conditions we obtain the constant density conformally flat solutions of Tolman, which are the zero pressure special cases of the Robertson–Walker perfect fluid solutions. They are the only members of this class of solutions which can be matched to an exterior Schwarzschild solution using Lichnerowicz boundary conditions.[16]

Putting $y = d$ and $\sigma_1 = \sigma_2 = \sigma_3 = -P = 0$, $\lambda = 0$, the conservation Eqs. (5.10) give the relations $(\rho c^2 + \epsilon)_{,4} = -(3d/2)(\rho c^2 + \epsilon) = -\theta(\rho c^2 + \epsilon)$, $\theta = \theta^a_{\ a}$, and $k = 0$ (i.e., zero acceleration in rest frame). These automatically imply g and d are integrable. For convenience we put $F = 2\kappa(\rho c^2 + \epsilon)$ to represent matter and energy density. Now $k = 0 \rightarrow k_4 = 0 \rightarrow d_1 = 0$ (by Gauge) $\rightarrow \lambda = 0$. Also

$k_1 = 0$ gives

$$F = 3h_2 - \frac{3h^2}{2} - \frac{3g^2}{2} + \frac{3d^2}{2},$$

so $\alpha = 0$ (conformally flat). Also

$$g_1 = h_2 - \frac{h^2}{2}, \quad g_4 = -\frac{dg}{2},$$

$$d_4 = -\frac{h_2}{2} + \frac{h^2}{4} + \frac{g^2}{4} - \frac{3d^2}{4}.$$

We have, of course,

$$h_1 = \frac{gh}{2}, \quad h_4 = -\frac{dh}{2},$$

$$h_{21} = gh_2, \quad h_{24} = -dh_2, \quad h_{22} = hh_2.$$

We can check that $F_{,4} = -(3/2)dF$ as expected, but also that $F_{,1} = 0$, so the density is a constant radially. Thus this would make a reasonable spherical cosmological model ($d_1 = 0, F_1 = 0$). We get the pair of equations

$$d_4 = -\frac{F}{6} - \frac{d^2}{2}, \quad F_4 = -\frac{3}{2}dF, \tag{8.1}$$

which contain all the physics of the problem. Since $T_a{}^4{}_b = 0$, i.e., $k = 0$ we have $du_a|_b = 0$ so there exists a local metric time coordinate τ with $\tau_{,a} = -u_a = \delta^4{}_a$, $\varphi_4 = d\varphi/d\tau$ for any φ. Thus d and F are functions only of τ.

To separate variables in (8.1), introduce $U = d^3/F$ so

$$\frac{dU}{d\tau} = -\frac{d^2}{2} \quad \text{and} \quad \frac{dF}{d\tau} = -\frac{3}{2}dF.$$

Dividing and integrating we obtain $F = (3/2)d^2 + CF^{2/3}$ for some constant C. If $C > 0$ we have an elliptic solution with ultimate gravitational collapse in finite time, $C = 0$ parabolic, and $C < 0$ hyperbolic, where the universe expands to infinity with energy to spare. The complete Robertson–Walker solutions can be handled very simply in the tetrad notation (including pressure) by the omnidirectional and unidirectional conditions (Ref. 8, pp. 167 and 172), though we will not give the solutions here.

9. SCHWARZSCHILD'S INTERIOR SOLUTION

In this section we shall consider static solutions in a frame with $y = d = 0$, and all four-derivatives zero. For a perfect fluid we put $\sigma_1 = \sigma_2 = \sigma_3 = -P = -f/2\kappa$ and $\rho c^2 + \epsilon = F/2\kappa$, $\lambda = 0$. Hence we obtain, writing prime for one-derivative,

$$f' = -\frac{k}{2}(f + F),$$

$$k' = \frac{3}{2}f + \frac{F}{2} - \frac{k^2}{2} + gk, \tag{9.1}$$

$$g' = \frac{f}{2} + \frac{F}{2} + \frac{g^2}{2} + \frac{gk}{2}.$$

This system of three equations can be reduced to two if desired. We can see that $(\ln|f + gk - \frac{1}{2}g^2|)' = g = -2(\ln r)'$, where the radial coordinate r replaces the metric coordinate x so that $r' = dr/dx = -gr/2$, $h = -(2/r)\cot\theta$, $\theta_2 = 1/r$. By appropriate boundary conditions at the center of the sphere (Ref. 8, p. 236) we obtain $f = -gk + \frac{1}{2}g^2 - 2/r^2$ and eliminating f, (9.1) becomes a set of two equations for k and g derivatives with respect to r.

Let us look for conformally flat interiors. Now $\alpha = C_{4141} = f/2 + F/6 + gk/2$, and putting $\alpha = 0$ we obtain $F = -3f - 3gk \rightarrow F' = 0$ so F (the density) is constant. We eliminate f to obtain

$$k' = -\frac{k^2}{2} - \frac{gk}{2}, \quad g' = \frac{F}{3} + \frac{g^2}{2}. \tag{9.2}$$

The boundary conditions on g require g to become infinite and behave like $-2/x$ (or $-2/r$) near $x = 0$ (Ref. 8, p. 236). Thus integrating gives $g = -(\frac{2}{3}F)^{1/2}\cot[(\frac{2}{3}F)^{1/2}(x/2)]$. From $r' = -gr/2$ we get $r = (2/(\frac{2}{3}F)^{1/2})\sin[(\frac{2}{3}F)^{1/2}x/2]$, since $r \sim x$ near $x = 0$.

Putting $U = kg$ we obtain

$$\frac{dU}{(F/3)U - U^2/2} = \frac{dx}{g(x)},$$

and integrating with $k(0) = 0$ we obtain k and f from $f = -F/3 - gk$. In terms of r this gives

$$g = -\frac{2}{r}\left(1 - F\frac{r^2}{6}\right)^{1/2},$$

$$k = \frac{aFr}{3[1 - a(1 - Fr^2/6)^{1/2}]}, \tag{9.3}$$

$$f = \frac{-F + 3Fa(1 - Fr^2/6)^{1/2}}{3[1 - a(1 - Fr^2/6)^{1/2}]},$$

where

$$a = \frac{f_0 + F/3}{f_0 + F},$$

and where $f_0 = f|_{x=0}$ (central pressure). Now $f = 0$ at $r = r_0$, the point of application of Lichnerowicz boundary conditions, and $r_0^2 = 6/F - 2/(3a^2F)$, $\frac{2}{3} > a > \frac{1}{3}$. Matching conditions give Schwarzschild mass $m = Fr_0^3/12 = \frac{4}{3}\pi r_0^3(\rho c^2 + \epsilon)(G/c^4)$ using $F = 2\kappa(\rho c^2 + \epsilon)$ and $\kappa = 8\pi G/c^4$. Now

$$g = -\frac{1}{C\sqrt{B}}\frac{\partial C}{\partial r}, \quad k = \frac{1}{A\sqrt{B}}\frac{\partial A}{\partial r},$$

$$h = -\frac{2\cot\theta}{\sqrt{C}},$$

so $C = r^2$, $B = (1 - Fr^2/6)^{-1}$, $A = [1 - a(1 - Fr^2/6)^{1/2}]^2$. Thus we obtain the metric as

$$ds^2 = -[1 - a(1 - Fr^2/6)^{1/2}]^2dt^2 + \frac{dr^2}{1 - Fr^2/6} + r^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{9.4}$$

The junction requirements are imposed across the $f = 0$ surface so that g, k are continuous and this implies h is continuous.

10. ISOTROPIC EXPANSION OF PERFECT FLUIDS

A general spherically symmetric time-dependent perfect fluid has $\lambda = 0$, $\sigma_1 = \sigma_2 = \sigma_3 = -P = -f/2\kappa$ and $\rho c^2 + \epsilon = F/2\kappa$ so from (5.8),

$$d_4 = -\frac{f}{2} - \frac{h_2}{2} + \frac{h^2}{4} + \frac{g^2}{4} - \frac{3d^2}{4} - \frac{gk}{2},$$

$$d_1 = \frac{g}{2}(d-y), \quad g_4 = -\frac{d}{2}(k+g),$$

$$g_1 = \frac{F}{2} - \frac{h_2}{2} + \frac{h^2}{4} + \frac{3g^2}{4} - \frac{d^2}{4} - \frac{dy}{2}, \qquad (10.1)$$

$$k_1 - y_4 = \frac{f}{2} + \frac{F}{2} - h_2 + \frac{h^2}{2} + \frac{g^2}{2} - \frac{d^2}{2} - \frac{k^2}{2} + \frac{y^2}{2},$$

$$\left(h_2 - \frac{h^2}{2}\right)_1 = g\left(h_2 - \frac{h^2}{2}\right),$$

$$\left(h_2 - \frac{h^2}{2}\right)_4 = -d\left(h_2 - \frac{h^2}{2}\right).$$

Integrability for g and d give conservation equations

$$F_4 = -\left(d + \frac{y}{2}\right)(f+F), \quad f_1 = -\frac{k}{2}(f+F). \qquad (10.2)$$

The Weyl tensor component $\alpha = C_{4141}$ is given by

$$\alpha = \frac{F}{6} - \frac{h_2}{2} + \frac{h^2}{4} + \frac{g^2}{4} - \frac{d^2}{4}, \qquad (10.3)$$

and satisfies

$$\alpha_{,4} = -\frac{3d\alpha}{2} + \frac{(d-y)}{12}(f+F),$$

$$\alpha_{,1} = \frac{3g\alpha}{2} + \frac{F_1}{6}. \qquad (10.4)$$

The objective in solving (10.1) is to find explicit expressions for the unknown derivatives such as to satisfy the general integrability conditions $\varphi_{14} - \varphi_{41} = (y/2)\varphi_1 - (k/2)\varphi_4$ for all the Ricci coefficients. This permits the equations to be integrated.

To do this, conditions may be imposed. If we require isotropic expansion $y = d$, dynamic motion $d \neq 0$ and canonical flow (i.e., Gauge condition) then $\alpha = 0$ and the space–time is conformally flat, so by (10.4) $F_1 = 0$, i.e., the density is radially constant. This is seen by imposing integrability for k. If we require, on the other hand, the existence of a one parameter equation of state $f = f(F), f' = df/dF$ so $f_1 = f'F_1, f_4 = f'F_4, f'' = d^2f/dF^2$, etc., then from (10.2),

$$f_4 = -f'\left(d + \frac{y}{2}\right)(f+F) \text{ and } F_1 = -\frac{k}{2f'}(f+F). \qquad (10.5)$$

Imposing integrability on f (equivalently on F) gives

$$0 = \frac{k}{2}\left(d + \frac{y}{2}\right)\left(\frac{f''}{f'}(f+F) - f'\right) - \frac{f'}{2}(gd - gy + y_1) + \frac{k_4}{2} + \frac{ky}{4}.$$

This equation involving y_1 and k_4 can be combined with the one in (10.1) for $k_1 - y_4$ imposing integrability on k and y and solving, in general quite difficult to do. If we impose the added condition $y = d$ (isotropic expansion) then $k_4 = -kd\psi/2$ where $\psi = 1 + 3[(f''/f')(f+F) - f']$. Also $k_1 = 3\alpha - (k/2)(g+k)$, so integrability for k gives

$$0 = -\tfrac{3}{2}(\psi - 2)\alpha + k^2\left(\frac{\psi'}{4f'}(f+F) - \frac{\psi}{2}\right)$$

for $d \neq 0$, and this means α/k^2 is a function of F. Imposing this condition, i.e., $(\alpha/k^2)_1 F_4 = (\alpha/k^2)_4 F_1$ we obtain

$$\frac{5}{2}\frac{g}{k} - \frac{f+F}{12\alpha f'}$$

is a function of F, or equivalently $[g + (\ln\alpha)_{,1}]/k$ is a function of F. Repeated applications give new successive constraint conditions, the apparent implication being that no solution is possible unless $\alpha = 0$. This is the content of a theorem proved by Mansouri[17] and Glass.[18] It states that for isotropic expansion $y = d$ subject to dynamic motion $d \neq 0$ under a barotropic equation of state, the only solution matching an exterior Schwarzschild is the Tolman dust solution.

It is possible to obtain solutions for $y = d$ (isotropic expansion) that is $B(r,t)/C(r,t)$ is independent of t, and the coordinate r may be chosen so that this ratio is $1/r^2$. A summary of such solutions is given by Chakravarty *et al.*[14] It is to be noted that for these the flow is not canonically oriented nor are they subject to a barotropic equation of state.

We can see that if $y = d$ and $F_1 = 0$ then (10.4) and (10.1) imply $\alpha/(h_2 - h^2/2)^{3/2}$ is a constant. This is related to the "gravitational field energy" discussed in Ref. 18.

ACKNOWLEDGMENTS

I am grateful to Professor E.N. Glass for helpful comments in the preparation of this work, and to Professor H.P. Künzle at the University of Alberta.

[1] C.H. Brans, J. Math. Phys. **18**, 1378 (1977).
[2] C.D. Ciubotariu, Bul Inst. Politeh. Iasi. Sec. I (Rumania) **23**, 77 (1977).
[3] A. Koppel, Toim. Eesti, N.S.V. Tead. Akad.Fuus. Mat. (U.S.S.R.) **27**, 123 (1978).
[4] C. Moller, Mat.-Fys. Medd. Dan. Vidensk Selsk (Denmark) **39**, 1 (1978).
[5] G.G. Asgekar and T.H. Date, G.R.G. **9**, 175 (1978).
[6] N.V. Mistkievic and A.I. Nesterov, Acta Phys. Polon. **B8**, 423 (1977).
[7] C. Hoenselaers, Prog. Theor. Phys. (Japan) **57**, 1223 (1977).
[8] P. Carlson, Ph.D. Thesis (Univ. of Alta., Edmonton, Canada, 1978).
[9] P. Carlson, Gen. Rel. Grav., to appear.
[10] F.J. Ernst, Phys. Rev. **167**, 1175 (1968).
[11] J. Ehlers, Abhan. Math.-Naturwiss. Klasse Akad. Wiss. Lit. Nr. **11**, 793 (1961).
[12] S. Hawking, Commun. Math. Phys. **18**, 301 (1970).
[13] H. Wahlquist, Phys. Rev. **172**, 1291 (1968).
[14] N. Chakravarty, S.B. Dutta Choudhury, and A. Banerjee, Australian J. Phys. **29**, 113 (1976).
[15] A.C. Hearn, REDUCE 2 Manual, Univ. of Utah, Salt Lake (1975).
[16] A. Lichnerowicz, Théories Relativistes de la Gravitation et de l'Électromagnetisme, Masson et Cie, Paris (1955).
[17] R. Mansouri, Ann. Inst. Henri Poincaré **27**, 175 (1977).
[18] E.N. Glass, "Shear-Free Gravitational Collapse," Phys. Dept., Univ. of Windsor, Ontario, Canada (1978).

Aesthetic field theory; the problem of spatial inversions

M. Muraskin
Physics Department, University of North Dakota, Grand Forks, N. Dak. 58201

(Received 19 October 1979; accepted for publication 25 January 1980)

In a previous paper we studied a set of origin point data that was invariant under three-dimensional rotations. This data, however, was not, in general, invariant under spatial inversions. We can obtain another set of data from the original data by performing a spatial inversion. Both sets of data can be incorporated into aesthetic field theory without compromising the basic ideas of the theory, by introducing complex fields. It is found that this combined set of data leads to computer solutions only slightly more complex than the corresponding real case.

I. INTRODUCTION

Field equations have been formulated based on mathematically aesthetic ideas.[1,2] All orders of derivatives of the fields as well as all tensors are treated in a uniform manner with respect to their change.

The aesthetic field equations yield different sorts of solutions depending on their origin point data. We have obtained a two-particle solution and have studied a scattering of the two particles.[3] We have found sine, cosine solutions analytically.[4] These latter solutions are exact for certain origin point data. Other choices of origin point data lead to constant Γ^i_{jk}. Some other solutions can be shown to be unbounded.[5] Still other solutions are more difficult to analyze on the computer. Numbers get very large so it is not clear whether singularities are developing or if we are just dealing with large numbers.

At any rate, it is clear that solutions of the aesthetic field equations depend on the origin point data. The question then is what sort of criterion should be used in order to obtain a satisfactory set of data.

II. CHOOSING ORIGIN POINT DATA

We have taken two points of view in the past:

(1) We have initiated searches for a general set of $\Gamma^\alpha_{\beta\gamma}$ that obeys no tensor restrictions other than integrability. It is not easy to prove that a $\Gamma^\alpha_{\beta\gamma}$ does not obey any set of tensor relations, since there are an infinite number of tensor restrictions that can be obtained by taking products and contractions of expressions involving the fields. To implement this program we have looked for $\Gamma^\alpha_{\beta\gamma}$ with no restrictions in mind at the outset. All manner of solutions to the integrability equations have been pursued.

(2) Our second viewpoint is motivated by the notion that not only should the field equations be formulated according to mathematically aesthetic ideas, but the origin point data should also be subject to aesthetic considerations.

Basically, the idea we have pursued in this context is that there should exist coordinate transformations that do not affect the nature of the solution.

An example of this approach is given in Ref. 2. For nonzero $\Gamma^\alpha_{\beta\gamma}$ we take

$$\begin{aligned}&\Gamma^1_{10}=\Gamma^2_{20}=\Gamma^3_{30}=\Gamma^0_{00}=\Gamma^1_{01}=\Gamma^2_{02}=\Gamma^3_{03}=A,\\&\Gamma^0_{11}=\Gamma^0_{22}=\Gamma^0_{33}=-B,\\&\Gamma^2_{13}=\Gamma^3_{21}=\Gamma^1_{32}=-\Gamma^1_{23}=-\Gamma^3_{12}=-\Gamma^2_{31}=C.\end{aligned}\tag{1}$$

For any choice of A, B, and C the integrability equations

$$A^i_{jkpl}\equiv\Gamma^i_{mk}R^m_{jpl}+\Gamma^i_{jm}R^m_{kpl}-\Gamma^m_{jk}R^i_{mpl}=0\tag{2}$$

are satisfied. For $A=B=C$ we find

$$R^i_{jkl}=0.\tag{3}$$

We have investigated the data for various choices of A, B, and C and found similar results. (This was true as long as A and B had the same sign. If A and B have opposite signs we have not been able, in the cases observed, to observe boundedness in axis runs.) We have chosen, A, B, and C to have same order of magnitude so as not to lead to any more errors than is necessary. Some investigations of a possible pronounced magnitude effect were looked for, but not found as yet.[6] The set of data (1) leads to maps a bit less complex than those found in Ref. 3. The x, y planar maps show only 2 planar maximum and minimum while those in Ref. 3 show 3 and 4.

What makes the data (1) so interesting from a group theory point of view is that this $\Gamma^\alpha_{\beta\gamma}$ is unchanged by a three-dimensional rotation. Thus we see that such a symmetry requirement is a powerful restriction on the nature of the origin point data.

We have also investigated a $\Gamma^\alpha_{\beta\gamma}$ that is unchanged by a Lorentz transformation in Ref. 7. This was a somewhat more unnatural situation since it led to a five-dimensional theory.

A limitation of our results using data (1) was that the fields went to zero far away from the two particles in what appears to be an uninteresting way. A realistic solution would be expected to give many particles and not just two. It is possible that with greater distances from the origin more structure could develop, although we have yet to see any of this.

Our aim at this point is to find more complicated solutions, by one way or another, to the aesthetic field equations, assuming that such complicated solutions to the equations exist.

Writing

$$\Gamma^i_{jk}=e_\alpha{}^i e^\beta{}_j e^\gamma{}_k\Gamma^\alpha_{\beta\gamma},\tag{4}$$

 0022-2488/80/051155-04$01.00

we note that even when $\Gamma^{\alpha}_{\beta\gamma}$ is invariant under three-dimensional rotations Γ^{i}_{jk} is not in general. The symmetry is thus an underlying one. However for various choices of A, B, and C in (1) the character of solutions was similar for different general choices of $e^{\alpha}{}_{i}$ in all the cases we investigated.

III. SPATIAL INVERSIONS

We have concerned ourselves thus far with three-dimensional rotations. What we have not looked into previously is the effect of spatial inversions. We can bring in spatial inversions by using

$$(e^{\alpha}{}_{i}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5}$$

in Eq. (4). We see that $\Gamma^{\alpha}_{\beta\gamma}$ in (1) is not invariant under this transformation. We obtain from (1), using (4) and (5),

$$\begin{aligned} &\Gamma^{1}_{10} = \Gamma^{2}_{20} = \Gamma^{3}_{30} = \Gamma^{0}_{00} = \Gamma^{1}_{01} = \Gamma^{2}_{02} = \Gamma^{3}_{03} = A, \\ &\Gamma^{0}_{11} = \Gamma^{0}_{22} = \Gamma^{0}_{33} = -B, \\ &\Gamma^{2}_{13} = \Gamma^{3}_{21} = \Gamma^{1}_{32} = -\Gamma^{1}_{23} = -\Gamma^{3}_{12} = -\Gamma^{2}_{31} = -C. \end{aligned} \tag{6}$$

Thus, only if we choose $C = 0$ at the outset would the data be invariant under rotations and inversions.

We have investigated $C = 0$ theory and found similar type maps as when $C \neq 0$. The restriction $C = 0$ implies that $g_{\alpha\beta}\Gamma^{\alpha}_{\sigma\gamma}$ has no totally antisymmetric part when

$$(g_{\alpha\beta}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7}$$

$C = 0$ theory means that $\Gamma^{\alpha}_{\beta\gamma}$ has no nonzero components with the indices α, β, and γ all taking on different values. We note that in our sine, cosine solution[4] such components played an important role.

The question then is whether we can satisfy the basic equations of aesthetic field theory and at the same time introduce *both* data (1) and (6) into the theory, such that both sets of data play a role of equal stature.

There are several ways to proceed at this point. What we need is a scheme that allows for some kind of doubling.

We note that in Dirac theory a doubling is necessary in going from the Pauli equation to the Dirac equation in order to take into account the existence of antiparticles. We should then look into various ways that a doubling can be discussed within aesthetic field theory.

Doubling to dimension eight was considered in Ref. 5. There we considered a $4 \oplus 4$ structure. Then we allowed coupling between the subuniverses by means of a general eight-dimensional $e^{\alpha}{}_{i}$. A problem in such a theory is to prove the "integrity" of the subuniverses. We would have to show that the coupling between the subuniverses remains small in some sense so that an observer would find himself restricted to one of the subuniverses. We have yet to be able to prove that this is the case. Also, comparing results in the eight-dimensional theory with a four-dimensional theory used as a control, we did not see significant changes in the nature of the results.

Another possibility is to take a $3 \oplus 3$ theory with a single time coordinate (rather than the two times in eight-dimensional theory). Thus we effectively have a seven-dimensional theory. Our nonzero $\Gamma^{\alpha}_{\beta\gamma}$ was then

$$\begin{aligned} &\Gamma^{1}_{14} = \Gamma^{2}_{24} = \Gamma^{3}_{34} = \Gamma^{4}_{44} = \Gamma^{1}_{41} = \Gamma^{2}_{42} = \Gamma^{3}_{43} = A, \\ &\Gamma^{4}_{11} = \Gamma^{4}_{22} = \Gamma^{4}_{33} = -B, \\ &\Gamma^{2}_{13} = \Gamma^{3}_{21} = \Gamma^{1}_{32} = -\Gamma^{1}_{23} = -\Gamma^{3}_{12} = -\Gamma^{2}_{31} = C, \\ &\Gamma^{5}_{54} = \Gamma^{6}_{64} = \Gamma^{7}_{74} = \Gamma^{5}_{45} = \Gamma^{6}_{46} = \Gamma^{7}_{47} = A, \\ &\Gamma^{4}_{55} = \Gamma^{4}_{66} = \Gamma^{4}_{77} = -B, \\ &\Gamma^{6}_{57} = \Gamma^{7}_{65} = \Gamma^{5}_{76} = -\Gamma^{5}_{67} = -\Gamma^{7}_{56} = -\Gamma^{6}_{75} = -C. \end{aligned} \tag{8}$$

Thus the data for the three-dimensional subuniverse is the spatially inverted data of the other three-dimensional subuniverse. The common fourth coordinate is the way that the two subuniverses are coupled. The problem that we have encountered with respect to data (8), is that integrability is not satisfied. As integrability is a basic requirement in aesthetic field theory we have not pursued the data (8). Note in Ref. 8 we have investigated a theory not based on integrability. However, we feel that dropping this restriction would make the theory rather unnatural.

We have not been able to find an interesting set of data to study in the seven-dimensional theory so we will turn to other considerations.

IV. COMPLEX FIELDS

There is another approach we can take toward combining (1) and (6) into aesthetic field theory without the necessity of bringing in higher dimensions. We can allow the fields to be complex. We note that the aesthetic field equations can be obtained in exactly the same way as in Ref. 2, whether or not the change function is real or complex. The basic equations thus have the same form. We can think of the introduction of complex fields as a mathematical tool that enables the aesthetic principles of treating higher derivatives of the field, as well as all tensors in a uniform way (with respect to change), to be satisfied and still we can make use of both the origin data (1) and (6). Thus we can use aesthetic principles towards assigning the origin point data while not compromising the other aesthetic ideas in the theory.

We write

$$\Gamma^{i}_{jk} = A^{i}_{jk} + iB^{i}_{jk}, \tag{9}$$

with A^{i}_{jk} and B^{i}_{jk} real. We choose the real part of $\Gamma^{\alpha}_{\beta\gamma}$ to be the data (1) and the imaginary part of $\Gamma^{\alpha}_{\beta\gamma}$ to be the data (6). For simplicity we shall take

$$\begin{aligned} &A = 0.1, \\ &B = 0.1, \\ &C = 0.1. \end{aligned} \tag{10}$$

The $R^{i}_{jkl} = 0$ integrability equation is then satisfied for either data (1) or (6) when taken by itself—but when we combine the two sets of data via complex fields this is no longer the case. This is because our equations are nonlinear and the real and imaginary parts are coupled. Thus we are again in danger of not having a theory consistant with integrability. However, we shall find that the integrability equations (2) are satisfied in this case. The basic field equations are as usual

$$\frac{\partial \Gamma^i_{jk}}{\partial x^l} + \Gamma^m_{jk}\Gamma^i_{ml} - \Gamma^i_{mk}\Gamma^m_{jl} - \Gamma^i_{jm}\Gamma^m_{kl} = 0. \tag{11}$$

But now Γ^i_{jk} is complex. $R^t{}_{imk}$ is defined by

$$R^t{}_{imk} = \frac{\partial \Gamma^t_{ik}}{\partial x^m} - \frac{\partial \Gamma^t_{im}}{\partial x^k} - \Gamma^j_{im}\Gamma^t_{jk} + \Gamma^j_{ik}\Gamma^t_{jm}. \tag{12}$$

Using the field equations (11), $R^t{}_{imk}$ becomes

$$R^t{}_{imk} = \Gamma^t_{ij}\Gamma^j_{km} - \Gamma^t_{ij}\Gamma^j_{mk} + \Gamma^j_{im}\Gamma^t_{jk} - \Gamma^j_{ik}\Gamma^t_{jm}. \tag{13}$$

Using (9) we see

$$\begin{aligned}\mathrm{Re}R^t{}_{imk} &= A^t_{ij}A^j_{km} - B^t_{ij}B^j_{km} - A^t_{ij}A^j_{mk} + B^t_{ij}B^j_{mk}\\ &\quad + A^j_{im}A^t_{jk} - B^j_{im}B^t_{jk} - A^j_{ik}A^t_{jm} + B^j_{ik}B^t_{jm}\end{aligned} \tag{14}$$

and

$$\begin{aligned}\mathrm{Im}R^t{}_{imk} &= A^t_{ij}B^j_{km} + B^t_{ij}A^j_{km} - A^t_{ij}B^j_{mk} - B^t_{ij}A^j_{mk}\\ &\quad + A^j_{im}B^t_{jk} + B^j_{im}A^t_{jk} - A^j_{ik}B^t_{jm} - B^j_{ik}A^t_{jm}.\end{aligned} \tag{15}$$

Choosing $e^\alpha{}_i = \delta^\alpha_i$ ($R^t{}_{imk} = 0$ is independent of the choice of $e^\alpha{}_i$ so long as it is not singular), we see for A^i_{jk} given by (1) and B^i_{jk} given by (6) that (14) is zero but (15) is not. Then

$$R^t{}_{imk} = \mathrm{Re}R^t{}_{imk} + i\,\mathrm{Im}R^t{}_{imk} \tag{16}$$

is not zero.

However, from (2) we can calculate $\mathrm{Re}A^i{}_{jk\,pl}$ and $\mathrm{Im}A^i{}_{jk\,pl}$ for the data (1) and (6). The results of the calculation show that both real and imaginary parts of $A^i{}_{jk\,pl}$ are zero. Hence we have

$$A^i{}_{jk\,pl} = 0. \tag{17}$$

Thus the fields Γ^i_{jk} are integrable. That is, the mixed derivatives of all fields constructed from Γ^i_{jk} (including contractions) are symmetric. Furthermore, equation (17) is maintained by a complex $e^\alpha{}_i$ transformation.

Putting (9) into (11) and separating the real and imaginary parts yields

$$\begin{aligned}\frac{\partial A^i_{jk}}{\partial x^l} &= A^i_{mk}A^m_{jl} - B^i_{mk}B^m_{jl} + A^i_{jm}A^m_{kl}\\ &\quad - B^i_{jm}B^m_{kl} - A^m_{jk}A^i_{ml} + B^m_{jk}B^i_{ml},\end{aligned} \tag{18}$$

$$\begin{aligned}\frac{\partial B^i_{jk}}{\partial x^l} &= A^i_{mk}B^m_{jl} + B^i_{mk}A^m_{jl} + A^i_{jm}B^m_{kl}\\ &\quad + B^i_{jm}A^m_{kl} - A^m_{jk}B^i_{ml} - B^m_{jk}A^i_{ml}.\end{aligned} \tag{19}$$

These equations were then programmed for the computer. We were unable to sum an infinite series of corrections as we did in Ref. 9, so instead we employed a fourth order Runge–Kutta method. The field equations maintain the integrability equations at all points which, as in our previous work, gives a check on the reliability of the numbers coming off the computer.

If we consider a real quantity at the origin we note, because the change function Γ^i_{jk} is complex, the real quantity no longer remains real as we move away from the origin. On the other hand A^i_{jk} and B^i_{jk} are defined to be real, so they remain real, as is evidenced from the right-hand side of (18) and (19).

If $B^i_{jk} = 0$ the theory collapses into what we have considered before in Refs. 1 and 2. The question then is whether the introduction of B^i_{jk} leads to new types of effects or not.

V. COMPUTER RESULTS

We investigated Eqs. (18) and (19) on the computer. We used the following $e^\alpha{}_i$

$$e^\alpha{}_i = f^\alpha{}_i + ih^\alpha{}_i, \tag{20}$$

with

$$\begin{aligned}
&f^1{}_1 = 0.88, && f^1{}_2 = -0.42, && f^1{}_3 = -0.32, && f^1{}_0 = 0.2,\\
&f^2{}_1 = 0.5, && f^2{}_2 = 0.7, && f^2{}_3 = -0.425, && f^2{}_0 = 0.3,\\
&f^3{}_1 = 0.2, && f^3{}_2 = -0.55, && f^3{}_3 = 0.89, && f^3{}_0 = 0.6,\\
&f^0{}_1 = -0.16, && f^0{}_2 = -0.35, && f^0{}_3 = 0.28, && f^0{}_0 = 1.01,
\end{aligned} \tag{21}$$

and

$$\begin{aligned}
&h^1{}_1 = 0.3, && h^1{}_2 = -0.2, && h^1{}_3 = 0.11, && h^1{}_0 = 0.15,\\
&h^2{}_1 = -0.24, && h^2{}_2 = -0.16, && h^2{}_3 = 0.09, && h^2{}_0 = 0.07,\\
&h^3{}_1 = 0.13, && h^3{}_2 = -0.26, && h^3{}_3 = 0.31, && h^3{}_0 = -0.086,\\
&h^0{}_1 = 0.05, && h^0{}_2 = 0.1, && h^0{}_3 = -0.26, && h^0{}_0 = -0.31.
\end{aligned} \tag{22}$$

$\mathrm{Re}\Gamma^\alpha_{\beta\gamma}$ was taken to be (1) and $\mathrm{Im}\Gamma^\alpha_{\beta\gamma}$ was taken to be (6). We compared the results with the case when $\mathrm{Im}\Gamma^\alpha_{\beta\gamma}$ and $h^\alpha{}_i$ were taken to be zero.

The results are as follows. For the noncomplex theory there were two turnabout points on the x axis for Γ^1_{11}, and one planar maximum and one planar minimum. In the complex theory the number of turnabout points for A^1_{11} along the x axis was now 4. For the $z = 0$, $t = 0$ map we saw again one planar maximum and one planar minimum. There is thus a slight overall increase in complexity but nothing we have not seen before. A^1_{23} was also studied. It showed one planar maximum and one planar minimum. We found no new structure after long runs made from the origin in any case. It is quite remarkable that $\Gamma^i_{jk} \to 0$ at infinity since this is not an input into the theory. However, we have not yet been able to obtain a many body system in our work.

We considered other sets of data as well. We considered the following

$$\Gamma^1_{23} = \Gamma^2_{31} = \Gamma^3_{12} = -\Gamma^1_{32} = -\Gamma^3_{21} = -\Gamma^2_{13} = 0.1, \tag{23}$$

with the other $\Gamma^\alpha_{\beta\gamma} = 0$. We also considered the case when

$$\Gamma^1_{23} = \Gamma^2_{10} = -\Gamma^1_{20} = -\Gamma^2_{13} = 0.1 \quad (\text{other } \Gamma^\alpha_{\beta\gamma} = 0) \tag{24}$$

and the case

$$\Gamma^1_{23} = \Gamma^1_{20} = -\Gamma^2_{13} = -\Gamma^2_{10} = 0.1 \quad (\text{other } \Gamma^\alpha_{\beta\gamma} = 0). \tag{25}$$

Equations (23) – (25) would each lead to a constant Γ^i_{jk} if the theory were real. Each of the sets (23), (24), and (25) were taken together with (6) ($A = B = C = 0.1$) to form a complex theory. The results are not very different from the case of data (1) taken with data (6).

We have also found in all the cases above the B^i_{jk} behaved similarly to A^i_{jk}.

VI. CONCLUSIONS

We have found that the complex theory used with the group theoretical data has led to only slightly greater complexity as compared to the case of data (1) alone. We have not observed new effects not previously observed. However, the complex theory gives us more parameters at our disposal and it is not clear that the new effects cannot be obtained for some other, as yet untried, data. We have seen that with the complex theory we can find, without too much difficulty, solutions to the integrability equations, and thus solutions to the aesthetic field equations exist (at least locally).

[1]M. Muraskin, Ann. Phys. (N.Y.) **59**, 27 (1970).
[2]M. Muraskin, Int. J. Theor. Phys. **13**, 303 (1975).
[3]M. Muraskin, Found. Phys. **5**, 13 (1975).
[4]M. Muraskin, 1980, to be published in Foundations of Physics.
[5]M. Muraskin and B. Ring, Int. J. Theor. Phys. **11**, 93 (1974).
[6]M. Muraskin, to be published in Int. J. Theor. Phys.
[7]M. Muraskin and B. Ring, Int. J. Theor. Phys. **12**, 157 (1975).
[8]M. Muraskin and B. Ring, Found. Phys. **7**, 451 (1971).
[9]M. Muraskin and B. Ring, Inter. J. Theor. Phys. **6**, 105 (1972).

Spherically symmetric space–times with vanishing curvature scalar

Hubert Goenner
Institute of Theoretical Physics, University of Göttingen, D-3400 Göttingen
Peter Havas[a]
Department of Physics, Temple University, Philadelphia, Pennsylvania 19122

(Received 10 October 1979; accepted for publication 14 December 1979)

A systematic study of spherically symmetric space–times with vanishing curvature scalar R is undertaken. The solutions of the equation $R = 0$ are classified according to the number of different eigenvalues of the Ricci tensor $R_\alpha{}^\beta$ (which for spherical symmetry necessarily possesses one double eigenvalue) as well as to the vanishing or nonvanishing of the conformal curvature tensor. The case of conformally flat spherically symmetric solutions is fully integrated, as well as the case of one quadruple or two double eigenvalues. For the case of one triple and one single eigenvalue various classes of solutions are obtained, and for that of one double and two single eigenvalues a number of particular solutions with two free functions of one variable are found. It is noted that the methods developed in this paper allow the full integration of the case of plane symmetric solutions of $R = 0$.

I. INTRODUCTION

In this paper we study spherically symmetric solutions (s.s.s.) of the equation

$$R = 0, \tag{1}$$

where R is the Ricci scalar of the four-dimensional pseudo-Riemannian metric $g_{\alpha\beta}$. Spherical symmetry is defined here by the vanishing of the Lie derivative of the components of the metric tensor with regard to the generators of the rotation group $O(3,R)$.[1,2] This problem is of course of intrinsic mathematical interest. Our own interest arose in the course of a larger program aiming to establish whether, in the framework of Riemannian geometry, Einstein's theory of gravitation is the only one derivable from a variational principle which admits Birkhoff's theorem.[3,4] This theorem states that the vacuum field equations of Einstein's theory are satisfied by a *unique* one-parameter family of s.s.s.; these are static (outside the horizon) and asymptotically flat. If a theory of gravitation does not allow Birkhoff's theorem, the description of isolated gravitating systems meets with difficulties.[5,6]

Any solution of Eq. (1) solves also the field equations following from the Lagrangian $(-g)^{1/2}R^2$, i.e.,

$$2g^{\mu\nu}R_{;\sigma}{}^{\sigma} - 2R_{;}{}^{\mu\nu} - 2RR^{\mu\nu} + \tfrac{1}{2}g^{\mu\nu}R^2 = 0, \tag{2}$$

where $R^{\mu\nu}$ is the contracted curvature (Ricci) tensor, g is the determinant of the metric, and the semicolon denotes covariant differentiation. Every conformally flat solution of Eq. (1) satisfied the field equations following from the Lagrangian $(-g)^{1/2}R^{\alpha\beta}R_{\alpha\beta}$[7]:

$$g^{\mu\nu}R^{\sigma\tau}{}_{;\sigma\tau} + R^{\mu\nu}{}_{;\sigma}{}^{\sigma} - R^{\mu\sigma}{}_{;}{}^{\nu}{}_{\sigma} - R^{\nu\sigma}{}_{;}{}^{\mu}{}_{\sigma} - 2R^{\mu\sigma}R^{\nu}{}_{\sigma} + \tfrac{1}{2}g^{\mu\nu}R^{\sigma\tau}R_{\sigma\tau} = 0. \tag{3}$$

The field equations following from the Lagrangian $(-g)^{1/2}R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$ (where $R_{\alpha\beta\gamma\delta}$ is the Riemann–Christoffel curvature tensor) are a linear combination of Eqs. (2) and (3).[8] Therefore, any linear combination of the above three quadratic Lagrangians leads to a field equations which is solved by every conformally flat solution of Eq. (1). Because of the multiplicity of such solutions (see Sec. III) no such field equation admits a Birkhoff theorem.[6] Of course, any Lagrangian $(-g)^{1/2}f(R)$ where f is a polynominal not containing terms independent of or linear in R also leads to a field equation which is solved by any solution of Eq. (1) and thus also does not admit a Birkhoff theorem.[6]

We were further motivated to look into the s.s.s. of Eq. (1) by the fact that this equation seems to be the weakest possible vacuum field equation of any theory of gravitation in the framework of Riemannian geometry, e.g., Eq. (1) contains Nordströms gravitational theory, for which we display all s.s. vacuum solutions. In general, all s.s. Einstein–Maxwell fields and null fluids are to be found among the solutions of Eq. (1).

The generic s.s. metric contains two free functions of two variables and leads to two independent curvature invariants of order 2: the Ricci scalar R and the conformal invariant $C := (\tfrac{3}{4}C_{\alpha\beta\gamma\delta}C^{\alpha\beta\gamma\delta})^{1/2}$, where $C_{\alpha\beta\gamma\delta}$ denotes the conformal curvature tensor.[9] Thus, Eq. (1) will lead to a (coupled) system of two nonlinear second-order hyperbolic partial differential equations with one free coefficient function C (of two variables). Obviously, such a system cannot be expected to yield to a complete integration in terms of known functions or their integral representations in general. Nevertheless, below we do give the generic solution of Eq. (1) in the case $C = 0$ and for a specific one-parametric family of values $C \neq 0$, $\neq$constant. In both cases, it depends on two arbitrary functions of one variable each. We also give the general solutions in the case $C \neq 0$ whose Ricci tensor has either one quadruple or two double eigenvalues. They also contain up to two free functions, but now of the same variable.

After some general remarks in Sec. II, in Sec. III we give the generic conformally flat s.s.s. of Eq. (1) and discuss its subclasses according to the eigenvalue structure of the

[a]Research supported in part by Grant No. 77-28356 of the National Science Foundation and through the award of a study leave for 1978–9 by Temple University.

Ricci tensor. In Sec. IV two classes of s.s.s. with nonvanishing conformal curvature tensor are displayed. In Sec. V a different canonical form for the metric is used to derive other classes of s.s.s. of Eq. (1) of the same generality as those found in Sec. IV, which do not entirely coincide with them. The equivalence problem has not been tackled in general, however. Section VI is concerned with conformal mappings used for the generation of new s.s.s. of Eq. (1) from those already known, and our results are summarized in Table III and discussed in Sec. VII, where we also note the complete solution of Eq. (1) for the case of plane symmetry.

II. METHODS OF SOLUTION AND BASIC FORMULAS

We use two methods for obtaining s.s.s. of Eq. (1): (i) direct integration and (ii) generation of new solutions by conformal mapping from those already known.

Unfortunately, for the direct integration, the use of a single form of the metric tensor may not suffice to produce the optimal set of solutions. We use the following canonical forms:

$$ds^2 = z^2(u,v)[2e^{w(u,v)}du\,dv - d\Omega^2], \tag{4}$$

$$ds^2 = c^2(r,t)dt^2 - a^4(r,t)[dr^2 + r^2\,d\Omega^2], \tag{5}$$

$$ds^2 = F(u,r)du^2 + 2du\,dr - r^2\,d\Omega^2, \tag{6}$$

where $d\Omega^2 := d\theta^2 + \sin^2\theta\,d\varphi^2$ is the surface element of the unit 2-sphere. Any s.s. metric can be brought into either form (4) or (5),[10] while Eq. (6) is especially suited for metrics whose Ricci tensor admits two double eigenvalues.[11] The main advantage of the canonical form (4) is that it allows a linearization of Eq. (1); it also facilitates the discussion of s.s. metrics of Takeno type S_2.

Obviously, coordinate transformations of the form $u = f(u')$, $v = g(v')$ are the most general ones preserving the canonical form (4) (see Ref. 1, p. 18). This means that the functions z, w and z', w' defined by

$$z' = z, \quad w' = w + \ln A(u) + \ln B(v), \tag{7}$$

describe the same metric. Similarly, the canonical form (5) is preserved among others by the transformations $t = f(t')$, $r = (r')^{-1}$ which implies

$$c' = A(t')c, \quad a' = r'^{-4}a. \tag{8}$$

If $\partial F/\partial u \neq 0$, the canonical form (6) does not allow any nontrivial transformations of u and r preserving it.

The components of the curvature tensor of the canonical forms (4) and (5) are given in Appendix A, while those for Eq. (6) can be taken from Ref. 11.

The main drawback in working with different canonical forms for the metric lies in the fact that one must prove the inequivalence of the solutions obtained. In general, this proof is as difficult to carry through as the construction of the solutions themselves. In order to circumvent the equivalence problems we classify the solutions with regard to the algebraic structure of the Ricci tensor $R_\alpha{}^\beta$. (The Petrov type of $C_{\alpha\beta\gamma\delta}$ is of little help here[12] because s.s. metrics may only have type D or 0.) As is well known, the Ricci tensor of any s.s. metric possesses one double eigenvalue

$$\lambda_{3,4} = R_2{}^2 = R_3{}^3, \tag{9a}$$

while the remaining ones are given by

$$\lambda_{1,2} = \tfrac{1}{2}(R_0{}^0 + R_1{}^1) \pm \Delta^{1/2},$$
$$\Delta := \tfrac{1}{4}(R_0{}^0 - R_1{}^1)^2 + R_{01}R^{01}. \tag{9b}$$

The second method departs from the observation that the Ricci scalar R^* of the metric $g^*_{\alpha\beta} = g_{\alpha\beta}e^{2\phi}$ is related to R of $g_{\alpha\beta}$ through

$$R^* = R\,e^{-2\phi} + \Xi, \tag{10}$$

where Ξ depends on ϕ, $g_{\alpha\beta}$, and their first and second derivatives. Although we have not been able to integrate the differential equation $\Xi = 0$ completely for any of the canonical forms (4)–(6), we give particular solutions and apply them to generate further solutions of Eq. (1). This method is of special interest because the algebraic type of $R^*{}_\alpha{}^\beta$ is different (and, in general, less degenerate) than the type of $R_\alpha{}^\beta$, as will be shown in Appendix B. Thus, solutions of Eq. (1) with the most general algebraic type of $R_\alpha{}^\beta$ can be generated from those with one quadruple or two double eigenvalues, which we know completely.

III. THE GENERIC CONFORMALLY FLAT s.s.s. of $R = 0$

Starting from the canonical form (4) we calculate the curvature invariants R and C. This leads to the system of partial differential equations for the unknown functions $z(u,v)$, $w(u,v)$:

$$e^{-w}w_{,uv} + 1 - z^2C = 0, \tag{11a}$$

$$z_{,uv} + \tfrac{1}{6}z^3e^w(C - \tfrac{1}{2}R) = 0, \tag{11b}$$

where R and C are considered to be given, and the comma denotes a partial derivative. For conformally flat s.s.s. of Eq. (1), Eq. (11b) implies $z = A(u) + B(v)$, while w must satisfy the Liouville equation

$$w_{,uv} + e^w = 0, \tag{12}$$

whose solution is given by[13] $\exp w = 2D_{,u}E_{,v}[D(u) - E(v)]^{-2}$. After successively transforming to $\tilde{u} := D(u)$, $\tilde{v} := E(v)$ and $\rho := 2^{-1/2}(\tilde{u} - \tilde{v})$, $\tau := 2^{-1/2}(\tilde{u} + \tilde{v})$ we arrive at the metric

$$g_{\alpha\beta} = [A(\rho + \tau) + B(\rho - \tau)]^2\rho^{-2}\eta_{\alpha\beta}, \tag{13}$$

where $\eta_{\alpha\beta}$ is the Minkowski metric in polar coordinates ρ, θ, φ and $A(\rho + \tau)$, $B(\rho - \tau)$ are arbitrary functions. This is the general s.s.s. of Nordström's theory of gravitation.[14]

The only solution (13) with a quadruple eigenvalue of $R_\alpha{}^\beta$ is Minkowski space, while for solutions with two double eigenvalues either $B = 0$, $A \neq 0$ or $A = 0$, $B \neq 0$ holds. The Robinson–Lovelock metric[15] $g_{\alpha\beta} = a^2\rho^{-2}\eta_{\alpha\beta}$ is the most elementary example of this set of solutions. Surprisingly, those solutions which possess one triple and one single eigenvalue of $R_\alpha{}^\beta$ are difficult to identify. The lengthy equation to be satisfied by the functions A and B of Eq. (13) in this case, which we have not been able to solve in general, is given in Appendix C. The following three metrics are particular solutions of Eqs. (C1) and (C2):

$$g_{\alpha\beta} = \tau^2\eta_{\alpha\beta}, \tag{14a}$$

with

$$A(\tilde{u}) = \tfrac{1}{2}\tilde{u}^2, \quad B(\tilde{v}) = -\tfrac{1}{2}\tilde{v}^2,$$
$$g_{\alpha\beta} = (\tau^2 - \rho^2)^{-2}\eta_{\alpha\beta}, \tag{14b}$$
with
$$A(\tilde{u}) = 2^{-3/2}\tilde{u}^{-1}, \quad B(\tilde{v}) = -2^{-3/2}\tilde{v}^{-1},$$
and
$$g_{\alpha\beta} = \beta_0\left(1 - \frac{\tau^2-\rho^2}{2}\right)^2 \times \{[1 + \tfrac{1}{2}(\rho+\tau)^2]^2[1+\tfrac{1}{2}(\rho-\tau)^2]^2\}^{-1}\eta_{\alpha\beta}, \tag{14c}$$
with
$$A(\tilde{u}) = (\tfrac{1}{2}\beta)^{1/2}\tilde{u}(1+\tilde{u}^2)^{-1}, \; B(\tilde{v}) = -(\tfrac{1}{2}\beta)^{1/2}\tilde{v}(1+\tilde{v}^2)^{-1},$$
and β_0 a constant.

Solution (14c) was given by Tolman[16] in the isotropic form
$$ds^2 = dt^2 - \frac{\beta_0 - t^2}{(R_0 + (r^2/4R_0))^2}(dr^2 + r^2\,d\Omega^2), \tag{14c$'$}$$
and describes a null fluid solution of Einstein's theory of gravitation.

IV. SPHERICALLY SYMMETRIC SOLUTIONS WHICH ARE NOT CONFORMALLY FLAT

A glance at Eq. (11) shows that the case $R = 0$, $C \neq 0$ leaves the system of partial differential equations for z and w coupled, in general. If C is eliminated from Eq. (11b) by means of Eq. (11a), Eq. (1) becomes linear in z with w given:
$$z_{,uv} + \tfrac{1}{6}z(e^w + w_{,uv}) = 0. \tag{11b$'$}$$
Before we explicitly solve this linear equation for specific values of C, we first treat the case of the Ricci tensor having one quadruple or two double eigenvalues.

A. Degenerate eigenvalue structure of $R_\alpha{}^\beta$

According to Ref. 4, of the five subcases of the canonical form (4) of metrics whose Ricci tensor has two double eigenvalues, only
$$ds^2 = 2G_{,u}\,du\,dv - G^2(u,v)d\Omega^2,$$
and
$$ds^2 = 2G_{,v}\,du\,dv - G^2(u,v)d\Omega^2,$$
lead to $R - 2C \neq 0$. If G is introduced as a new variable, the canonical form (6) turns up as the most general canonical form of a s.s. metric leading to a Ricci tensor with two double eigenvalues and $R - 2C \neq 0$. The general solution of Eq. (1) with $C \neq 0$ and a quadruple eigenvalue of $R_\alpha{}^\beta$ is given by[11]
$$ds^2 = \left[1 - \frac{2m(u)}{r}\right]du^2 + 2du\,dr - r^2\,d\Omega^2, \tag{15}$$
with one arbitrary function $m(u) \neq 0$.

Two double eigenvalues of $R_\alpha{}^\beta$ occur for[11]
$$ds^2 = \left[1 - \frac{2m(u)}{r} + \frac{e^2(u)}{r^2}\right]du^2 + 2du\,dr - r^2\,d\Omega^2, \tag{16}$$
where $m(u)$, $e(u) \neq 0$ are arbitrary functions. Among the metrics (15) and (16) are many well known solutions of Einstein's field equations, including all Einstein–Maxwell fields.

Again, it is difficult to pin down solutions with one triple eigenvalue of $R_\alpha{}^\beta$ and one single one. For the case $C \neq 0$ we have found only one such s.s.s. of Eq. (1):
$$ds^2 = r^{\pm 2/\sqrt{3}}dt^2 - r^{-2\mp 4/\sqrt{3}}(dr^2 + r^2\,d\Omega^2). \tag{17}$$

B. Generic eigenvalue structure of $R_\alpha{}^\beta$

We now consider metrics whose Ricci tensor has one double and two single eigenvalues. Of the various possibilities to decouple the system (11a) and (11b), a promising one is to put
$$C = (1 - \sigma_0)z^{-2}, \tag{18}$$
where σ_0 is an arbitrary constant. We assume $\sigma_0 \neq 1$, as the case $C = 0$ was treated exhaustively in Sec. III. Two subcases are distinguished:

1. $\sigma_0 = 0$

Here, Eq. (11a) reduces to the wave equation $w_{,uv} = 0$. Due to the freedom of introducing new independent variables [see Eq. (7)] we may set w equal to zero without loss of generality. Then Eq. (11b) becomes
$$z_{,uv} + \tfrac{1}{6}z = 0. \tag{19}$$
The general solution of this equation (essentially the telegrapher's equation) can be found in many books[17] in the form
$$z = \tfrac{1}{2}[\varphi(r+t) + \varphi(r-t)] + \tfrac{1}{2}\int_{r-t}^{r+t} d\xi\,\psi(\xi)J_0\left(\sqrt{\tfrac{1}{6}[t^2-(r-\xi)^2]}\right) + \frac{t}{2\sqrt{6}}\int_{r-t}^{r+t} d\xi\,\varphi(\xi)\frac{J_1\left(\sqrt{\tfrac{1}{6}[t^2-(r-\xi)^2]}\right)}{\sqrt{t^2-(r-\xi)^2}}, \tag{20}$$
where φ and ψ are arbitrary functions, J_0 and J_1 Bessel functions, and
$$r := 2^{-1/2}(u - v), \quad t := 2^{-1/2}(u+v).$$
The function z of Eq. (20) solves the boundary value problem
$$z(r,0) = \varphi(r), \quad z_{,t}(r,0) = \psi(r). \tag{21}$$
The corresponding s.s.s. of Eq. (1) is
$$ds^2 = z^2(2du\,dv - d\Omega^2). \tag{22}$$
In place of the integral representation (20) we may use another one:
$$z = z_1 + z_2 + z_3 + (a_1 r + a_2)(b_1\sin(t/\sqrt{3}) + b_2\cos(t/\sqrt{3})) + (c_1 t + c_2) \times (d_1\sinh(r/\sqrt{3}) + d_2\cosh(r/\sqrt{3})), \tag{23}$$
where $a_1,\ldots,d_2$ are constants and
$$z_1 = \int_0^\infty d\lambda\,[A(\lambda)\sinh t\sqrt{\lambda} + B(\lambda)\cosh t\sqrt{\lambda}] \times\left[C(\lambda)\sinh r\sqrt{\lambda+\tfrac{1}{3}} + D(\lambda)\cosh r\sqrt{\lambda+\tfrac{1}{3}}\right], \tag{24a}$$

$$z_2 = \int_0^{1/3} d\lambda\, [A(\lambda)\sin t\sqrt{\lambda} + B(\lambda)\cos t\sqrt{\lambda}\,] \times\left[C(\lambda)\sinh r\sqrt{\tfrac{1}{3}-\lambda} + D(\lambda)\cosh r\sqrt{\tfrac{1}{3}-\lambda}\right], \tag{24b}$$

$$z_3 = \int_0^{\infty} d\lambda \left[A(\lambda)\sin t\sqrt{\lambda+\tfrac{1}{3}} + B(\lambda)\cos t\sqrt{\lambda+\tfrac{1}{3}}\right] \times [C(\lambda)\sin r\sqrt{\lambda} + D(\lambda)\cos r\sqrt{\lambda}\,]. \tag{24c}$$

Among the functions $A(\lambda)$, $B(\lambda)$, $C(\lambda)$, and $D(\lambda)$ only two are free in the sense of leading to different solutions. By the special choice of

$$A = C = 0,\quad B(\lambda) = (3\lambda)^{-1/2},\quad D(\lambda) = (1-3\lambda)^{-1/2},$$

we obtain from Eq. (24b)

$$z_2 = \frac{\pi}{3} J_0\left(\frac{t^2 - r^2}{\sqrt{3}}\right), \tag{25}$$

or the Riemann function of Eq. (19) which was used in the integral representation (20).

2. $\sigma_0 \neq 0, \neq 1$

Here, Eq. (11a) reduces to $w_{,uv} + \sigma_0 e^{-w} = 0$ with the general solution

$$w = \ln\{(2/\sigma_0)(dA(u)/du)(dB(v)/dv)[A(u) - B(v)]^{-2}\}. \tag{26}$$

By a coordinate transformation we arrive at the metric

$$ds^2 = 2z^2(u,v)(u-v)^{-2}[2\sigma_0^{-1}\, du\, dv - \tfrac{1}{2}(u-v)^2\, d\Omega^2], \tag{27}$$

where now u and v denote the transformed variables, or, equivalently,

$$ds^2 = r^{-2}z^2(r,t)\,[\sigma_0^{-1}(dt^2 - dr^2) - r^2 d\Omega^2]. \tag{27'}$$

Equation (11b) now reads

$$(u-v)^2 z_{,uv} + \frac{1-\sigma_0}{3\sigma_0} z = 0. \tag{28}$$

Introducing a new dependent variable χ by

$$z := (u-v)^{\lambda}\chi \quad \text{with} \quad \lambda := \tfrac{1}{2}\left(1 \pm \sqrt{\frac{4-\sigma_0}{3\sigma_0}}\right),\quad 0 \leqslant \sigma_0 \leqslant 4,$$

Eq. (28) is transformed into

$$\chi_{,uv} - \lambda(u-v)^{-1}(\chi_{,u} - \chi_{,v}) = 0. \tag{29}$$

This is a special case of the Euler–Darboux equation whose general solution is known.[18] For $\sigma_0 > 1$, $\sigma_0 \neq 4$ we obtain as the solution of Eq. (28)

$$\begin{aligned} z = {} & -(u-v)^{(1/2)(1 \mp \Sigma)} \times \int_0^1 d\xi\, \varphi\,[u + (v-u)\xi\,]\,[\,\xi(1-\xi)]^{-\lambda} \\ & + (u-v)^{(1/2)(1 \pm \Sigma)} \times \int_0^1 d\xi\, \psi[u + (v-u)\xi\,]\,[\,\xi(1-\xi)]^{\lambda-1}, \end{aligned} \tag{30a}$$

where $\Sigma := [(4-\sigma_0)/3\sigma_0]^{1/2}$ and φ, ψ are arbitrary functions. If $\sigma_0 = 4$, the solution is given by

$$\begin{aligned} z = (u-v)^{1/2}\Big\{ & \int_0^1 d\xi\, \varphi\,[u + (v-u)\xi\,]\,[\,\xi(1-\xi)]^{-1/2} \\ & + \int_0^1 d\xi\, \psi[u + (v-u)\xi\,]\,[\,\xi(1-\xi)]^{-1/2} \times \ln[\,\xi(1-\xi)(v-u)]. \end{aligned} \tag{30b}$$

The solution of Eq. (29) for values $\sigma_0 \leqslant 1$ is known as well (see Ref. 18). For particular values of σ_0 the general solution of Eq. (28) may be written without use of an integral representation. For example, if $\sigma_0 = \frac{1}{7}$, the general solution of Eq. (28) is given by

$$\begin{aligned} z = {} & (d^2/dx^2)f_1(x) - 10^{-1/2}(d/dx)f_1(x) \\ & + (d^2/dy^2)f_2(y) + 10^{-1/2}(d/dy)f_2(y), \end{aligned} \tag{31}$$

where $x := 20^{1/2}u$, $y := 20^{1/2}v$, and $f_1(x)$ and $f_2(y)$ are arbitrary functions.

Of course, the classes of solutions given by Eq. (20) or (23) with Eq. (24) and by Eqs. (30a), (30b), and (31) may also contain solutions leading to one quadruple or two double eigenvalues.

V. ALTERNATIVE APPROACH TO SOLUTIONS WHICH ARE NOT CONFORMALLY FLAT

Up to here the equivalence problem was avoided altogether, because we applied invariant criteria (eigenvalue structure of $R_\alpha{}^\beta$) to distinguish the solutions found as well as the same canonical form for the metric within each algebraic type. We know that the algebraically most general type of s.s.s. of Eq. (1) with $C \neq 0$ contains many more solutions than the rather rich classes exhibited in Sec. IV, since evidently Eq. (11b') possesses classes of solutions containing one free function w of two variables and two free functions of one variable each. In order to see if another canonical form facilitates the integration or even leads to still more general classes of solutions we now use the canonical form (5) for the s.s. metric (isotropic coordinates).

In place of the functions c, a of the form (5) it is convenient to introduce other functions $\psi(r,t)$, $\chi(r,t)$ defined by

$$c = \psi/\chi,\quad a = \chi/r. \tag{32}$$

The metric then becomes

$$ds^2 = (\psi/\chi)^2 dt^2 - (\chi/r)^4(dr^2 + r^2\, d\Omega^2), \tag{33}$$

or, equivalently, due to Eq. (8),

$$ds^2 = (\psi/\chi)^2 dt^2 - \chi^4(d\rho^2 + \rho^2\, d\Omega^2), \tag{33'}$$

where $\rho := r^{-1}$. A straightforward calculation of the Ricci tensor belonging to Eq. (33) leads to

$$\begin{aligned} R = {} & -(2r^4/\chi^4)(3(\chi''/\chi) + (\psi''/\psi)) + (12/\psi^2)[(\ddot{\chi}/\chi) \\ & + 4(\dot{\chi}/\chi)^2 - (\dot{\psi}/\psi)(\dot{\chi}/\chi)], \\ & \chi' := \chi_{,r},\quad \dot{\chi} := \chi_{,t}\,. \end{aligned} \tag{34}$$

Equation (34) is valid for any s.s. metric as χ and ψ are arbitrary functions of two variables. We now reduce the generality by assuming

$$\chi_{,t} = 0, \tag{35}$$

i.e., $\chi = \chi(r)$. Then a class of solutions of Eq. (1) is given by *arbitrary* $\chi(r)$ and $\psi(r,t)$ satisfying the linear equation

TABLE I. Explict solutions of Eq. (37) for some simple functions $f(r)$; J_ν and Y_ν are Bessel functions; $\xi := \alpha_0 + \beta_0 r$.

$f(r)$		$\chi(r) = a_0\chi_1 + b_0\chi_2$		$\psi(r,t) = c(t)\psi_1 + d(t)\psi_2$	
		$\chi_1(r)$	$\chi_2(r)$	$\psi_1(r)$	$\psi_2(r)$
0		1	r	1	r
$3\alpha_0^2$		$\cosh\alpha_0 r$	$\sinh\alpha_0 r$	$\cos\sqrt{3}\alpha_0 r$	$\sin\sqrt{3}\alpha_0 r$
$-3\alpha_0^2$		$\cos\alpha_0 r$	$\sin\alpha_0 r$	$\cosh\sqrt{3}\alpha_0 r$	$\sinh\sqrt{3}\alpha_0 r$
$\alpha_0 + \beta_0 r$					
$\alpha_0, \beta_0 > 0$		$\xi^{1/2} J_{1/3}\left(\frac{2}{\beta_0}\xi^{3/2}\right)$	$\xi^{1/2} Y_{1/3}\left(\frac{2}{\beta_0}\xi^{3/2}\right)$	$\xi^{1/2} J_{1/3}\left(-\frac{2}{3\beta_0}\xi^{3/2}\right)$	$\xi^{1/2} Y_{1/3}\left(-\frac{2}{3\beta_0}\xi^{3/2}\right)$
$\frac{\alpha_0}{r^2}$	$\alpha_0 > \frac{1}{4}$	$r^{1/2+\sqrt{1/4+\alpha_0/3}}$	$r^{1/2-\sqrt{1/4+\alpha_0/3}}$	$r^{1/2}\cos\left[\sqrt{\alpha_0 - \frac{1}{4}}\ \ln r\right]$	$r^{1/2}\sin\left[\sqrt{\alpha_0 - \frac{1}{4}}\ \ln r\right]$
	$\alpha_0 = \frac{1}{4}$	$r^{1/2+3^{-1/2}}$	$r^{1/2+3^{-1/2}}$	$r^{1/2}$	$r^{1/2}\ln r$
	$-\frac{3}{4} < \alpha_0 < \frac{1}{4}$	$r^{1/2+\sqrt{1/4+\alpha_0/3}}$	$r^{1/2-\sqrt{1/4+\alpha_0/3}}$	$r^{1/2+\sqrt{1/4-\alpha_0}}$	$r^{1/2+\sqrt{1/4-\alpha_0}}$
	$\alpha_0 = -\frac{3}{4}$	$r^{1/2}$	$r^{1/2}\ln r$	$r^{3/2}$	$r^{-1/2}$
	$\alpha_0 < -\frac{3}{4}$	$r^{1/2}\cos\left[\sqrt{-\alpha_0/3 - \frac{1}{4}}\ln r\right]$	$r^{1/2}\sin\left[\sqrt{-\alpha_0/3 - \frac{1}{4}}\ln r\right]$	$r^{\sqrt{1/4-\alpha_0}}$	$r^{1/2-\sqrt{1/4-\alpha_0}}$

TABLE II. Some classes of mathematical functions occurring as solutions ψ and χ of Eq. (37). $p(r)$ is the Weierstrass elliptic function.

$f(r)$	Special functions of mathematical physics
$\beta_0 r^m$	Bessel functions
$\alpha_0 \cos 2r + \beta_0$	Mathieu functions
$\alpha_0 p(r) + \beta_0$	Lamé functions
$1 + \frac{\alpha_0}{r}$	confluent hypergeometric function
$\frac{\beta_0}{\cosh^2\alpha_0 r}$	hypergeometric function
$-\frac{\nu(\nu+1)}{\sin^2 r}$	Legendre functions

$$\psi'' + 3(\chi''(r)/\chi(r))\psi = 0. \tag{36}$$

For every chosen $\chi(r)$, Eq. (36) is of the type $\psi'' + f(r)\psi = 0$ with known coefficient function $f(r)$. However, as we do want both functions $\chi(r)$ and $\psi(r,t)$ explicitly, it is convenient to solve the system

$$\begin{aligned} &\psi''(r,t) + f(r)\psi(r,t) = 0, \\ &3\chi''(r) - f(r)\chi(r) = 0, \end{aligned} \tag{37}$$

for ψ and χ with arbitrarily chosen parametric function $f(r)$. This is easily done for large classes of functions $f(r)$. In fact, among the differential equations which can be solved explicitly are the equations of Bessel, Mathieu, Lamé, and Hill and the hypergeometric differential equation. Also each of the two equations (37) can be put into the form of the one-dimensional time-independent Schrödinger equation in Cartesian coordinates (or the radial equation in polar coordinates). Table I lists some of the more obvious results while Table II collects some of the differential equations mentioned above. Of the two free functions of the coordinate t [the integration "constants" of the first of Eqs. (37)], one can always be transformed to 1. If $\dot{\psi} \neq 0$, the only solution leading to two double eigenvalues of $R_\alpha{}^\beta$ is given by $\chi = a_0 r^{1/2}$.

A second set of solutions of Eq. (1) is obtained by assuming

$$\psi = \psi(r), \quad \chi(r,t) = t^{1/5}\eta(r), \tag{38}$$

where $\psi(r)$ and $\eta(r)$ (in place of χ) satisfy the system (37) and no free functions of t do occur. The corresponding metric is

$$ds^2 = t^{-2/5}[\psi(r)/\eta(r)]^2 dt^2 - t^{4/5}[\eta(r)/r]^4(dr^2 + r^2 d\Omega^2), \tag{39}$$

which after a coordinate transformation and a rescaling of ψ and η becomes

$$ds^2 = [\psi(r)/\eta(r)]^2 d\tau^2 - \tau[\eta(r)/r]^4(dr^2 + r^2\, d\Omega^2). \tag{39'}$$

The conformal curvature invariant C for the general s.s. metric (33) is given by

$$\begin{aligned} C = (r/\chi)^4[&3(\chi''/\chi) + (9/r)(\chi'/\chi) - 12(\chi'/\chi)^2 \\ &+ 6(\chi'/\chi)(\psi'/\psi) + (3/r)(\psi'/\psi) - (\psi''/\psi)]. \end{aligned} \tag{40}$$

It should be noted that this expression does not contain any time derivatives. Thus, C depends, in general, on one free function of t and one free function of r, i.e., $f(r)$. This is the same degree of generality which was achieved in Sec. IV when departing from the canonical form (4). The advantage here is that many of the solutions are more familiar than the integral representations (20), (24), (30), and (30′).

As the canonical forms (4) and (5) can always be transformed into each other there is a distinct possibility that all of the solutions found in Sec. IV and all of Sec. V are equivalent. However, from the following example one can convince oneself that at least some solutions do not belong to both classes. Take the s.s.s of Eq. (1) obtained from the first entry in Table I, i.e., with

$$ds^2 = \left[\frac{c(t)+rd(t)}{b+ar}\right]^2 dt^2 - \left(a+\frac{b}{r}\right)^4 (dr^2 + r^2 d\Omega^2). \tag{41}$$

Let $ab \neq 0$ and $c(t)b^{-1} + d(t)a^{-1} \neq 0$. Then the coordinate transformation

$$\rho = r\left(a+\frac{b}{r}\right)^2, \quad \tau = \tau(t), \quad 2\frac{d\tau}{dt} = \frac{c(t)}{b} + \frac{d(t)}{a}, \tag{42}$$

brings the metric into the form

$$ds^2 = \left[1 \pm e(\tau)\sqrt{1-\frac{4ab}{\rho}}\right]^2 d\tau^2 - \frac{d\rho^2}{1-4ab/\rho} - \rho^2 d\Omega^2, \tag{41′}$$

where $e(t) := [bd(t) - ac(t)]/[bd(t) + ac(t)]$. To simplify matters we assume

$$e(\tau) = 0. \tag{43}$$

It is easy to show that the Ricci tensor of Eq. (41) with Eq. (43) has three different eigenvalues. A further transformation of the spacelike coordinate ρ into $\bar\rho$ by

$$\bar\rho = 4ab \ln\left[\sqrt{\frac{\rho}{4ab}} + \sqrt{\frac{\rho}{4ab}-1}\right] + \sqrt{\rho(\rho-4ab)}, \tag{44}$$

leads to

$$ds^2 = \rho^2(\bar\rho)[\rho^{-2}(\bar\rho)(d\tau^2 - d\bar\rho^2) - d\Omega^2], \tag{45}$$

i.e., a metric of the canonical form (4) dealt with in Sec. IV. From Eq. (45) we have

$$\rho(\bar\rho) = z\left(\frac{u-v}{\sqrt{2}}\right), \quad e^w = z^{-2}. \tag{46}$$

The system (11a) and (11b) with $R = 0$, $z = z(u - v)$, and $\exp w = z^{-2}$ can be solved explicitly (see Appendix D). The solutions satisfy

$$C = \tfrac{3}{2}z^{-2}(1 + 2z_{,u}z_{,v}), \tag{47}$$

with $1 + 2z_{,u}z_{,v} \neq$ const. Consequently, they cannot be contained within the s.s.s. of Eq. (1) given in Sec. IV.

VI. GENERATION OF SOLUTIONS BY CONFORMAL MAPPING

We now apply the method of conformal mapping described in Sec. II for the construction of new solutions $g^*_{\alpha\beta} = g_{\alpha\beta}\exp 2\phi$ of

$$R^* = 0 \tag{48}$$

from given ones $g_{\alpha\beta}$ of Eq. (1). For the canonical forms (4), (33), and (6) we have, respectively, with $\Psi := \exp\phi$,

$$R^* = \Psi^{-2}R + 12\Psi^{-3}e^{-w}z^{-2} \times (\Psi_{,uv} + z^{-1}z_{,u}\Psi_{,v} + z^{-1}z_{,v}\Psi_{,u}), \tag{49}$$

$$R^* = \Psi^{-2}R + 6\Psi^{-3}\left\{\left(\frac{\chi}{\psi}\right)^2\left(\Psi_{,tt} - \Psi_{,t}\frac{\psi_{,t}}{\psi}\right) - \left(\frac{r}{\chi}\right)^4\left[\Psi_{,rr} + \Psi_{,r}\left(\frac{\psi_{,r}}{\psi} + \frac{\chi_{,r}}{\chi}\right)\right]\right\}, \tag{50}$$

$$R^* = \Psi^{-2}R + 6\Psi^{-3}[2\Psi_{,ur} - 2\Psi_{,rr}F + (2/r)\Psi_{,u} - \Psi_{,r}(F_{,r} + (2/r)F)]. \tag{51}$$

As a first application all conformally flat s.s. metrics generated from the Minkowski metric will be determined. With $\psi = \chi = r$ in Eq. (33) we obtain from Eq. (50)

$$R^* = \Psi^{-2}R + 6\Psi^{-3}(\Psi_{,tt} - \Psi_{,rr} - (2/r)\Psi_{,r}). \tag{52}$$

With the new dependent variable $y := r\Psi$ we obtain

$$R^* = r^2y^{-2}R + 6r^2y^{-3}(y_{,tt} - y_{,rr}). \tag{53}$$

Thus, the Minkowski metric is carried into a solution of Eq. (48) if y satisfies the wave equation. This leads back to the result (13) of Sec. III.

As a second application we set in Eq. (50)

$$\Psi = a + bt + (c + dt)\int^r \frac{dx}{\psi(x)\chi(x)}, \tag{54}$$

with $\psi_{,t} = 0$, and a, b, c, d constants. Obviously $g^*_{\alpha\beta} = \Psi^2 g_{\alpha\beta}$ solves Eq. (48) if $g_{\alpha\beta}$ solves Eq. (1). To each of the multitude of solutions obtained in Sec. V in which the free function of t is replaced by a constant we obtain a further two- or three-parametric set of solutions of Eq. (1):

$$ds^{*2} = \left[a + bt + (c + dt)\int^r \frac{dx}{\psi(x)\chi(x)}\right]^2 \times\{[\psi(r)/\chi(r)]dt^2 - [\chi(r)/r]^4(dr^2 + r^2 d\Omega^2)\}, \tag{55}$$

which in general is not equivalent to the metrics we started from, as can be seen from the special subcase $c = d = 0$. By a coordinate transformation the metric transforms into Eq. (39) of Sec. V. Thus, a time-dependent solution of Eq. (1) has been generated by a time-independent one. To give a specific example we start from the isotropic form of the Schwarzschild metric

$$ds^2 = \left(\frac{2r-m}{2r+m}\right)^2 dt^2 - \left(\frac{r+\frac{1}{2}m}{r}\right)^4(dr^2 + r^2 d\Omega^2). \tag{56}$$

Now $\psi = r - \frac{1}{2}m$, $\chi = r + \frac{1}{2}m$, and from Eq. (55) we obtain the following solution of Eq. (48):

$$ds^{*2} = \left[a + bt + (c + dt)m^{-1}\ln\frac{2r-m}{2r+m}\right]^2 ds^2. \tag{57}$$

Using the Schwarzschild coordinate $\rho := r(1 + \frac{1}{2}m/r)^2$, the s.s. metric (57) takes the form

$$ds^{*2} = \left[a + bt + (c + dt)\frac{1}{2m}\ln\left(1 - \frac{2m}{\rho}\right)\right]^2 \times\left[\left(1 - \frac{2m}{\rho}\right)dt^2 - \left(1 - \frac{2m}{\rho}\right)^{-1} d\rho^2 - \rho^2 d\Omega^2\right]. \tag{58}$$

This is a generalization of the solutions (vi) and (vii) of Wynne and Derrick[19] which follow for $b = d = 0$.

In our last application we generate solutions of Eq. (1) with the most general algebraical structure of $R_\alpha{}^\beta$ from those with two double eigenvalues. From Eq. (51) it follows that the metric

$$ds^{*2} = \left[u + \int^r dx \frac{x^2 + a_0}{x^2 F(x)}\right]^2 ds^2, \qquad (59)$$

where a_0 is a constant, satisfies Eq. (48) if $ds^2 = F(r)du^2 + 2du\,dr - r^2 d\Omega^2$ solves Eq. (1). From Eq. (16) the most general such solution implies

$$F(r) = 1 - 2m/r + e^2/r^2, \qquad (60)$$

with constants m and e. As a special case we consider

$$-e = m = \alpha_0, \qquad (61)$$

and obtain from Eq. (59)

$$ds^{*2} = [u + r - (\alpha_0{}^2 + a_0)(r + \alpha_0)^{-1} - 2\alpha_0 \ln(r + \alpha_0)]^2 ds^2. \qquad (62)$$

If we set $a_0 = 0$ and introduce new coordinates τ, ρ by

$$\tau = \tfrac{1}{2}[u + r - \alpha_0{}^2(r + \alpha_0)^{-1} - 2\alpha_0 \ln(r + \alpha_0)]^2,$$

$$\rho = 2^{1/2}(r + \alpha_0),$$

Eq. (62) becomes, after a rescaling of α_0,

$$ds^{*2} = \left(1 - \frac{\alpha_0}{\rho}\right)^{-2} d\tau^2 - \tau\left(1 - \frac{\alpha_0}{\rho}\right)^2 (d\rho^2 + \rho^2 d\Omega^2). \qquad (62')$$

Equation (62′) is contained within the class of solutions (33) where $\chi = t^{1/4}\mu(r)$ and $\psi = t^{1/4}\nu(r)$, while $\nu(r)$ and $\mu(r)$ satisfy

$$\nu_{,rr} + f(r)\nu = 0, \quad 3\mu_{,rr} - f(r)\mu = 0,$$

with

$$f(r) = \tfrac{3}{4}\alpha_0{}^2 r^{-2}(r - \alpha_0)^{-1}.$$

VII. DISCUSSION

In our investigation of spherically symmetric solutions of $R = 0$ we have obtained a considerable number of explicitly given solutions depending on two free functions of one variable each. As far as we know all solutions presented in the literature are contained within the classes found in this paper and collected in Table III.

However, we were unable to obtain the general s.s.s. of Eq. (1) and had to content ourselves with showing how s.s.s. can be constructed directly from different canonical forms of the s.s. metric or with the help of conformal mappings. It might be feasible to obtain the general solution for the case of one triple and one single root, but at present the general case of one double and two single roots appears to be beyond reach. Investigations of this question are continuing.

Although we have not succeeded in establishing to what extent the classes of solutions belonging to different canonical forms are equivalent or not, the discussion given in Secs. IV and V shows that they are neither mutually exclusive nor fully equivalent.

In Sec. VI a method was developed to generate new s.s.s. of Eq. (1) with equal or more general eigenvalue structure from known ones. This structure is not readily apparent, however, but needs detailed investigation in each particular case. Various examples were given which show that the method does yield inequivalent new solutions.

It is interesting to note that in the case of plane symmetry Eq. (1) can be integrated completely. Then the metric is of the form[20]

$$ds^2 = z^2(2e^w du\,dv - dx^2 - dy^2), \qquad (63)$$

where z and w are functions of u and v; instead of the system (11a) and (11b) one obtains

$$w_{,uv} - z^2 e^w C = 0, \qquad (64a)$$

$$z_{,uv} + \tfrac{1}{6} z^3 e^w (C - \tfrac{1}{2}R) = 0. \qquad (64b)$$

If $C = 0$, this system leads to the general form of the metric (63):

$$ds^2 = [A(\bar{u}) + B(\bar{v})]^2 (2d\bar{u}\,d\bar{v} - dx^2 - dy^2). \qquad (65)$$

If $C \neq 0$, z is arbitrary, and w is the solution of the inhomogeneous wave equation

$$w_{,uv} = -(6/z) z_{,uv}, \qquad (66)$$

for which the integral representation is known.[18]

In the course of working on $R = 0$, we have also obtained a number of s.s.s. of $R = \text{const.} \neq 0$. A paper on this problem is in preparation.

APPENDIX A

Using Takeno's notation[1] for the six nonvanishing components of the curvature tensor of a s.s. space–time

$$\begin{aligned} &\alpha := R_{12}{}^{12} = R_{13}{}^{13}, \quad \beta := R_{02}{}^{02} = R_{03}{}^{03}, \\ &\gamma := R_{12}{}^{20} = R_{13}{}^{30}, \quad \delta := R_{20}{}^{12} = R_{30}{}^{13}, \\ &\xi := R_{01}{}^{01}, \quad \eta := R_{23}{}^{23}, \end{aligned} \qquad (A1)$$

we obtain for the canonical form (4)

$$\begin{aligned} &\alpha = \beta = -z^{-3} e^{-w} z_{,uv}, \\ &\gamma = z^{-1} e^{-w}[-(z^{-1})_{,vv} + (z^{-1})_{,v} w_{,v}], \\ &\delta = z^{-1} e^{-w}[-(z^{-1})_{,uu} + (z^{-1})_{,u} w_{,u}], \\ &\xi = z^{-1} e^{-w}[(z^{-1})_{,uv} - z^{-2} z_{,uv} + z^{-1} w_{,uv}], \\ &\eta = -z^{-2} - 2e^{-w}(z^{-1})_{,u}(z^{-1})_{,v}, \end{aligned} \qquad (A2)$$

from which Eqs. (11a) and (11b) follow.

For the canonical form (5) we have

$$\begin{aligned} &\alpha = 2a^{-4}\left[\frac{a''}{a} - \left(\frac{a'}{a}\right)^2 + \frac{1}{r}\frac{a'}{a}\right] - 4c^{-2}\left(\frac{\dot{a}}{a}\right)^2, \\ &\beta = a^{-4}\left[\frac{1}{r} + 2\frac{a'}{a}\right]\frac{c'}{c} - 2c^{-2}\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 - \frac{\dot{a}}{a}\frac{\dot{c}}{c}\right], \\ &\gamma = -a^4 c^{-2}\delta = 2c^{-2}\left[\frac{\dot{a}'}{a} - \frac{\dot{a}}{a}\frac{a'}{a} - \frac{\dot{a}}{a}\frac{c'}{c}\right], \\ &\xi = a^{-4}\left[\frac{c''}{c} - 2\frac{a'}{a}\frac{c'}{c}\right] - 2c^{-2}\left[\left(\frac{\dot{a}}{a}\right)^2 + \frac{\ddot{a}}{a} - \frac{\dot{a}}{a}\frac{\dot{c}}{c}\right], \\ &\eta = 4a^{-4}\left[\left(\frac{a'}{a}\right)^2 + \frac{1}{r}\frac{a'}{a}\right] - 4c^{-2}\left(\frac{\dot{a}}{a}\right)^2, \end{aligned} \qquad (A3)$$

where the prime and dot denote derivatives with respect to r and t, respectively.

TABLE III. Collection of s.s.s. of $R=0$ obtained. If not stated otherwise, the entries form the general solution of the case considered. The solutions (55) and (59) obtained by conformal mapping overlap this classification and therefore are not listed in the table.

	Eigenvalues of $R_\alpha{}^\beta$: Quadruple	Two double	One triple, one single	One double, two single
$C=0$	Minkowski space	Eq. (13) with $AB=0$, $A^2+B^2\neq 0$	Eq. (14a)–(14c) (incomplete)	Eq. (13) with $AB\neq 0$
$C\neq 0$	Eq. (15) with $m\neq 0$	Eq. (16) with $e\neq 0$	Eq. (17) (incomplete)	Canonical form (4) of metric: Eq. (22) with Eq. (20) or (23), Eqs. (24) and (27′) with Eqs. (30a) and (30b) or (31) (particular solutions) Canonical form (5) of metric: Eq. (33) with Eq. (37), Eq. (39) with Eq. (37) }see also Tables I and II (particular solutions)

From Eq. (A3)

$$R=-2a^{-4}\left[4\frac{a''}{a}+\frac{8}{r}\frac{a'}{a}+2\frac{a'}{a}\frac{c'}{c}+\frac{2}{r}\frac{c'}{c}+\frac{c''}{c}\right]+12c^{-2}\left[3\left(\frac{\dot a}{a}\right)^2+\frac{\ddot a}{a}-\frac{\dot a}{a}\frac{\dot c}{c}\right],\tag{A4}$$

and

$$C=a^{-4}\left[2\frac{a''}{a}-\frac{2}{r}\frac{a'}{a}-6\left(\frac{a'}{a}\right)^2+4\frac{a'}{a}\frac{c'}{c}+\frac{1}{r}\frac{c'}{c}-\frac{c''}{c}\right].\tag{A5}$$

APPENDIX B

We express the eigenvalues (9a) and (9b) of the Ricci tensor by the components of the curvature tensor

$$\lambda_{1,2}=-(\alpha+\beta+\xi)\pm\Delta^{1/2},$$
$$\lambda_{3,4}=-(\alpha+\beta+\eta),\tag{B1}$$
$$\Delta:=(\alpha-\beta)^2+4\gamma\delta.$$

Thus, we have the cases

Eigenvalues of $R_\alpha{}^\beta$	Necessary and sufficient condition
two double	$\Delta=0,\quad \xi\neq\eta,$
one quadruple	$\Delta=0,\quad \xi=\eta,$
one triple, one single	$\Delta\neq 0,\quad \Delta=(\xi-\eta)^2.$

If $ds^{*2}=ds^2\exp 2\phi$, where ds^2 stands for the line element (4), a straightforward calculation using Eq. (A2) leads to

$$\eta^*-\xi^*=\Lambda^2(\eta-\xi)-2\Lambda e^{-w}z^{-2}\times[\Lambda_{,uv}-z^{-1}z_{,u}\Lambda_{,v}-z^{-1}z_{,v}\Lambda_{,u}],$$
$$\Delta^*=4\gamma^*\delta^*,\qquad \Lambda:=e^{-\phi},\tag{B2}$$

and

$$\gamma^*=\Lambda^2\gamma-\Lambda e^{-w}z^{-2}[\Lambda_{,vv}-2z^{-1}z_{,v}\Lambda_{,v}-w_{,v}\Lambda_{,v}],\tag{B3}$$
$$\delta^*=\Lambda^2\delta-\Lambda e^{-w}z^{-2}[\Lambda_{,uu}-2z^{-1}z_{,u}\Lambda_{,u}-w_{,u}\Lambda_{,u}].\tag{B4}$$

Obviously, $\Delta=0$ and/or $\xi-\eta=0$ do give $\Delta^*\neq 0$ and/or $\xi^*-\eta^*\neq 0$ in general. Similarly, in general the condition for a triple eigenvalue is not preserved by the conformal mapping.

APPENDIX C

From Eq. (A2) with $\exp w=2(u-v)^{-2}$ we obtain as the condition for a triple eigenvalue of $R_\alpha{}^\beta$ for a conformally flat s.s. metric

$$4\kappa^2+4(u-v)^2[\kappa\kappa_{,uv}+\kappa_{,u}\kappa_{,v}]+2(u-v)^3[\kappa_{,uu}\kappa_{,v}-\kappa_{,vv}\kappa_{,u}]+(u-v)^4[(\kappa_{,uv})^2-\kappa_{,uu}\kappa_{,vv}]=0,\tag{C1}$$

$\kappa:=z^{-1}=[A+B]^{-1}$.

In addition to have $\Delta\neq 0$, the inequality

$$[\kappa_{,vv}-2\kappa_{,v}(u-v)^{-1}][\kappa_{,uu}+2\kappa_{,u}(u-v)^{-1}]\neq 0,\tag{C2}$$

is required to hold.

APPENDIX D

With $w=-2\ln z$ and Eq. (1) the system (11a) and (11b) reduces to

$$zz_{,uv}+\tfrac{1}{2}z_{,u}z_{,v}+\tfrac{1}{4}=0,\tag{D1a}$$
$$C-\tfrac{3}{2}z^{-2}(1+2z_{,u}z_{,v})=0.\tag{D1b}$$

Introduction of the new dependent variable y by $y:=z^{3/2}$ leads to the nonlinear wave equation

$$y_{,uv}+\tfrac{3}{8}y^{-1/3}=0.\tag{D2}$$

The solutions of the form $y=y(u-v)$ are easily obtained. Three subcases arise:

(a) $z=\pm 2^{-1/2}(u+v)+\kappa_2$,

(b) $\pm 2^{-1/2}(u-v)=\kappa_2+[z(z+\kappa_1)]^{1/2}-\kappa_1\sinh^{-1}(z/\kappa_1)^{1/2}$,

(c) $\pm 2^{-1/2}(u-v)=\kappa_2+[z(z+|\kappa_1|)]^{1/2}+|\kappa_1|\cosh^{-1}(z/\kappa_1)^{1/2}$,

where κ_1, κ_2 are constants.

The identification $\kappa_1=-4ab$, $\bar\rho=2^{-1/2}(u-v)$, $z=\rho$ shows that case (c) corresponds to Eqs. (44) and (45) of Sec. V. In this case C cannot have the form $z^{-2}(1-\sigma_0)$. Case (a) corresponds to Minkowski space.

[1]H. Takeno, "The Theory of Spherically Symmetric Space–Times" (Revised edition), Reports of the Research Institute for Theoretical Physics No. 5, Hiroshima University (1966), Chap. 1.
[2]This definition of spherical symmetry is mathematically but not physically sufficient (cf. s.s. space–times of Takeno type S_2).
[3]G.D. Birkhoff, *Relativity and Modern Physics* (Harvard University, Cambridge, Mass., 1923), p. 255; J.T. Jebson, Arch. Mat. Astron. Fys. **15**, 1 (1921); W. Alexandrow, Ann. Phys. (Leipzig) **72**, 141 (1923).
[4]H. Goenner, Commun. Math. Phys. **16**, 34 (1970).
[5]P. Havas, Gen. Rel. Grav. **8**, 631 (1977).
[6]P. Havas and H. Goenner, "Theories of Gravitation with Higher-Order Field Equations and Birkhoff's Theorem," Paper presented at the Third Latin-American Symposium for General Relativity and Gravitation, Mexico City, August 1978 (in press).
[7]One of us (H.G.) wishes to thank Hans Buchdahl for pointing this out to him.
[8]R. Bach, Math. Z. **9**, 110 (1921); C. Lanczos, Ann. Math. **39**, 842 (1938).
[9]This definition of the conformal invariant C differs from the one used in Ref. 4 by a factor $\frac{1}{2}$.
[10]H. Takeno, Ref. 1, pp. 7 and 20.
[11]J. Plebański and J. Stachel, J. Math. Phys. **9**, 269 (1967).
[12]However, one could study the algebraic type of $R_{\alpha\beta\gamma\delta}$, which in general differs from that of $C_{\alpha\beta\gamma\delta}$.
[13]E. Goursat, *Cours d'analyse mathématique* (Gauthier–Villars, Paris 1927), 4th edition, Vol. 3, p. 71.
[14]G. Nordström, Phys. Z. **13**, 1126 (1912); A. Einstein and A.D. Fokker, Ann. Physi. (Leipzig) **44**, 321 (1914); the solution (13) seems to have been given first by W.-T. Ni, Phys. Rev. Lett. **35**, 319 and 1748 (1975), who also showed that it is a solution of Yang's theory of gravitation.
[15]I. Robinson. Bull. Acad. Pol. Sci. 7, 351 (1959); D. Lovelock, Commun. Math. Phys. **5**, 205 (1967).
[16]R.C. Tolman, *Relativity, Thermodynamics and Cosmology* (Clarendon, Oxford, 1966), Sec. 163, p. 413.
[17]N.N. Lebedev, I.P. Shalskaya, and Y.S. Uflyand, *Problems of Mathematical Physics* (Prentice Hall, Englewood Cliffs, 1965), Chap. 2.
[18]N.S. Koshlyakov, M.M. Smirnov, and E.B. Gliner, *Differential Equations of Mathematical Physics* (North-Holland, Amsterdam, 1964), Chap. 1.
[19]V.A. Wynne and G.H. Derrick, Nuovo Cimento B **15**, 181 (1973). All s.s.s. of $R = 0$ given there except Nos. (vi) and (vii) belong to the class of conformally flat solutions (13).
[20]A.H. Taub, Ann. Math. **53**, 472 (1951).

Conformal coupling of gravitational wave field to curvature

L.P. Grishchuk [a)]
Shternberg Astronomical Institute, Moscow 117234, USSR
W.K. Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, California 91125
V.M. Yudin
Shternberg Astronomical Institute, Moscow 117234, USSR

(Received 20 November 1979; accepted for publication 7 December 1979)

Conformal properties of the equations for weak gravitational waves in a curved space–time are investigated. The basic equations are derived in the linear approximation from Einstein's equations. They represent, in fact, the equations for the second-rank tensor field $h_{\alpha\beta}$, restricted by the auxiliary conditions $h_\alpha{}^\beta{}_{;\alpha} = 0$, $h \equiv \gamma_{\alpha\beta} h^{\alpha\beta} = 0$, and embedded into the background space–time with the metric tensor $\gamma_{\alpha\beta}$. It is shown that the equations for $h_{\alpha\beta}$ are not conformally invariant under the transformations $\hat{\gamma}_{\alpha\beta} = e^{2\sigma}\gamma_{\alpha\beta}$ and $\hat{h}_{\alpha\beta} = e^{\sigma} h_{\alpha\beta}$, except for those metric rescalings which transform the Ricci scalar $\hat{R}$ of the original background space–time into $e^{-2\sigma}R$, where R is the Ricci scalar of the conformally related background space–time. The general form of the equations for $h_{\alpha\beta}$ which are conformally invariant have been deduced. It is shown that these equations cannot be derived in the linear approximation from any tensor equations which generalize the Einstein equations.

I. INTRODUCTION

Conformal symmetry appears to be important in contemporary physics (see, for example, Refs. 1 and 2). Usually, the invariance with respect to the 15-parameter Lie group of conformal transformations C_0 which generalizes the Poincaré group is meant by a conformal symmetry.[3,4] A more general kind of conformal transformation is a conformal rescaling of the metric tensor: $\tilde{g}_{\mu\nu} = e^{-2\sigma} g_{\mu\nu}$. Conformal rescalings are mostly applied to the equations that are written down in a generally covariant form. The conformal transformations, viewed as conformal rescaling, correspond to that particular case in which a flat space–time Minkowski metric transforms into another flat space–time metric. Even this simplest kind of conformal symmetry which is important for high-energy physics may have some relevance to gravity as well (not curvature!) since a conformal transformation may be interpreted as a transformation to a constantly accelerated frame of reference.[3] Therefore, a property of a physical system with respect to C_0 may describe the behavior of the system in a constant homogeneous gravitational field.

It has been known already for a long time[3,5] that some basic equations of theoretical physics, among them the equations for massless fields, are invariant not only with respect to C_0 but also with respect to the group C_g of conformal rescalings. For instance, the field equations for massless fields with integer spins remain unchanged under a replacement of $g_{\mu\nu}$ and field variables $\varphi_{\alpha\beta\cdots\nu}$ according to the rule $\tilde{g}_{\mu\nu} = e^{-2\sigma} g_{\mu\nu}$, $\tilde{\varphi}_{\alpha\beta\cdots\nu} = e^{-\sigma(s-1)} \varphi_{\alpha\beta\cdots\nu}$, where s is the spin of the field. It is important to notice that the field variables transform with different powers of the conformal factor $e^{-2\sigma}$, depending on the spin of the field. For the scalar field it is $\tilde{\varphi} = e^{\sigma}\varphi$, and for the Maxwell equations ($s = 1$) $\tilde{A}_\alpha = A_\alpha$ or $\tilde{F}_{\alpha\beta} = F_{\alpha\beta}$. For the gravitational field ($s = 2$) the conformal invariance is usually referred to the vacuum Bianchi identities with the Weyl tensor transforming as $\tilde{C}_{\alpha\beta\mu\nu} = e^{-\sigma} C_{\alpha\beta\mu\nu}$.

Conformal symmetry of the field equations with respect to C_g is important from the physical point of view since it describes the particular way of coupling of the physical system to the external gravitational field (curvature).[6] The role of conformal invariance in the context of quantum field theory in curved space–time has been emphasized many times.[7,8] It was shown in Ref. 9 that the Einstein linearized equations for weak gravitational waves in nonvacuum conformally flat metrices do not transform into the usual flat space–time wave equations under the conformal transformation of the metric tensor and gravitational-wave variables. Thus, graviton creation in the early Universe is possible[9] while other massless particles such as photons, neutrinos, and gravitinos (spin $s = 3/2$ massless particle) cannot be created. (For the properties of the pure supergravity theory in this context see Ref. 10.) This fact seems to be fundamental enough in order to see to which extent it is inevitable.

The purpose of this article is to investigate the conformal property of the gravitational-wave equations in more detail. In particular, we are trying to find such equations which could be conformally invariant.

It is necessary to clarify the difference between our approach and that which was used in other works, devoted to conformal gravitation on the classical and quantum levels.[11-15] We intend to treat the gravitational-wave variables on the same footing as all other fields embedded in a curved space–time. It means that under conformal rescaling the field variables should transform according to their spin $s = 2$ weight. It might be $\tilde{C}_{\alpha\beta\mu\nu} = e^{-\sigma} C_{\alpha\beta\mu\nu}$ according to Penrose's suggestion or $\tilde{h}_{\alpha\beta} = e^{-\sigma} h_{\alpha\beta}$ in a linearized approximation to the Einstein equations which is considered here.

[a)] Supported in part by the Ministry of Higher Education, USSR, and by the National Science Foundation, USA (AST76-80801 A02).

(They are obviously consistent in the same way as the conformal transformation rules are consistent for electrodynamical field components $\widehat{F}_{\alpha\beta} = F_{\alpha\beta}$ and electrodynamical potentials $\widehat{A}_\alpha = A_\alpha$.) The linearized version of Einstein's equations provides a natural framework for treating spin $s = 2$ fields in an external gravitational field since these equations describe, in fact, the second-rank symmetric tensor field embedded in a curved space–time. On the other hand, the works[11-15] are concerned with the action and the field equations which are invariant under conformal rescaling of the metric tensor $g_{\alpha\beta}$ and some scalar function. If the components of $g_{\alpha\beta}$ are to be interpreted as spin $s = 2$ field variables, then they transform according to the wrong rule; this rule includes the factor $e^{-2\sigma}$ instead of $e^{-\sigma}$. The same rule is prescribed for the second-rank tensor $h_{\alpha\beta}$ at the linearized level. Although this kind of symmetry may be useful for some purposes, it is certainly not what is meant by conformal invariance for other massless field equations.

In Sec. II from Einstein's equations we derive the basic equations for graviational-wave perturbations $h_{\alpha\beta}$. These equations have the same form both in vacuum space–time and in space–time filled with matter. We introduce also the usual auxiliary conditions $h_\alpha{}^\beta{}_{;\beta} = 0$, $h = 0$ which are similar to that used in a flat space–time for separating spin $s = 2$ states.[16] We investigate the conformal properties of the field equations and show that they are not conformally invariant except for those transformations which transform the Ricci scalar $\widehat{R}$ of the original background space–time into $e^{-2\sigma}R$, where R is the Ricci scalar of the conformally related background space–time. Although we believe that the chosen field equations, the auxiliary conditions, and the transformation law for $h_{\alpha\beta}$ are well motivated, one should not think that the conformal noninvariance is a consequence of these assumptions. The formulas presented in the Appendix show that any other choice of the auxiliary conditions (if any) and of a transformation law cannot improve the situation. Moreover, there is an indication that the prescribed auxiliary conditions and the transformation law emerge in a natural way under an attempt to make the basic equations conformally invariant. Having proved conformal noninvariance of the equations derived from Einstein's equations, we were interested in the formulation of the equations for spin $s = 2$ field which are conformally invariant.

Essentially, we look for a conformally invariant second-order differential operator which acts on a symmetric second-rank tensor field restricted by some auxiliary conditions. In other words, we generalize the flat space–time equations in such a way that the coupling of the tensor field to curvature is conformally invariant. To clarify the method used, we start from the simplified problem of finding conformally invariant equations for a scalar field φ (see Sec. III). It is known that conformal coupling of the φ field to curvature can be represented by the equation

$$\varphi_{;\alpha}{}^{;\alpha} - \frac{R}{6}\varphi = 0. \tag{1}$$

For the sake of generality we take into account some other fields to which the φ field can be coupled (other than the curvature) and which transform according to definite rules under a conformal rescaling. It is shown that the most general conformally invariant coupling to curvature is expressed by Eq. (1) while coupling to other fields can also be conformally invariant and then Eq. (1) contains additional terms. The same method of searching for conformally invariant equations was applied to spin $s = 2$ field (Sec. IV). Since in this case the equations are more complicated, we restricted the search to the coupling of this field to curvature. The general form of such conformally invariant equations is deduced. It is seen that these equations could not be derived from Einstein's equations in the linearized approximation. The next step is to look for exact tensorial equations from which conformally invariant equations can follow in the linearized approximation (Sec. V). If such a theory existed it might be interesting to investigate it and compare its predictions with the predictions of Einstein's theory. Quite surprisingly, it turns out that such an exact theory does not exist, at least within those restrictions which were imposed on it. It is also shown that there exists a conformally invariant equation describing the coupling of the second-rank tensor field to curvature and some additional scalar field. The possibility of finding an exact theory which would yield this equation in the linearized limit is not clear. In conclusion (Sec. VI), we give a discussion of the presented results.

II. CONFORMAL NONINVARIANCE OF THE LINEARIZED EINSTEIN EQUATIONS

First we will derive the equations which we will be working with. Let us start from the vacuum Einstein equations with the cosmological term

$$R_{\mu\nu} = \lambda g_{\mu\nu}.$$

Assume that $g_{\mu\nu} = \gamma_{\mu\nu} + h_{\mu\nu}$, where $\gamma_{\mu\nu}$ is the metric tensor of a background space–time, and assume that the background field equations $R^{(0)}_{\mu\nu} = \lambda\gamma_{\mu\nu}$ are fulfilled. The linearized equations $-(\gamma_{\mu\alpha}\delta R_\nu{}^\alpha + \gamma_{\nu\alpha}\delta R^\alpha{}_\mu) = 0$ or $-2\delta R_{\mu\nu} + 2\lambda h_{\mu\nu} = 0$ both lead to the same equation

$$\begin{aligned} h_{\mu\nu;\alpha}{}^{;\alpha} &- 2R_{\mu\alpha\beta\nu}h^{\alpha\beta} - (h_\mu{}^\alpha - \tfrac{1}{2}\delta_\mu{}^\alpha h)_{;\alpha;\nu} \\ &- (h_\nu{}^\alpha - \tfrac{1}{2}\delta_\nu{}^\alpha h)_{;\alpha;\mu} = 0, \end{aligned} \tag{2}$$

where, as usual, all operations are performed in background space–time. $R_{\mu\alpha\beta\nu}$ denotes the background curvature tensor; here and below we will not especially mark the background quantities. Equation (2) can be regarded as a generally covariant equation for a symmetric second-rank tensor field $h_{\mu\nu}$.

Choose the solutions to Eq. (2) which are subject to the auxiliary conditions

$$h \equiv h_{\alpha\beta}\gamma^{\alpha\beta} = 0, \tag{3}$$

$$h_\mu{}^\alpha{}_{;\alpha} = 0. \tag{4}$$

In analogy to what is known for analogous equations in flat space–time, Eqs. (3) and (4) can be interpreted as the necessary conditions for removing the spin $s = 0$, and $s = 1$ contributions to $h_{\mu\nu}$. For these solutions Eq. (2) takes the form

$$h_{\mu\nu;\alpha}{}^{;\alpha} - 2R_{\mu\alpha\beta\nu}h^{\alpha\beta} = 0. \tag{5}$$

[Of course, Eq. (2) can be reduced to Eq. (5) under the simpler condition

$$\chi_\mu \equiv (h_\mu{}^\alpha - \tfrac{1}{2}\delta_\mu{}^\alpha h)_{;\alpha} = 0. \tag{4'}$$

We are going to work with Eqs. (3)–(5) but before this let us see which part takes the solutions restricted by Eqs. (3) and (4) (we will call them spin $s = 2$ solutions), among all solutions to Eq. (2).

It can be easily checked using the background field equations that if $h^*_{\mu\nu}$ is a solution to Eq. (2) then

$$h_{\mu\nu} = h^*_{\mu\nu} + \xi_{\mu;\nu} + \xi_{\nu;\mu} \tag{6}$$

for arbitrary ξ_μ is also a solution to Eq. (2). This fact is frequently referred to as a gauge freedom.[17–19] For any given solution $h^*_{\mu\nu}$ one can find a vector ξ_μ which will map this solution into the class of solutions restricted by Eqs. (3) and (4).[20] Therefore, the spin $s = 2$ solutions represent in a sense all the solutions to Eq. (2). Moreover, the spin $s = 2$ solutions map into themselves by the gauge transformations with $\xi_{\mu;\nu}{}^{;\nu} = 0$, $\xi^\nu{}_{;\nu} = 0$. The remaining gauge freedom can be used to impose the initial conditions $(h_{\mu\nu}u^\nu)_{|\Sigma} = 0$, $(h_{\mu\nu}u^\nu)_{;\alpha}n^\alpha{}_{|\Sigma} = 0$ on some hypersurface Σ with the normal vector n^α, where u^α is a vector field. It was shown in Ref. 19 that the sufficient condition for $h_{\mu\nu}u^\nu$ to be equal to zero not only on Σ but also off Σ is the existence of u^α obeying the equation

$$u_{\mu;\nu} = u_\mu a_\nu + b\gamma_{\mu\nu}, \tag{7}$$

where a_ν and b are arbitrary vector and scalar fields, respectively. In flat space–time such a vector u^α does exist and therefore all solutions to Eq. (2) can be mapped into a class of solutions which fulfill Eqs. (3) and (4) and

$$h_{\mu\nu}u^\nu = 0 \tag{8}$$

(TT gauge, according to Ref. 17). We will call this class of solutions the spin $s = 2$ solutions with definite helicity.

As for the curved space–time in general Eq. (7) is not integrable, except for a certain class of background metrics, among them the important case of conformally flat metrics.[21] So in these cases the spin $s = 2$ solutions with definite helicity represent all solutions to the wave equations, similarly to what we have in flat space–time.

Let us turn now to gravitational-wave equations in a nonvacuum space–time. The Einstein equations $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \lambda g_{\mu\nu} = T_{\mu\nu}$ in the linearized approximation $-\gamma_{\mu\alpha}(\delta R^\alpha_\nu - \frac{1}{2}\delta^\alpha{}_\nu\delta R) - \gamma_{\nu\alpha}(\delta R_\mu{}^\alpha - \frac{1}{2}\delta^\alpha{}_\mu\delta R) = -(\gamma_{\mu\alpha}\delta T^\alpha{}_\nu + \gamma_{\nu\alpha}\delta T^\alpha{}_\mu)$ have the following form:

$$\begin{aligned} &h_{\mu\nu;\alpha}{}^{;\alpha} - 2R_{\mu\alpha\beta\nu}h^{\alpha\beta} - \chi_{\mu;\nu} - \chi_{\nu;\mu} \\ &\quad + \gamma_{\mu\nu}(-h_{\alpha\beta}R^{\alpha\beta} + \chi_{\alpha;}{}^\alpha - \tfrac{1}{2}h_{;\alpha}{}^{;\alpha}) \\ &\quad = -(\gamma_{\mu\alpha}\delta T^\alpha{}_\nu - \gamma_{\nu\alpha}\delta T^\alpha{}_\mu). \end{aligned} \tag{9}$$

Equation (9), similarly to Eq. (2), is gauge invariant. If $h^*_{\mu\nu}$ and $\delta T^*_{\mu\nu}$ are a solution to Eq. (9), then $h_{\mu\nu}$, defined by Eq. (6), and $\delta T_{\mu\nu}$, defined by

$$\delta T_{\mu\nu} = \delta T^*_{\mu\nu} + T_\mu{}^\alpha\xi_{\alpha;\nu} + T_\nu{}^\alpha\xi_{\alpha;\mu},$$

are also a solution to Eq. (9). For any given solution one can find ξ_α which will map this solution into a class of solutions subjected to Eq. (4') (see Ref. 20). Moreover, there still remains some gauge freedom $\xi_{\mu;}{}^\alpha{}_{;\alpha} + \xi^\alpha R_{\alpha\mu} = 0$ which can be used to impose the zero-initial conditions for h, $h_{;\alpha}n^\alpha$, $h_{\mu\nu}u^\nu$, $(h_{\mu\nu}u^\nu)_{;\alpha}n^\alpha$ on some initial hypersurface Σ. However, in the general case, h and $h_{\mu\nu}u^\nu$ will not vanish off Σ.

Equation (9) includes metric perturbations as well as perturbations of $T_{\mu\nu}$. It is clear, however, that the source-free gravitational-wave perturbations should be associated in some sense with the perturbations of the gravitational field itself and not the matter. We shall define the spin $s = 2$ solutions in a nonvacuum background as a class of solutions for which Eqs. (3) and (4) are valid together with $\delta T_\alpha{}^\beta = 0$. The last condition reduces Eq. (9) to the form of Eq. (2) and $\delta T_\alpha{}^\beta = 0$ together with Eq. (4') reduces it to the equation

$$h_{\mu\nu;\alpha}{}^{;\alpha} - 2R_{\mu\alpha\beta\nu}h^{\alpha\beta} = 0, \tag{10}$$

which is exactly the form of Eq. (5) (but $R_{\mu\nu} - \lambda g_{\mu\nu} \neq 0$ now). The fact that Eqs. (5), (10), (3), and (4) formally coincide in a vacuum and in a nonvacuum background corresponds to an intuitive feeling that a free gravitational wave should be "sensitive" to a curvature in the same way, independently of what is the source of that curvature. The other argument in favor of Eqs. (10), (3), and (4) is that for those space–times (for instance, for Robertson–Walker background metrics) for which a unique decomposition of all perturbations into proper modes is possible, the tensorial (gravitational wave) modes obey these equations (cf. Ref. 22).

It is important that we impose the condition $\delta T_\alpha{}^\beta = 0$ (variation of $T_\alpha{}^\beta$ with mixed indices.) Other authors sometimes define the gravitational-wave perturbations as the set of conditions $\delta T_{\mu\nu} = 0$ or $\delta(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T) = 0$ together with Eqs. (3) and (4). These alternative equations differ from Eq. (10) and do not lead to tensorial proper modes in symmetric backgrounds, which we regard as an unsatisfactory drawback. As far as conformal invariance is concerned, these alternative equations are not conformally invariant.

Thus, we regard Eqs. (10), (3), and (4) as those whse conformal property should be investigated. We will mark by carets all quantities in an original space–time and apply the conformal transformation

$$\hat{\gamma}_{\mu\nu} = e^{2\sigma}\gamma_{\mu\nu}, \tag{11}$$

$$\hat{h}_{\mu\nu} = e^{\sigma}h_{\mu\nu}, \tag{12}$$

to careted Eqs. (10), $(\hat{3})$, and (4). Under the transformation (11) the Ricci scalar $\hat{R}$, Ricci tensor $\hat{R}_{\mu\nu}$, and Weyl tensor $\hat{C}^\alpha{}_{\mu\nu\beta}$ transform as follows:

$$\hat{R} = e^{-2\sigma}[R - 6(\sigma_{\alpha;}{}^\alpha + \sigma_\alpha\sigma^\alpha)], \tag{13}$$

$$\hat{R}_{\mu\nu} = R_{\mu\nu} - 2\sigma_{\mu;\nu} + 2\sigma_\mu\sigma_\nu - (\sigma_{\alpha;}{}^\alpha + 2\sigma_\alpha\sigma^\alpha)\gamma_{\mu\nu}, \tag{14}$$

$$\hat{C}^\alpha{}_{\beta\mu\nu} = C^\alpha{}_{\beta\mu\nu}, \quad \sigma_\alpha \equiv \sigma_{,\alpha}. \tag{15}$$

Recall also the relation between $C^\alpha{}_{\beta\mu\nu}$ and the curvature tensor $R^\alpha{}_{\beta\mu\nu}$:

$$\begin{aligned} C^\alpha{}_{\beta\mu\nu} = R^\alpha{}_{\beta\mu\nu} &+ \tfrac{1}{2}(\delta^\alpha{}_\nu R_{\beta\mu} - \delta^\alpha{}_\mu R_{\beta\nu} + g_{\beta\mu}R^\alpha{}_\nu \\ &- g_{\beta\nu}R^\alpha{}_\mu) - \frac{R}{6}(\delta^\alpha{}_\nu g_{\beta\mu} - \delta^\alpha{}_\mu g_{\beta\nu}), \end{aligned}$$

which helps to restore the transformation rule for $R^\alpha{}_{\beta\mu\nu}$. One can see that $\hat{h} = e^{-\sigma}h$ and $\hat{h}_\mu{}^\nu{}_{;\nu} = e^{-\sigma}(h_\mu{}^\nu{}_{;\nu}$

$+ 3h_\mu{}^\nu\sigma_\nu - h\sigma_\mu)$. Thus, Eq. (3) is conformally invariant. To keep Eq. (4) conformally invariant as well one needs

$$h_\mu{}^\nu\sigma_\nu = 0 = \hat{h}_\mu{}^\nu\sigma_\nu. \tag{16}$$

This condition is analogous to one which keeps the electrodynamic Lorentz gauge $A_{\alpha;}{}^\alpha = 0$ conformally invariant.

With Eqs. (3), (4), and (16) valid the left-hand side of Eq. (10) transforms as follows (one can consult formulas in the Appendix with $k = 1$):

$$e^{-\sigma}\,[h_{\mu\nu;\alpha}{}^{;\alpha} - 2R_{\mu\alpha\beta\nu}h^{\alpha\beta} - (\sigma_{\alpha;}{}^\alpha + \sigma_\alpha\sigma^\alpha)h_{\mu\nu}] = 0.$$

It follows that Eq. (10) is not conformally invariant, unless

$$\sigma_{\alpha;}{}^\alpha + \sigma_\alpha\sigma^\alpha = 0. \tag{17}$$

Equation (17) severely restricts conformal transformations with respect to which set of Eqs. (10), (3), and (4) is conformally invariant. For a given $\hat{h}_\mu{}^\nu$, Eq. (16) also restricts σ; however, the origin and the meaning of the restrictions (17) and (16) are completely different. Equation (17) represents, so to say, the "genuine" noninvariance of the wave equation (10), while Eq. (16) is a necessary condition for keeping the auxiliary condition conformally invariant. Conformal invariance of the wave equations and auxiliary conditions for potentials seems to be a more significant property than just a conformal invariance of the wave equations in terms of field components.

Equation (16), together with Eqs. (3) and (4), selects the spin $s = 2$ solutions with definite helicity as those which could be conformally transformed. In general, for a given σ_ν, Eq. (16) restricts $\hat{h}_\mu{}^\nu$; however, in some cases, the conditions (16), (3), and (4) can be achieved at the expense of the gauge freedom and therefore do not, in fact, restrict the transformed solutions. Again, this is true for the Friedmann universes—the case which we are most interested in. For example, in a background metric

$$ds^2 = a^2(\eta)(d\eta^2 - dx^2 - dy^2 - dz^2), \tag{18}$$

the gravitational-wave perturbations obey the auxiliary conditions (3), (4), and (8), where $u^\nu = (1/a,0,0,0)$. The metric (18) transforms into the flat space–time metric by $\sigma = \ln a$ and therefore Eq.(16) is automatically fulfilled.

Notice, that gravitational-wave equations and auxiliary conditions in Minkowski space–time are conformally invariant with respect to C_0. Really, a solution to Eq. (17) in Minkowski space–time is

$$\sigma = -\ln(1 + 2a_\alpha x^\alpha + a^2 x_\alpha x^\alpha), \tag{19}$$

where $a^2 = a_\alpha a^\alpha$, and a_α are constants. Conformal rescaling with the σ factor (19) corresponds to a group of conformal transformations C_0. This rescaling transforms the Minkowski line element $ds^2{}_m$ into the line element $ds^2 = e^{2\sigma}ds^2_m$:

$$ds^2 = \frac{1}{(1 + 2a_\alpha x^\alpha + a^2 x_\alpha x^\alpha)^2}\,(c^2dt^2 - dx^2 - dy^2 - dz^2). \tag{20}$$

Due to the gauge freedom in Minkowski space–time one can introduce Eqs. (3) and (4) and reduce the field equations to $h_{\mu\nu,\alpha}{}^{,\alpha} = 0$. The gauge freedom which remains is described by $\xi_{\mu,\alpha}{}^{,\alpha} = 0$, $\xi^\alpha{}_{,\alpha} = 0$. For σ, given by Eq. (19), we introduce $u^\nu = a^\nu + a^2x^\nu \equiv -\frac{1}{2}e^{-\sigma}\sigma^\nu$. Note that $(h_{\mu\nu}u^\nu)_{,\alpha}{}^{,\alpha} = 0$. The remaining gauge freedom can be used to impose $(h_{\mu\nu}u^\nu)_{|\Sigma} = 0$ $(h_{\mu\nu}u^\nu)_{,\alpha}n^\alpha{}_{|\Sigma} = 0$, and hence $h_{\mu\nu}u^\nu = 0 = h_{\mu\nu}\sigma^\nu$. Thus, the field equations, the auxiliary conditions, and the spin $s = 2$ solutions with definite helicity in the conformally related space–times (20) transform into each other.

III. COUPLING OF A SCALAR FIELD TO CURVATURE AND OTHER FIELDS

The aim of this section is to find the covariant and conformally invariant, second order, linear, homogeneous differential equations for a scalar field φ. As far as this equation will be formulated in a curved space–time it will describe the coupling of φ to curvature, but, for the sake of generality, we also allow for coupling of φ to other (nongravitational) scalar, vector, and tensor fields, which could be present in the given space–time.

The general form of the equation is

$$\hat{C}^{\alpha\beta}\hat{\varphi}_{\alpha\hat{;}\beta} + \hat{B}^\alpha\hat{\varphi}_\alpha + \hat{A}\hat{\varphi} = 0, \tag{21}$$

where $\varphi_\alpha \equiv \varphi_{,\alpha}$; A, B^α, and $C^{\alpha\beta}$ are some scalar, vector, and tensor fields, respectively, $C^{\alpha\beta} = C^{\beta\alpha}$, since $\varphi_{\alpha;\beta} = \varphi_{\beta;\alpha}$. We make two additional assumptions also. Firstly, we assume that the coupling to the external gravitational field can be realized only through the metric tensor, the curvature tensor, and their different algebraic combinations, so that the gravitational part of coefficients, A, B^α, and $C^{\alpha\beta}$ should be constructed from them. Secondly, we note that the first term in Eq.(21) contains a piece $\hat{\varphi}_{\alpha\hat{;}}{}^\alpha$ among all other possible contributions. We want this piece to be present in the original and in the transformed equation.

Under a conformal rescaling of the metric tensor the gravitational part of the coefficients A, B^α, and $C^{\alpha\beta}$ transform according to the law which is basically determined by Eqs. (13) and (15). As for the transformation laws for the nongravitational contributions to A, B^α, and $C^{\alpha\beta}$ we will derive them from the condition of conformal invariance of Eq. (21).

First we will transform the φ field and its derivatives in Eq. (21). Under the transformation rules

$$\hat{\gamma}_{\mu\nu} = \Omega^2\gamma_{\mu\nu}, \quad \hat{\varphi} = \Omega^{-1}\varphi, \tag{22}$$

Eq. (21) takes the form

$$\begin{aligned}\Omega^{-1}\Big\{&\hat{C}^{\alpha\beta}[\varphi_{\alpha;\beta} - 2\Omega^{-1}(\varphi_\alpha\Omega_\beta + \Omega_\alpha\varphi_\beta)\\ &+ \Omega^{-1}\gamma_{\alpha\beta}\varphi_\sigma\Omega^\sigma - \Omega^{-1}\Omega_{\alpha;\beta}\varphi\\ &+ 4\Omega^{-2}\Omega_\alpha\Omega_\beta\varphi - \Omega^{-2}\gamma_{\alpha\beta}\Omega_\sigma\Omega^\sigma\varphi\,]\\ &+ \hat{B}^\alpha\Big(\varphi_\alpha - \frac{\Omega_\alpha}{\Omega}\varphi\Big) + \hat{A}\varphi\Big\} = 0.\end{aligned} \tag{23}$$

For Eq. (21) to be conformally invariant, we need the left-hand side of Eq. (23) to be equal to the left-hand side of Eq. (21) (without "carets") multiplied by Ω in some power: Ω^n. Since $\hat{C}^{\alpha\beta}\hat{\varphi}_{\alpha\hat{;}\beta}$ includes $\hat{\gamma}^{\alpha\beta}\hat{\varphi}_{\alpha\hat{;}\beta}$ and this term transforms as $1/\Omega^3(\gamma^{\alpha\beta}\varphi_{\alpha;\beta} + \cdots)$, we want $n = -3$. Comparing the coefficients in front of $\varphi_{\alpha;\beta}$, $\varphi_{;\alpha}$, and φ in between Eqs. (23) and (21) (without "carets") multiplied by Ω^{-3}, one can obtain the transformation laws

$$\hat{C}_{\alpha\beta} = \Omega^2 C_{\alpha\beta}, \quad \hat{\gamma}^{\alpha\beta}\hat{C}_{\alpha\beta} \equiv \hat{C} = C \equiv \gamma^{\alpha\beta}C_{\alpha\beta}, \tag{24}$$

$$\hat{B}_\alpha = B_\alpha + 4\Omega^{-1}(\Omega_\beta C_\alpha{}^\beta - \tfrac{1}{4}\Omega_\alpha C), \tag{25}$$

$$\hat{A} = \Omega^{-2}A + \Omega^{-3}(\Omega^\alpha B_\alpha + \Omega_{\alpha;\beta}C^{\alpha\beta}). \tag{26}$$

Let us see what the gravitational contributions to $\hat{C}_{\alpha\beta}$ could be. This coefficient could contain $\hat{\gamma}_{\alpha\beta}$, $\hat{R}_{\alpha\beta}$, $\hat{R}_\alpha{}^\sigma \hat{R}_{\sigma\beta}$, etc. However, only $\hat{\gamma}_{\alpha\beta}$ can meet the transformation law (24). Hence,

$$\hat{C}_{\alpha\beta} = \hat{\gamma}_{\alpha\beta} + \hat{c}_{\alpha\beta}, \tag{27}$$

where $\hat{c}_{\alpha\beta}$ is some tensor field which does not depend on the metric, but is connected by a relation

$$\hat{c}_{\alpha\beta} = \Omega^2 c_{\alpha\beta} \tag{28}$$

in the conformally related space–time. Since the coefficient $\hat{B}_\alpha$ has an odd number of indices, it cannot have any gravitational contribution, so $\hat{B}_\alpha = \hat{b}_\alpha$, where $\hat{b}_\alpha$ is a vector field, independent of metric. Substituting $C_{\alpha\beta} = \gamma_{\alpha\beta} + c_{\alpha\beta}$ into Eq. (25), one obtains the transformation law for $\hat{b}_\alpha$:

$$\hat{b}_\alpha = b_\alpha + 4\Omega^{-1}\Omega_\beta(c^\beta{}_\alpha - \tfrac{1}{4}c\delta^\beta{}_\alpha). \tag{29}$$

The gravitational contributions to $\hat{A}$ could be of the form, $\hat{R}$, $\hat{R}^2$, $\hat{R}_{\alpha\beta}\hat{R}^{\alpha\beta}$, etc. However, only the first term can meet the condition $\hat{R} = \Omega^{-2}R + \cdots$ which is dictated by Eq. (26). Hence, the general form of $\hat{A}$ is

$$\hat{A} = a\hat{R} + \hat{m}, \tag{30}$$

where a is some function and $\hat{m}$ is a scalar field. Substituting Eq. (30) and $B_\alpha = b_\alpha$ and $C_{\alpha\beta} = \gamma_{\alpha\beta} + c_{\alpha\beta}$ into Eq. (26), we obtain the equation

$$a\Omega^{-2}(R - 6\Omega^{-1}\Omega_{\alpha;}{}^\alpha) + \hat{m}$$
$$= \Omega^{-2}[aR + m + \Omega^{-1}(\Omega^\alpha b_\alpha + \Omega_{\alpha;}{}^\alpha + \Omega_{\alpha;\beta}c^{\alpha\beta})],$$

which gives rise to the relations

$$-6a = 1, \quad \hat{m} = \Omega^{-2}m + \Omega^{-3}(\Omega^\alpha b_\alpha + \Omega_{\alpha;\beta}c^{\alpha\beta}). \tag{31}$$

Thus, the most general conformally invariant equation of the form of Eq. (21) is

$$\varphi_{;\alpha}{}^{;\alpha} - \tfrac{1}{6}R\varphi + c^{\alpha\beta}\varphi_{;\alpha;\beta} + b^\alpha\varphi_\alpha + m\varphi = 0, \tag{32}$$

where $c_{\alpha\beta}$, b_α, and m transform according to Eqs. (28), (29), and (31), respectively, unless all of them or some of them are equal to zero. Notice that if there exists any other conformally invariant equation for the φ field, it cannot contain the operator $\varphi_{;\alpha}{}^{;\alpha}$. The first two terms in Eq. (32) give the familiar equation for a scalar field in a curved space–time.

IV. CONFORMALLY INVARIANT EQUATIONS FOR A SECOND-RANK SYMMETRIC TENSOR FIELD

Having proved that the equations which follow from the Einstein equations for the second-rank tensor field $h_{\alpha\beta}$ are not conformally invariant, we will try now to find certain equations which are conformally invariant. For simplicitly, we will consider only the coupling of $h_{\alpha\beta}$ to the external gravitational field and not to other fields. The general strategy will be similar to the one used in the previous section. However, in the case of $h_{\alpha\beta}$ there is a complication related to the fact that $h_{\alpha\beta}$ should obey not only the field equations but also the auxiliary conditions.

We are looking for conformally invariant equations within the following class of equations. They should be covariant, second-order, homogeneous differential equations. Coefficients in these equations can contain the metric tensor $\gamma_{\alpha\beta}$, the curvature tensor, and their different algebraic combinations.

The general form of these equations can be written as follows;

$$\hat{F}_{\mu\nu}{}^{\sigma\alpha\beta\rho}\hat{h}_{\alpha\beta\hat{;}\sigma\hat{;}\rho} + \hat{P}_{\mu\nu}{}^{\alpha\beta\gamma}\hat{h}_{\alpha\beta\hat{;}\gamma} + \hat{U}^\alpha{}_{\mu\nu}{}^\beta\hat{h}_{\alpha\beta} = 0. \tag{33}$$

We also assume the validity of the auxiliary conditions

$$\hat{h} = 0, \tag{34}$$

$$\hat{h}_\mu{}^\nu{}_{\hat{;}\nu} = 0, \tag{35}$$

and will demand their conformal invariance. In fact, we should put $\hat{P}_{\mu\nu}{}^{\alpha\beta\gamma} = 0$ since a tensor with an odd number of indices cannot be constructed as an algebraic combination of metric tensor and curvature tensor.

Some properties of the symmetry of tensors $\hat{F}$ and $\hat{U}$ follow from the fact that $\hat{h}_{\alpha\beta}$ is a symmetric tensor and Eq. (33) is assumed to be symmetric with respect to the free indices μ and ν. To the same end, since

$$h_{\alpha\beta;\sigma;\rho} - h_{\alpha\beta;\rho;\sigma} = h_{\alpha\epsilon}R^\epsilon{}_{\beta\sigma\rho} + h_{\epsilon\beta}R^\epsilon{}_{\alpha\sigma\rho},$$

we may assume that the first term in Eq. (33) contains only the symmetric (with respect to σ,ρ) part of $h_{\alpha\beta\hat{;}\sigma\hat{;}\rho}$ while the antisymmetric part is included in the last term in Eq. (33). This assumption also determines, in part, the symmetry properties of the tensor $\hat{F}$.

As we know from Sec. II, under the transformations (11) and (12) the gauge conditions (34) and (35) transform into Eqs. (3) and (4), respectively, if Eq. (16) is satisfied. With the use of Eq. (16) one obtains the following transformation rules for $\hat{h}_{\mu\nu\hat{;}\alpha}$ and $\hat{h}_{\mu\nu\hat{;}\alpha\hat{;}\beta}$:

$$\hat{h}_{\mu\nu\hat{;}\alpha} = e^\sigma(h_{\mu\nu;\alpha} - h_{\mu\alpha}\sigma_\nu - h_{\nu\alpha}\sigma_\mu - h_{\mu\nu}\sigma_\alpha), \tag{36}$$

$$\hat{h}_{\mu\nu\hat{;}\alpha\hat{;}\beta} = e^\sigma(h_{\mu\nu;\alpha;\beta} + h_{\mu\gamma;\delta}Y_{\nu\alpha\beta}{}^{\gamma\delta} + h_{\nu\gamma;\delta}Y_{\mu\alpha\beta}{}^{\gamma\delta}$$
$$+ h_{\mu\gamma}Z_{\nu\alpha\beta}{}^\gamma + h_{\nu\gamma}Z_{\nu\alpha\beta}{}^\gamma + h_{\gamma\delta}V_{\mu\nu\alpha\beta}{}^{\gamma\delta}), \tag{37}$$

where

$$Y_{\nu\alpha\beta}{}^{\gamma\delta} \equiv \tfrac{1}{2}\gamma_{\alpha\beta}\delta_\nu{}^\gamma\sigma^\delta - \sigma_\nu(\delta_\alpha{}^\gamma\delta_\beta{}^\delta + \delta_\beta{}^\gamma\delta_\alpha{}^\delta)$$
$$- \delta_\nu{}^\gamma(\sigma_\alpha\delta_\beta{}^\delta + \sigma_\beta\delta_\alpha{}^\delta), \tag{38}$$

$$Z_{\nu\alpha\beta}{}^\gamma = \tfrac{1}{2}\delta_\nu{}^\gamma(-\sigma_{\alpha;\beta} + 3\sigma_\alpha\sigma_\beta - \gamma_{\alpha\beta}\sigma_\delta\sigma^\delta)$$
$$+ \delta_\alpha{}^\gamma(-\sigma_{\nu;\beta} + 3\sigma_\nu\sigma_\beta - \gamma_{\beta\nu}\sigma_\delta\sigma^\delta)$$
$$- \gamma_{\beta\nu}\sigma^\gamma{}_{;\alpha} + 2\sigma_\nu\sigma_\alpha\delta_\beta{}^\gamma, \tag{39}$$

$$V_{\mu\nu\alpha\beta}{}^{\gamma\delta} = 2\sigma_\mu\sigma_\nu\delta_\alpha{}^\gamma\delta_\beta{}^\delta. \tag{40}$$

The conformal invariance of Eq. (33) implies that after substituting Eqs. (12) and (37) and transformation laws for $\hat{F}$ and $\hat{U}$ into Eq. (33), the left-hand side of Eq. (33) transforms into the same expression (without "carets") multiplied by some power of e^σ. The power is determined from considerations similar to the ones used in Sec. III. The first term in Eq. (33) contains a piece $\hat{h}_{\mu\nu\hat{;}\sigma}{}^{\hat{;}\sigma}$. This arises from the contribution to $\hat{F}_{\mu\nu}{}^{\sigma\alpha\beta\rho}$ of the following form:

$$\hat{F}_{\mu\nu}{}^{\sigma\alpha\beta\rho} = \delta_\mu{}^\alpha\delta^\beta{}_\nu\hat{\gamma}^{\delta\rho}. \tag{41}$$

We want to save the term $h_{\mu\nu;\sigma}{}^{;\sigma}$ in the transformed equation. So we will sacrifice all other contributions to $\hat{F}_{\mu\nu}{}^{\sigma\alpha\beta\rho}$ if they do not fit into the transformation law for $\hat{h}_{\mu\nu;\sigma}{}^{;\sigma}$. Since

$$\hat{h}_{\mu\nu;\sigma}{}^{;\sigma} = e^{-\sigma}(h_{\mu\nu;\sigma}{}^{;\sigma} + \cdots),$$

we will demand that Eq. (33) be transformed into

$$e^{-\sigma}(F_{\mu\nu}{}^{\sigma\alpha\beta\rho}h_{\alpha\beta;\sigma;\rho} + U^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta}) = 0. \tag{42}$$

The right-hand side of Eqs. (36) and (37) already contains the factor e^{σ}; therefore, Eq. (42) implies that the transformation law for the tensors $\hat{F}$ and $\hat{U}$ must have the following form:

$$\hat{F}_{\mu\nu}{}^{\sigma\alpha\beta\rho} = e^{-2\sigma}(F_{\mu\nu}{}^{\sigma\alpha\beta\rho} + \cdots), \tag{43}$$

$$\hat{U}^{\alpha}{}_{\mu\nu}{}^{\beta} = e^{-2\sigma}(U^{\alpha}{}_{\mu\nu}{}^{\beta} + \cdots). \tag{44}$$

Unlike what was done in Sec. III, here, we could not obtain the transformation rules for $\hat{F}$ and $\hat{U}$ directly as a result of comparing the coefficients in front of $h_{\alpha\beta;\sigma;\rho}$, $h_{\alpha\beta;\sigma}$, and $h_{\alpha\beta}$ in both Eqs. (33) and (42). This is because there may be additional terms in these transformation rules which after multiplying them by $h_{\alpha\beta;\sigma;\rho}$ and $h_{\alpha\beta}$ can vanish due to the auxiliary conditions (3) and (4).

Let us see what are the possible contributions to $\hat{F}$ and $\hat{U}$ and whether they can satisfy the conditions (43) and (44). As far as the tensor $\hat{F}$ is concerned, the only contribution which meets the condition (43) is Eq. (41). All other contributions which can contain different combinations of Kronecker symbols, metric tensor, Ricci tensor, and curvature are not appropriate. Some of them, $\delta^{\sigma}{}_{\mu}\delta_{\nu}{}^{\rho}\hat{\gamma}^{\alpha\beta}$ or $\delta^{\sigma}{}_{\mu}\delta^{\alpha}{}_{\gamma}\hat{\gamma}^{\beta\rho}$, though they have the correct transformation property, do not play any role because they disappear due to eqs. (34) and (35). The other terms, like $\delta^{\sigma}{}_{\mu}\delta_{\nu}{}^{\rho}\hat{R}^{\alpha\beta}$ or $\gamma_{\mu\nu}\hat{R}^{\alpha\beta}\hat{R}^{\sigma\rho}$ or $\hat{R}_{\mu\nu}\hat{R}^{\sigma\alpha}\hat{R}^{\beta\rho}$, which do not disappear due to Eqs. (34) and (35), transform with the wrong dependence on $e^{-2\sigma}$; they acquire coefficients $e^{-4\sigma}$, $e^{-6\sigma}$, or even $e^{-8\sigma}$.

Substituting Eqs. (12), (37), and (41) into Eq. (33) and comparing the result with Eq. (42), one can derive the transformation rule for $\hat{U}$. In the course of the calculation it is important to notice that because of Eq. (4) the following relaton is valid:

$$\delta^{\alpha}{}_{\mu}\delta^{\beta}{}_{\nu}\gamma^{\sigma\rho}(h_{\alpha\gamma;\delta}Y_{\beta\sigma\rho}{}^{\gamma\delta} + h_{\beta\gamma;\delta}Y_{\alpha\sigma\rho}{}^{\gamma\delta}) = 0.$$

The other terms give the equation

$$h_{\mu\gamma}Z_{\nu\alpha}{}^{\alpha\gamma} + h_{\nu\gamma}Z_{\mu\alpha}{}^{\alpha\gamma} + h_{\gamma\delta}V_{\mu\nu\alpha}{}^{\alpha\gamma\delta} + \hat{U}^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta}e^{2\sigma} = U^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta},$$

which in more detail reads as

$$\hat{U}^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta} = e^{-2\sigma}\,[U^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta} + 2h_{\mu\alpha}\sigma_{\nu;}{}^{\alpha} + 2h_{\nu\alpha}\sigma_{\mu;}{}^{\alpha} + h_{\mu\nu}(\sigma_{\alpha;}{}^{\alpha} + 3\sigma_{\alpha}\sigma^{\alpha})]. \tag{45}$$

The general form of $\hat{U}^{\alpha}{}_{\mu\nu}{}^{\beta}$ which might be consistent with Eq. (44) is

$$\hat{U}^{\alpha}{}_{\mu\nu}{}^{\beta} = a\delta^{\alpha}{}_{\mu}\delta^{\beta}{}_{\nu}\hat{R} + b\,(\delta^{\alpha}{}_{\mu}\hat{R}_{\nu}{}^{\beta} + \delta^{\alpha}{}_{\nu}\hat{R}_{\mu}{}^{\beta}) + c\hat{\gamma}_{\mu\nu}\hat{R}^{\alpha\beta} + d\hat{C}^{\alpha}{}_{\mu\nu}{}^{\beta}, \tag{46}$$

where a, b, c, and d, are arbitrary functions of the space–time variables. Since $\hat{C}^{\alpha}{}_{\mu\nu}{}^{\beta} = e^{-2\sigma}C^{\alpha}{}_{\mu\nu}{}^{\beta}$ and $\hat{\gamma}_{\mu\nu}\hat{R}^{\alpha\beta}h_{\alpha\beta} = e^{-2\sigma}\gamma_{\mu\nu}(R^{\alpha\beta} - 2\sigma^{\alpha;\beta} + 2\sigma^{\alpha}\sigma^{\beta} - \sigma_{\rho;}{}^{\rho}\gamma^{\alpha\beta} - 2\sigma_{\rho}\sigma^{\rho}\gamma^{\alpha\beta})h_{\alpha\beta} = e^{-2\sigma}\gamma_{\mu\nu}R^{\alpha\beta}h_{\alpha\beta}$, the functions c and d are not restricted by Eq. (45) and they can remain arbitrary. As for the functions a and b, they are determined after substitution of Eq. (46) into Eq. (45) and they must be $a = 1/6$, $b = -1$.

Thus, the general form for conformally invariant equations for a second-rank symmetric tensor field $h_{\mu\nu}$ is

$$h_{\mu\nu;\alpha}{}^{;\alpha} + \frac{R}{6}h_{\mu\nu} - h_{\mu\alpha}R^{\alpha}{}_{\nu} - h_{\nu\alpha}R^{\alpha}{}_{\mu} + c\gamma_{\mu\nu}R^{\alpha\beta}h_{\alpha\beta} + dC^{\alpha}{}_{\mu\nu}{}^{\beta}h_{\alpha\beta} = 0. \tag{47}$$

These equations and the auxiliary conditions (3) and (4) are conformally invariant if Eq. (16) is satisfied.

For an easier comparison of Eq. (47) with Eq. (10) we can rewrite the former one in the form

$$h_{\mu\nu;\alpha}{}^{;\alpha} - 2R_{\alpha\mu\nu\beta}h^{\alpha\beta} - \tfrac{1}{6}Rh_{\mu\nu} - \tfrac{1}{2}k\gamma_{\mu\nu}R^{\alpha\beta}h_{\alpha\beta} + lC_{\alpha\mu\nu\beta}h^{\alpha\beta} = 0, \tag{48}$$

where k and l are arbitrary functions. Multiplying this equation by $\gamma_{\mu\nu}$ we obtain its consequence

$$2(1-k)h^{\alpha\beta}R_{\alpha\beta} = 0,$$

which says that either $k = 1$ and then $R_{\alpha\beta}h^{\alpha\beta}$ is not necessarily equal to zero, or $h^{\alpha\beta}R_{\alpha\beta} = 0$ (what has been true for the linearized Einstein equations) and then we can put $k = 0$. In any case, it is seen from Eqs. (48) and (10) that the most important difference between them is the term $\frac{1}{6}Rh_{\mu\nu}$. The lack of this term was the cause of conformal noninvariance of Eq. (10).[23]

One should remember that the conditions (43) and (44) were obtained as a consequence of a desire to keep the operator $h_{\mu\nu;\alpha}{}^{;\alpha}$ in the equations. So if there exist any other conformally invariant equations, different from Eq. (48), it does not include this operator.

We have considered the conformally invariant coupling of $h_{\alpha\beta}$ to curvature. There must exist conformally invariant equations which describe the coupling of $h_{\alpha\beta}$ to curvature and other fields. Derivation of the general form of such equations is a complicated problem, so we shall restrict ourselves to a specific example. This is provided by the equation

$$\hat{h}_{\mu\nu;\alpha}{}^{;\alpha} - 2\hat{R}_{\alpha\mu\nu\beta}\hat{h}^{\alpha\beta} - \frac{\hat{\varphi}_{;\alpha}{}^{;\alpha}}{\hat{\varphi}}\hat{h}_{\mu\nu} = 0, \tag{49}$$

where φ is a scalar field. This equation transforms into

$$e^{-\sigma}\left(h_{\mu\nu;\alpha}{}^{;\alpha} - 2R_{\alpha\mu\nu\beta}h^{\alpha\beta} - \frac{\varphi_{;\alpha}{}^{;\alpha}}{\varphi}h_{\mu\nu}\right) = 0, \tag{50}$$

under the transformation rules $\hat{\varphi} = e^{-\sigma}\varphi$ and Eqs. (11) and (12), and the conditions (16), (34), and (35). Since the scalar field φ transforms with the correct dependence on the conformal factor, it may obey the conformally invariant equation as well. In that case Eq. (50) can be represented in terms of the background variables only, since $(\varphi_{;\alpha}{}^{;\alpha})/\varphi = R/6$.

It is interesting to note that Eq. (49) can be obtained from Eq. (10) as a result of applying the conformal transformation $h_{\mu\nu} = \hat{\varphi}\hat{h}_{\mu\nu}$, $\gamma_{\mu\nu} = \hat{\varphi}^{2}\hat{\gamma}_{\mu\nu}$.

V. NONEXISTENCE OF A GRAVITATIONAL THEORY WITH CONFORMALLY INVARIANT LINEARIZED WAVE EQUATIONS

Equation (10) was derived from the Einstein equations in the linear approximation. Let us see if Eq. (48) can be derived in a similar way from some exact equations which generalize Einstein's equations.

Suppose that the generalized equations have the following form:

$$N_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R + \lambda g_{\mu\nu} + F_{\mu\nu} = T_{\mu\nu}, \tag{51}$$

where $F_{\mu\nu}$ is a symmetric tensor, constructed in an arbitrary way from the exact metric $g_{\mu\nu}(g_{\mu\nu} \simeq \gamma_{\mu\nu} + h_{\mu\nu})$ and its derivatives. In principle, among possible contributions of $F_{\mu\nu}$ could be terms like $RR_{\mu\nu}$, $R_{\alpha\beta} R^{\alpha}{}_{\mu\nu}{}^{\beta}$, etc. We do not assume that $N_{\mu}{}^{\nu}{}_{;\nu} \equiv 0$ should hold necessarily, at least for the time being.

In analogy to the way in which Eq. (10) was derived from the first variation of the Einstein equations, Eq. (48) should follow from the equations

$$-(\gamma_{\mu\alpha} \delta N_{\nu}{}^{\alpha} + \gamma_{\nu\alpha} \delta N_{\mu}{}^{\alpha}) = 0 \tag{52}$$

and the auxiliary conditions (3) and (4). Since the first two terms of Eq. (48) follow from the expression $-\gamma_{\mu\alpha}(\delta R_{\nu}{}^{\alpha} + \lambda \delta_{\nu}{}^{\alpha}) - \gamma_{\nu\alpha}(\delta R_{\mu}{}^{\alpha} + \lambda \delta_{\mu}{}^{\alpha})$, Eq. (52) can be reduced to

$$2\delta F_{\mu\nu} - h^{\alpha}{}_{\mu} F_{\alpha\nu} - h_{\nu}{}^{\alpha} F_{\alpha\mu} = \tfrac{1}{6} R h_{\mu\nu} + (\tfrac{1}{2}k - 1)\gamma_{\mu\nu} h_{\alpha\beta} R^{\alpha\beta} - l C_{\alpha\mu\nu\beta} h^{\alpha\beta}. \tag{53}$$

[Obviously, the background, or "unperturbed," values of the curvature enter the right-hand side of Eq. (53) and background values of $F_{\mu\nu}$ enter the last two terms on the left-hand side of this equation.] The question is whether there exists a tensor $F_{\mu\nu}$ which is a solution to Eq. (53). First of all, one can notice that since the right-hand side of Eq. (53) is linear in the background curvature, the tensor $F_{\mu\nu}$ can only consist of terms which are not higher than quadratic order in curvature, or otherwise the variation of $F_{\mu\nu}$ would give rise to quadratic and higher order terms, which are not present at the right-hand side of Eq. (53). Secondly, since the right-hand side of Eq. (53) does not contain derivatives of $h_{\mu\nu}$, we should exclude the contributions to $F_{\mu\nu}$ which could lead to them, unless they disappear due to Eq. (4).

Then, the general form of $F_{\mu\nu}$ which could meet these restrictions is $F_{\mu\nu} = a g_{\mu\nu} + bR g_{\mu\nu} + c R_{\mu\nu} + m R R_{\mu\nu} + n R_{\mu}{}^{\alpha} R_{\alpha\nu} + p R_{\alpha\beta} R^{\alpha}{}_{\mu\nu}{}^{\beta} + q \Psi(R^2) g_{\mu\nu} + r R_{;\mu;\nu}$, where all the coefficients are arbitrary functions of space–time and $\Psi(R^2)$ symbolizes any quadratic function of scalars constructed out of the curvature tensor. A more detailed analysis shows that, in fact, none of the terms with coefficients, c, m, n, p, q and r is useful because the variation of each of them gives either (i) the second (or higher) derivatives of $h_{\mu\nu}$, which cannot be cancelled out, or (ii) the terms which are quadratic in the background curvature. Both these cases contradict the form of the right-hand side of Eq. (53). Thus, we should seek among the terms with coefficients a and b. The term with coefficient a (like the cosmological term in the Einstein equations) does not play any role because the left-hand side of Eq. (53) calculated from this term is identically equal to zero. The term with coefficient b, for $b = \frac{1}{2} - \frac{1}{4}k$, can give rise to the second term on the right-hand side of Eq. (53). However, this term cannot explain the appearance of the first and the third term. Since l is an arbitrary coefficient, we can choose $l = 0$. However, the presence of the term $\frac{1}{6} R h_{\mu\nu}$ is a real obstacle.

To demonstrate this in a more straightforward way we will choose $l = k = 0$ in Eq. (48). Then, a consequence of this equation and Eq. (3) is $h^{\alpha\beta} R_{\alpha\beta} = 0$, and hence δR is equal to zero, since $\delta R = -h_{\alpha\beta} R^{\alpha\beta} - h_{;\alpha}{}^{;\alpha} + h_{\alpha}{}^{\beta}{}_{;\beta}{}^{;\alpha}$. Equation (53) can now be rewritten in the following form:

$$\gamma_{\mu\alpha} \delta F^{\alpha}{}_{\nu} + \gamma_{\nu\alpha} \delta F^{\alpha}{}_{\mu} = \delta(\tfrac{1}{6} R g_{\mu\nu}). \tag{54}$$

One can solve Eq. (54) with respect to quantities $\delta F_{\beta}{}^{\alpha}$ which, by assumption, should be variations of some tensor. However, from the very way of constructing the solution to Eq. (54) it is clear that this solution is not a variation of a tensor. Thus, conformally invariant Eqs. (48) cannot follow in a linear approximation from any tensor equations of the form of Eq. (51).

VI. CONCLUSIONS

It seems that one is left with two options, though each of them looks interesting. One of them is to agree that the equations which govern weak gravitational waves in a curved space–time are not conformally invariant. Then, it means that on both levels—classical and quantum—gravitons behave drastically different from other massless particles. Classical gravitational waves can be amplified and gravitons can be created (contrary to other massless fields and particles) in a nonstationary isotropic gravitational field, particularly in the strong gravitational field of the early Universe. The other option is to try to endow gravitons with the same kind of coupling to the external gravitational field that other massless particles have. Then one has to find some nontrivial generalization of the Einstein equations.

ACKNOWLEDGMENT

Thanks are due to B. Mashhoon for reading the manuscript.

APPENDIX

We will give here the transformation rules for different terms which could enter the linearized Einstein equations. The background metric tensor $\hat{\gamma}_{\alpha\beta}$ and the field variables $\hat{h}_{\alpha\beta}$ transform as follows:

$$\hat{\gamma}_{\alpha\beta} = e^{2\sigma} \gamma_{\alpha\beta}, \quad \hat{h}_{\alpha\beta} = e^{k\sigma} h_{\alpha\beta},$$

where k is an arbitrary constant. We denote $\sigma_{\alpha} \equiv \sigma_{,\alpha}$; notice also that $\sigma_{\mu;\nu} = \sigma_{\nu;\mu}$. Then,

$$e^{-(k-2)\sigma} \hat{h}_{\mu\nu\hat{;}\alpha}{}^{\hat{;}\alpha} = h_{\mu\nu;\alpha}{}^{;\alpha} + 2(k-1) h_{\mu\nu;\alpha} \sigma^{\alpha} + (k^2 - 2k - 2) h_{\mu\nu} \sigma_{\alpha} \sigma^{\alpha} + (k-2) h_{\mu\nu} \sigma_{\alpha;}{}^{\alpha} + 2\sigma_{\alpha}(h_{\mu}{}^{\alpha}{}_{;\nu} + h_{\nu}{}^{\alpha}{}_{;\mu}) - 2(h_{\mu}{}^{\alpha}{}_{;\alpha} \sigma_{\nu} + h_{\nu}{}^{\alpha}{}_{;\alpha} \sigma_{\mu}) - 4\sigma_{\alpha}(h_{\mu}{}^{\alpha} \sigma_{\nu} + h_{\nu}{}^{\alpha} \sigma_{\mu}) + 2\gamma_{\mu\nu} h_{\alpha\beta} \sigma^{\alpha} \sigma^{\beta} + 2h \sigma_{\mu} \sigma_{\nu},$$

$$
\begin{aligned}
-e^{-(k-2)\sigma}\widehat{R}_{\alpha\mu\nu\beta}\hat{h}^{\alpha\beta} = & -R_{\alpha\mu\nu\beta}h^{\alpha\beta} + h^{\alpha\beta}(\sigma_\alpha\sigma_\beta - \sigma_{\alpha;\beta})\gamma_{\mu\nu} + h_{\mu\nu}\sigma_\alpha\sigma^\alpha + h_\mu{}^\alpha\sigma_{\nu;\alpha} \\
& + h_\nu{}^\alpha\sigma_{\mu;\alpha} - \sigma_\alpha(h_\mu{}^\alpha\sigma_\nu + h_\nu{}^\alpha\sigma_\mu) + h(\sigma_\mu\sigma_\nu - \sigma_{\mu;\nu} - \gamma_{\mu\nu}\sigma_\alpha\sigma^\alpha),
\end{aligned}
$$

$$
\begin{aligned}
-e^{-(k-2)\sigma}&\hat{h}^{\alpha}{}_{\mu\hat{;}\alpha\hat{;}\nu} \\
= & -h^{\alpha}{}_{\mu;\alpha;\nu} - (k-3)h^{\alpha}{}_{\mu;\alpha}\sigma_\nu + h^{\alpha}{}_{\nu;\alpha}\sigma_\mu - h^{\alpha\beta}{}_{;\beta}\sigma_\alpha\gamma_{\mu\nu} - (k+2)h^{\alpha}{}_{\mu;\nu}\sigma_\alpha - (k^2-k-6)h^{\alpha}{}_{\mu}\sigma_\alpha\sigma_\nu \\
& -(k+2)h^{\alpha}{}_{\mu}\sigma_{\alpha;\nu} + (k+2)h^{\alpha}{}_{\nu}\sigma_\alpha\sigma_\mu - (k+2)h^{\alpha\beta}\sigma_\alpha\sigma_\beta\gamma_{\mu\nu} + h_{,\nu}\sigma_\mu + h\,[\sigma_{\mu;\nu} + \sigma_\alpha\sigma^\alpha\gamma_{\mu\nu} + (k-4)\sigma_\mu\sigma_\nu],
\end{aligned}
$$

$$
e^{-(k-2)\sigma}\hat{h}_{\hat{;}\mu\hat{;}\nu} = h_{;\mu;\nu} + (k-3)(h_{,\mu}\sigma_\nu + h_{,\nu}\sigma_\mu) + \gamma_{\mu\nu}h_{,\alpha}\sigma^\alpha + (k-2)h\,[(k-4)\sigma_\mu\sigma_\nu + \sigma_{\mu;\nu} + \sigma_\alpha\sigma^\alpha\gamma_{\mu\nu}],
$$

$$
\begin{aligned}
e^{-(k-2)\sigma}(\hat{h}_\mu{}^\alpha\widehat{R}_{\alpha\nu} + \hat{h}_\nu{}^\alpha\widehat{R}_{\alpha\mu}) = & -(h_\mu{}^\alpha R_{\alpha\nu} + h_\nu{}^\alpha R_{\alpha\mu}) + 2(h^{\alpha}{}_{\mu}\sigma_{\nu;\alpha} + h^{\alpha}{}_{\nu}\sigma_{\mu;\alpha}) \\
& -2\sigma_\alpha(h_\mu{}^\alpha\sigma_\nu + h_\nu{}^\alpha\sigma_\mu) + 2h_{\mu\nu}(\sigma_{\alpha;}{}^{\alpha} + 2\sigma_\alpha\sigma^\alpha).
\end{aligned}
$$

[1]*Scale and Conformal Symmetry in Hadron Physics*, edited by R. Gatto (Wiley, New York, 1973).

[2]F. Bayen, in *Differential Geometry and Relativity* (A. Lichnerowicz honour volume), edited by M. Cohen, and M. Flato (Reidel, Dordrecht, 1976).

[3]T. Fulton, F. Rohrlich, and L. Witten, Rev. Mod. Phys. **34**, 442 (1962).

[4]V.I. Fushchich and A.G. Nikitin, Lett. Math. Phys. **2**, 471 (1978).

[5]R. Penrose, Proc. R. Soc. (London) Ser. A **284**, 159 (1965); in *Batelle Rencontres*, edited by C.M. DeWitt and J.A. Wheeler (Benjamin, New York, 1968).

[6]Elementary particles in De-Sitter space are considered in Refs. 24–26.

[7]L. Parker, Phys. Rev. **183**, 1057 (1969); in *Relativity, Fields, Strings and Gravity*, edited by Aragone (1975).

[8]Ya.B. Zel'dovich and I.D. Novikov, *Structure and Evolution of the Universe* (Nauka, Moscow, 1975).

[9]L.P. Grishchuk, Sov. Phys. JETP **40**, 409 (1975); Ann. N.Y. Acad. Sci. **302**, 439 (1977).

[10]L.P. Grishchuk and A.D. Popova, Sov. Phys. JETP (submitted for publication) (Caltech preprint).

[11]P.A.M. Dirac, Proc. R. Soc. (London) Ser. A **333**, 403 (1973).

[12]V. Canuto, P.J. Adams, S.-H. Hsiah, and E. Tsiang, Phys. Rev. D **16**, 1643 (1977).

[13]F. Englert, C. Truffin, and R. Gastmans, Nucl. Phys. B **117**, 407 (1976).

[14]E.S. Fradkin and G.A. Vilkovisky, Phys. Lett. B **77**, 262 (1978).

[15]S. Deser, Ann. Phys. (N.Y.) **59**, 248 (1970).

[16]H. Umezawa, *Quantum Field Theory* (Interscience, New York, 1956).

[17]C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[18]S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).

[19]We would like to note that the gauge freedom in a curved background space–time does not have exactly the same meaning as for flat space–time. It is true that Eq. (2) is gauge invariant. However, if one chooses, say, perturbations of the Weyl tensor to be "physical" variables (or field components) versus "potentials" $h_{\mu\nu}$, then they are not in general gauge invariant. This is essentially because the Lie derivative of the background Weyl tensor does not in general vanish.

[20]To find ξ_μ one has to solve a set of differential equations. Strictly speaking, the existence of ξ_μ is only guaranteed locally, in a vicinity of some initial hypersurface.

[21]L.P. Grishchuk and A.D. Popova, in Abstracts of IV Soviet Gravitation Conference Minsk (1976).

[22]E.M. Lifshitz, Zh. Eksp. Teor. Fiz. **16**, 587 (1946).

[23]As a curious fact, we should point out that Eq. (9) reduces to Eq. (48) (with $k = l = 0$) under the conditions $\chi_{\mu;\nu} + \chi_{\nu;\mu} - \gamma_{\mu\nu}(-h_{\alpha\beta}R^{\alpha\beta} + \chi_{\alpha;}{}^{\alpha} - \frac{1}{2}h_{;\alpha}{}^{;\alpha}) - \frac{1}{6}Rh_{\mu\nu} - (\gamma_{\mu\alpha}\delta T^{\alpha}{}_{\nu} + \gamma_{\nu\alpha}\delta T^{\alpha}{}_{\mu}) = 0$. However, this equation cannot be used as a gauge condition for χ_μ since it is not, in general, integrable. On the other hand, if this equation is to be interpreted as a definition of $\delta T_\mu{}^\nu$ (together with $\chi_\mu = 0$), then it picks up a mixture of tensor, scalar, and vector harmonics in the case of Robertson–Walker background metrics, which does also seem to be unsatisfactory.

[24]N.I. Chernikov and E.A. Tagirov, Ann. Inst. H. Poincaŕe **9**, 109 (1968).

[25]C. Fronsdal, Phys. Rev. D **12**, 3819 (1975).

[26]J. Fang and C. Fronsdal, Lett. Math. Phys. **2**, 391 (1978).

On a class of solutions of the Krook–Tjon–Wu model of the Boltzmann equation

M. Barnsley and H. Cornille
CEN-Saclay, Boite Postale N°2, 91190 GiF Sur Yvette, France

(Received 21 June 1979; accepted for publication 15 November 1979)

The construction of the solutions of the Krook–Wu model of the nonlinear Boltzmann equation (elastic differential cross sections inversely proportional to the relative speed of the colliding particles) is reduced to the resolution of a nonlinear first order differential system for functions which are related to the moments of the Boltzmann distribution functions. These functions depend upon only one variable, the time, and are the coefficients in the Laguerre expansion of the Tjon-Wu distribution function. We explicitly show how to construct the solutions of the nonlinear differential system and study the properties of the corresponding Tjon-Wu distribution function.

1. INTRODUCTION

Recently Krook and Wu[1] and Tjon and Wu[2] have provided a model of the nonlinear Boltzmann equation for the relaxation towards equilibrium of a spatially homogeneous and isotropic gas with one species of molecules. As usual only binary elastic scattering is taken into account and they further assume that the elastic cross section is inversely proportional to the relative speed. In their model the Boltzmann distribution function $F(x,t)$ (t being the time and $x = v^2/2$, v being the velocity) can be obtained from two independent although equivalent formalisms: firstly, by solving directly the nonlinear integral differential Boltzmann equation[2]; secondly, they have shown both that the Boltzmann generating function satisfies a well defined nonlinear equation[1] and that $F(x,t)$ can be obtained from the Inverse Laplace transform of the Boltzmann generating functional.[2] They were able to exhibit an explicit particular solution which was also found by Bobylev[3] and they have conjectured and investigated whether some features of this particular solution are also present in the more general case. In this paper, using mainly the second formalism, we extend their recalled results and give the method in order to get the solutions while in a companion paper,[4] an independent study is performed with the help of the first formalism.

The existence of the two above formalisms means that the resolution of either the nonlinear integrodifferential Tjon–Wu model or of the nonlinear partial differential Boltzmann generating function are two complementary views of the same underlying problem. This view is strengthened when we realize that the resolution of the same nonlinear differential equation (with only one variable) solves both problems and we sketch here very briefly the key equations to be considered in order to understand this point.

There exists a straightforward connection between the normalized moments $M_n(t)$ of the Boltzmann distribution function and the distribution function $F(x,t)$ of the Tjon–Wu model. Let us define

$$a_n(t) = \sum_{q=0}^{n+2} (-1)^q C^q_{n+2} M_{n+2-q}(t), \tag{1}$$

where C^q_p are the usual binomial coefficients; then the $\{a_n\}$ satisfy a nonlinear differential system

$$(n+3)\frac{da_n}{dt} + (n+1)\,a_n = \sum_{M+N=n-2} a_M a_N,\quad n = 0,1,\ldots,\quad a_n(t) \underset{t\to\infty}{\longrightarrow} 0, \tag{2}$$

whereas Fe^x has an expansion in Laguerre polynomials with coefficients $(-1)^n a_n$:

$$e^x F(x,t) = 1 + \sum_{n=0}^{\infty} (-1)^n a_n(t) L_{n+2}(x). \tag{3a}$$

On the other hand, the $\{a_n\}$ enter into the Taylor expansion of a function $H(u,t)$ associated to the generating functional of the Boltzmann moments

$$H(u,t) = \sum_0^{\infty} u^{n+2} a_n(t) = -1 + \sum_0^{\infty} u^n (1+u)^{-(n+1)} M_n(t), \tag{4}$$

which satisfies a nonlinear partial differential equation (n.l.p.d.e.)

$$\left(\frac{\partial}{\partial u} + \frac{\partial^2}{\partial u\,\partial t}\right)(u + uH) = (1+H)^2 \tag{5}$$

and which is also connected to a power series expansion of $e^x F$:

$$e^x F(x,t) = 1 + \sum_{n=0}^{\infty} \frac{x^n}{(n!)^2}\left(\frac{\partial^n}{\partial x^n} H(u,t)\right)_{u=-1}. \tag{3b}$$

In other words, if we solve Eq. (2) and substitute the a_n's into the expansion (3.4), we have both a class of solutions of the Tjon–Wu model and of the Boltzmann generating functional n.l.p.d.e. [Eq. (5)].

The aim of this paper is to solve the system (2) for the $\{a_n(t)\}$ and to study the corresponding properties for $F(x,t)$. In the second section we recall briefly the general formalism from which we have taken here in Eqs. (1)–(5), the key equations. In Sec. 3 and 4 we study the general structure of the solutions of Eq. (2) and some general properties of $F(x,t)$.

We quote here the more salient features:

(i) Due to the particular form of the linear part of Eq.

 0022-2488/80/051176-18$01.00

(2) (l.h.s.) and the fact that the r.h.s. can be recursively determined we easily see that $a_n(t)$ (n fixed) is a superposition of a finite number of terms carrying only discrete well-defined time dependence, the less decreasing one being $\exp\{-[(n+1)/(n+3)]t\}$. This property is fundamental for the reconstructed $F(x, t)$ because the only time dependence is provided by the set $\{a_n(t)\}$. Consequently, $a_n(t) \to 0$ when $t \to \infty$ and the Maxwellian behavior equivalently $e^x F \to 1$ at large x (but *a priori* finite) is automatically satisfied at equilibrium. Investigating the reason for this important property of the formalism, we find that it is a consequence of the conservation of mass and energy and allowing a violation of energy conservation we show a corresponding violation of the asymptotic Maxwellian behavior.

(ii) Due to the explicit form of the expansion (3b) we see that the initial conditions on $e^x F(x, 0)$ can equivalently be replaced by the yield of a set of values of $H(-1, 0)$ and of its derivatives. We establish a set of sufficient conditions in order that the series in Eq. (3b) be absolutely convergent and define, for any t value, an integer function in the x plane. From Eq. (3a) we remark that $e^{x/2}F$ is a series of Laguerre orthogonal functions. For these orthogonal functions we show that a finite number of terms in their Taylor series provide lower and upper bounds which can be carried out in $e^{x/2}F$ when the coefficient $a_n(t)$ possess well defined positivity properties.

(iii) Investigating the properties of the integrodifferential Boltzmann equation corresponding to the Tjon–Wu model, we study under what conditions the positivity of $F(x, t)$ at $t = 0$ can propagate at positive t values.

(iv) Still due to the fact that the r.h.s. of Eq. (2) can be recursively determined, it is clear that $a_n(t)$ depends at most upon n independent arbitrary constants which are introduced by integrating either the first, the second,..., the nth equation of the system. These constants appearing either at the origin $a_n(0)$ or at infinity $\bar{a}_n = \lim_{t\to\infty} a_n(t) \times \exp[(n+1)/(n+3)]t$, this gives us the possibility to classify the solutions in two different classes.

In the first class (see Sec. 5), the fundamental solutions called "pure solutions" are those having all constants $\bar{a}_n \equiv 0$ except one $\bar{a}_{n_0} \neq 0$. They possess very interesting properties. For any n there exists only one term in a_n and so only one time dependence multiplied by a constant which can be determined recursively. They lead in Eq. (5) to "solitonlike" solutions with the meaning that these solutions for H depend in fact upon only one variable linear combination of $\log u$ and t. We can also define the mixing of a *finite* number of such solutions requiring that all $\lim_{t\to\infty} a_n(t) \times \exp[(n+1)/(n+3)]t = 0$ except for two (n_0, n_1), three (n_0, n_1, n_2),...values. The particular Krook–Wu solution[1] is a particular mixing solution for $n_0 = 0$ and $n_1 = 1$ having the interesting property that for t larger than some well defined t_0 value, the $F(x, t)$ remains positive. Unfortunately, when the $\{a_n\}$, corresponding to other solutions than the Krook–Wu one, are substituted into the expansion (3), we have not found any other example where $F(x, t)$ remains positive for t large. Consequently, we must look at other classes of solutions of the system (2) or at an infinite mixing of pure solutions.

In the second class (see Sec. 6), the fundamental solutions called "positive solutions" are those having all $a_n(0) = 0$ except one $a_n(0) \neq 0$. Then we can control the positivity property of the Laguerre expansion at $t = 0$. Fortunately, these positivity properties subsist at higher t values in such a way that we can construct positive $F(x, t)$ solutions. However, their explicit construction is cumbersome. For a_n given there is not a unique time dependence but a finite number of terms carrying different time dependences. However, we can still provide a recursive scheme in order to construct explicitly the solutions. We can also define the mixing of such positive solutions having only two, three,... $a_n(0) \neq 0$ and carrying positivity properties for the $F(x, t)$. For an infinite mixing it is not very easy to find at $t = 0$ the conditions on the set $[a_n(0)]$ ensuring the positivity of the sum of Laguerre polynomials $e^x F(x, 0)$. However, with the help of the generating functional of the Laguerre polynomials we can construct exemples of expansions (3a) at $t = 0$ where the sum is given in closed form so that we control easily the positivity. Finally, as an illustration of our method we quote some numerical examples of solutions for $t = 0$ and $t \neq 0$.

2. THE FORMALISM

Here we recall briefly the Krook–Tjon–Wu formalism while in the companion paper[4] a more general derivation is performed. We consider the Boltzmann equation for a spatially homogeneous and isotropic gas and assume that the elastic cross sections of binary collisions are inversely proportional to the relative speed. We consider the Boltzmann distribution function $f(v, t)$ (v being the velocity, t the time, and $x = v^2/2$ the energy) and the units are such that $\exp(-x)$ corresponds to a Maxwellian distribution. To $f(v, t)$ we associate the moments $M_n(t)$ normalized in such a way that $\lim_{t\to\infty} M_n(t) \equiv 1$ if $f(v, t)$ tends to a Maxwellian distribution at infinite time. We define G, a generating functional of these moments, which can also be considered as a transform of $f(v, t)$:

$$G(\zeta,t) = \sum_0^\infty \zeta^n M_n(t), \tag{6}$$

$$G(\zeta,t) = 4\pi \int_0^\infty \left(v^2 \sum_{n=0}^\infty \frac{2^n n!}{(2n+1)!} (\zeta v^2)^n \right) f(v, t)\, dv. \tag{7}$$

The moments M_n must satisfy well-defined constraints: $M_0 \equiv 1$ for the conservation of masses, $M_1 \equiv 1$ for the conservation of energy, and the requirement of a Maxwellian distribution at equilibrium

$$M_0(t) \equiv M_1(t) \equiv 1, \quad \lim_{t\to\infty} M_n(t) \equiv 1. \tag{8}$$

Taking into account these constraints we define a new generating functional H, a new variable u, and a set $a_n(t)$ deduced from the $M_n(t)$ following the relation (1):

$$H(u,t) = -M_0 + (1-\zeta)G(\zeta,t) = a_{-1}(t)u + \sum_0^\infty u^{n+2} a_n(t), \tag{4'}$$

$M_0 \equiv 1,\ \zeta(1+u^{-1}) = 1,\ a_{-1} = M_1 - M_0 \equiv 0,\ a_n(t) \underset{t\to\infty}{\longrightarrow} 0,$

$$M_n(t) = 1 + \sum_{q=0}^{n+2} C_n^q a_{n-2-q}(t),\ n \geqslant 2. \qquad (1')$$

Krook and Wu have shown that the moments $M_n(t)$ satisfy a nonlinear system of equations

$$\frac{d}{dt} M_n + M_n \frac{(n-1)}{(n+1)} = \frac{1}{n+1} \sum_{k=1}^{n-1} M_k M_{n-k}, \quad (n \geqslant 2)$$

from whichs follows that G satisfies an n.l.p.d.e.

$$\left(\frac{\partial^2}{\partial \zeta\, \partial t} + \frac{\partial}{\partial \zeta} \right) \zeta G = G^2, \qquad (5')$$

which reduces to the n.l.p.d.e. (5) for the $H(u,t)$ generating functional. If we substitute $H(u,t)$ having the particular expansion (4) into Eq. (5):

$$\sum u^{(2+n)} \left[(3+n) \frac{da_n(t)}{dt} + (n+1)\ a_n(t) \right] = \left(\sum u^{2+n} a_n(t) \right)^2. \qquad (2')$$

we are reduced to a problem *depending upon only one variable t*. We have to solve the nonlinear system (2). Once this system has been solved we substitute the $a_n(t)$ into G:

$$G(\zeta,t) = (1-\zeta)^{-1} \left[1 + \sum^{\infty} a_n(t) (\zeta(1-\zeta)^{-1})^{n+2} \right], \qquad (4'')$$

and the Boltzmann distribution $f(v,t)$ can be obtained by inverting the transform (7).

Due to the complexity of this transform (7), Tjon and Wu have defined a new distribution function $F(x = v^2/2, t)$ through inverse Laplace tranfsorm:

$$G(\zeta,t) = \int_0^\infty e^{\zeta x} F(x,t)\, dx. \qquad (9)$$

In order to compare the two distribution functions f and F let us define new moments $\widetilde{M}_n(t) = \int_0^\infty x^n F(x,t) dx$ and from Eqs. (9) and (6) we get $n! M_n = \widetilde{M}_n$. It follows that $M_0 = \widetilde{M}_1 \equiv 1$ and the same conservation laws hold. Furthermore, from Eqs. (5′) and (9) it follows that F satisfies an integrodifferential nonlinear equation

$$\frac{\partial}{\partial t} (e^t F(x,t)) = e^{-t} \int_x^\infty \frac{dx'}{x'} \int_0^{x'} dx''\, F(x'-x'',t) F(x'',t)\, dx''. \qquad (10)$$

Recently Tjon and Wu[2] have given the necessary assumptions in order to derive Eq. (10) directly as a particular case of the nonlinear Boltzmann equation.

In this paper we shall not try to solve Eq. (10) directly although our main interest will be the study of the Tjon–Wu distribution $F(x,t)$. Our strategy will be to solve the system (2) for the (a_n), substitute the a_n into G given by Eq. (4″), and get $F(x,t)$ throught the inverse Laplace transform of Eq. (9). Let us remark that the inverse Laplace transform of $(-\zeta)^n (1-\zeta)^{-(n+1)}$ is just the Laguerre polynomial $L_n(x) = \Sigma_{q=0}^n (-1)^q x^q C_n^q (q!)^{-1}$ multiplied by e^{-x} and it follows that

$$e^x F(x,t) = 1 + \sum_0^\infty a_n(t)(-1)^n L_{n+2}(x). \qquad (3a)$$

From the explicit expression of the $L_n(x)$ in power of x we can formally rewrite Fe^x as a power series (assuming that we can interchange the order of summations)

$$e^x F(x,t) = 1 + \sum_{q=0}^\infty \frac{x^q}{(q!)^2} \left(\frac{\partial^q H(u,t)}{\partial u^q} \right)_{u=-1}, \qquad (3b)$$

$$\left.\frac{\partial^q}{\partial u^q} H(u,t)\right|_{u=-1} = (-1)^q \sum (n+2)(n+1)n \cdots (n+3-q)\ \ (-1)^n a_n(t),$$

which can be justified only after a study of the properties of the $\{a_n\}$ or of the $\{H^{(q)}(u=-1)\}$. Sufficient conditions for absolute convergence in Eq. (3b) will be established later. If when $t\to\infty$, $a_n(t)\to 0$, then for x large but finite, $F(x,t)$ tends to the Maxwellian distribution e^{-x}. The power series (3b) explicits the link between the two formalisms discussed in the Introduction. We have the relation

$$\frac{\partial^q}{\partial u^q} (F(x,t)e^x)_{x=0} = (q!)^{-1} \frac{\partial^q}{\partial u^q} H(u,t)_{u=-1}. \qquad (11)$$

3. GENERAL STRUCTURE OF THE SOLUTIONS OF THE NONLINEAR DIFFERENTIAL SYSTEM (2) FOR THE $\{a_n(t)\}$

If in Eq. (2′) we require for each power u^{n+2} that the coefficient (which depends only on t) is identically zero, we get two linear differential equations for a_0 and a_1 and the nonlinear system (2) for $n \geqslant 2$:

$$3\frac{da_0}{dt} + a_0 = 0,\ 2\frac{da_1}{dt} + a_1 = 0,$$

$$(n+3)\ \frac{da_n}{dt} + (n+1)\ a_n = \sum_{m+p=n-2} a_m a_p,\ n = 2,3,\cdots. \qquad (2)$$

For $n=0$ and 1 we integrate directly and get $a_0(t) = \bar{a}_0 \exp(-t/3)$, $a_1(t) = \bar{a}_1 \exp(-t/2)$, where $\bar{a}_0$ and $\bar{a}_1$ are arbitrary constants. Let us define two sets of constants $\bar{a}_n = \lim_{t\to\infty} a_n(t) \exp[-t(n+1)/(n+3)]$ and $a_n(t=0)$. For $n=0$ and 1 these constants are identical although this is not true for $n \geqslant 2$. For $n \geqslant 2$ the r.h.s. of Eq. (2) gives a known contribution when the a_m with $m \leqslant n-2$ have been previously determined. In this way we get

$$\begin{aligned} a_2 &= \bar{a}_2 \exp(-3t/5) - 3\bar{a}_0^2 \exp(-2t/3) \\ &= a_2(0)\exp(-3t/5) + 3\bar{a}_0^2 \\ &\quad \times [\exp(-3t/5) - \exp(-2t/3)], \\ a_3 &= \bar{a}_3 \exp(-2t/3) - 2\bar{a}_0\bar{a}_1 \exp(-5t/6) \\ &= a_3(0)\exp(-2t/3) - 2\bar{a}_0\bar{a}_1 \\ &\quad \times [\exp(-5t/6) - \exp(-2t/3)], \end{aligned}$$

and we verify that $\bar{a}_n \neq a_n(0)$. We could go on recursively for higher n values and write explicitly a_n; however, a general framework for the solutions can be established. Let us define $\tilde{a}_n(t)$ and integrate formally Eq. (2):

$$\tilde{a}_n(t) = (n+3)^{-1} \left[\exp\left(\frac{n+1}{n+3} t \right) \sum_0^{n-2} a_m(t) a_{n-2-m}(t) \right],$$

$$a_n(t) = \exp\left[-\left(\frac{n+1}{n+3}\right)t\right]\left[\tilde{a}_n + \int_\infty^t \tilde{a}_n(t')dt'\right], \tag{12}$$

$$a_n(t) = \exp\left[-\left(\frac{n+1}{n+3}\right)t\right]\left[a_n(0) + \int_0^t \tilde{a}_n(t')dt'\right],$$

where the constants $\bar{a}_n$ and $a_n(0)$ verify the relation $\bar{a}_n = a_n(0) + \int_0^\infty \tilde{a}_n(t)dt$. The validity of the first representation in Eq. (12) requires $\lim_{t\to\infty} \tilde{a}_n(t) = 0$ and we shall first prove this property. If this is true, then $\lim_{t\to\infty} a_n(t) = 0$ and from Eq. (1′) we see that the asymptotic conditions $\lim_{t\to\infty} M_n(t) = 1$ are automatically satisfied. Secondly, the most general solution $a_n(t)$ depends upon n arbitrary constants and we shall classify different classes of solutions. Thirdly, we emphasize that the property $M_n(t)\to 1$ [or $a_n(t)\to 0$] when $t\to\infty$ is in the present formalism a consequence of the conservation laws $M_0 \equiv M_1 = 1$. We explicitly show that a violation $M_1 \not\equiv 1$ leads to $a_n \not\to 0$ and consequently to a violation of the asymptotic Maxwellian behavior in Eq. (3).

A. $a_n(t)$ when $t\to\infty$ decreases at least like $\exp\{-[(n+1)/(n+3)]t\}$

We shall prove this property by induction. It is true for $n = 0, 1$; let us assume that the property holds for $n = 0, 1, 2, ..., n-2$, and try to prove it for n. a_m decreases at least like $\exp[-t(m+1)/(m+3)]$ for $m \leqslant n-2$; then $\tilde{a}_n$ decreases at least like $\exp(-bt)$ with

$$b = \frac{m+1}{m+3} + \frac{n-1-m}{n+1-m} - \frac{n+1}{n+3}$$

or

$$b = \frac{(n-2)\,(m(n-2-m)+n+6\ )-7m(n-2-m)+3}{(n+3)(m+3)(n+1-m)}$$
$$> 0 \quad \forall m \leqslant n-2.$$

Consequently, we can integrate $\tilde{a}_n(t)$ when $t\to\infty$, the representation (12) is valid, and $a_n(t)$ decreases at least like $\exp[-(n+1)t/(n+3)]$. As a by-product $\bar{a}_n$ is really defined as $\lim_{t\to\infty} a_n(t)\exp[((n+1)/(n+3))t]$; further, $a_n(t)\to 0$ and thus the boundary conditions (8) for the moments are satisfied.

B. Classification of the solutions (a_n) following the yield of the arbitrary integration constants

a_n depends upon n arbitrary constants $\bar{a}_0, \bar{a}_1, ..., \bar{a}_n$, or equivalently $a_0(0), a_1(0), a_n(0)$, or equivalently by n constants where for each p value $p \leqslant n$ we choose either $\bar{a}_p$ or $a_p(0)$. If the solution is known for $m \leqslant n-2$, then we determine $\tilde{a}_n$ and deduce a_n. Because the number of arbitrary constants increases with n, it follows that there exists an infinite number of solutions and if we substitute these a_n into Eq. (4) or (3) we get an infinite number of solutions for Eq. (3)–(10). In order to clarify the discussion we shall define two bases characterized either by $\{\bar{a}_n\}$ or by $\{a_n(0)\}$.

(i) In the first basis, the fundamental solutions, called pure solutions, are defined by having only one $\bar{a}_{n_0} \neq 0$. In that case the $a_n(t)$ have only one term, only one time dependence, and the constants can be determined recursively. In this way we can also mix two, three, ..., an infinite number of such solutions with only two, three, ..., an infinite number of $\bar{a}_n \neq 0$. If for all n we take $\bar{a}_n \neq 0$, we reconstruct of course the general solution. These pure solutions are in some way similar to "solitons" with the meaning that for Eq. (5) they depend in fact upon only one variable linear combination of u and t. These solutions are studied in Sec. 5. The particular solution found by Krook and Wu is a particular mixing of two such pure positive solutions which leads to a positive $F(x, t)$ for t higher than a well-defined t_0. Unfortunately, when these pure solutions (or a finite mixing of them) are substituted into the expansions (3a) for $F(x, t)$ we have not found any other example where $F(x, t)$ remains positive for t sufficiently large.

(ii) In the second basis, the fundamental solutions, called positive solutions, are defined by having only one $a_n(0) \neq 0$. In that case for any finite n, a_n is a sum of a finite number of terms with different time behavior and always one term decreases like $\exp[(-(n+1)/(n+3))t]$. Then the computation of a_n, which can also be done recursively, is not so easy as in the previous case. We can also define the mixing of two, three, or an infinite number of such solutions requiring that only two, three, or an infinite number of $a_n(0)$ are different of zero. Clearly, also if $a_n(0) \neq 0$ for all n, we reconstruct also the general solution. The great advantage is that we control the positivity at $t = 0$ and by substitution into Eq. (3a) we can start with solutions $F(x, t) > 0$ at $t = 0$ and which remain positive for positive t. These solutions are studied in Sec. 6. In fact, the positive solutions can be built up from an infinite mixing of pure solutions and conversely. Consider for instance the positive solution $a_0(0) \neq 0$ [$a_n(0) = 0$ for $n \neq 0$]. From Eq. (12) we get $\bar{a}_{2p+1} = 0$, $\bar{a}_2 = 3\bar{a}_0^2$ and more generally $\bar{a}_{2p} = \int_0^\infty \tilde{a}_{2p}(t, \bar{a}_0, \bar{a}_2, ..., \bar{a}_{2n-2})dt \neq 0$. Conversely, the pure solution $\bar{a}_0 \neq 0$ ($\bar{a}_n = 0$ for $n \neq 0$) can be obtained requiring $a_2(0) = -3\bar{a}_0^2$, $a_{2p+1}(0) = 0$, and $a_{2p}(0) = -\int_0^\infty \tilde{a}_{2p}(t, \bar{a}_0, a_2(0), a_{2n-2}(0))dt$.

(iii) More generally we could consider another basis in the following way: For any n value we associate either the pure solution $\bar{a}_n \neq 0$ or the positive solution $a_n(0) \neq 0$, the rule for such a basis being that the more general solution built-up by mixing all the fundamental solutions of the basis must contain for $a_n(t)$, n arbitrary really independent constants.

C. Violation of the energy conservation law $M_1 \equiv 1$

As we have seen, the structure of the system (2) is such that the property $a_n(t)\to 0$ when $t\to\infty$ is automatically satisfied. Consequently, in Eq. (3a) at large but finite x and $t\to\infty$, then $F(x, t)$ tends to the Maxwellian distribution e^{-x}. In order to understand more clearly this important property we relax the moments conditions $M_i(t) \equiv 1$ for $i = 1, 2$ and concentrate our attention on the second moment condition. If $M_0(0) = 1$ but $M_1(0) \neq 1$ as we shall see, F tends to $[M_1(0)]^{-1} \times \exp[-x/M_1(0)]$ when $t\to\infty$ and so leads to a violation of the Maxwellian distribution at equilibrium.

Let us start with the representation (4′) where $M_i \not\equiv 1$, $i = 1, 2$ and substituting into Eq. (5′) we obtain a generalization of the system (2):

$$(n+3)\left(\frac{da_n}{dt} + a_n\right) = \sum_{m=-2}^{m=n-2} a_m a_{n-m-2},$$

$a_{-2} = M_0, a_{-1} = M_1 - M_0, n = -2, -1, 0, 1, 2, \cdots, (2'')$

which reduces to Eq, (2) when $M_0 \equiv M_1 \equiv 1$. If we solve the two first nonlinear equations for $n = -2$ and $n = -1$, we get that the two first moments are proportional

$$\frac{M_0(t)}{M_0(0)} = \frac{M_1(t)}{M_1(0)} = [M_0(0) + (1 - M_0(0))e^t]^{-1}.$$

We have two different possibilities:

(i) if $M_0(0) = 1$, we see that $M_0(t) \equiv 1$ and $M_1(t) = M_1(0)$;

(ii) if $M_0(0) \neq 1$, we consider first $M_0(0) > 1$ and get that $M_0(t), M_1(t)$ are increasing functions with a discontinuity at $t = \log(M_0(0)/(M_0(0) - 1))$ becoming negative for larger t values. Secondly, if $M_0(0) < 1$, we see that $M_0(t)$ and $M_1(t)$ are always decreasing, remaining positive and going to zero when $t \to \infty$. In the following we always consider the first possibility $M_0(0) = 1$ which implies $M_0(t) \equiv 1$, $M_1(t) \equiv M_1(0) = 1 + \lambda$, with λ not necessarily zero. We put $a_{-2} = 1, a_{-1} = \lambda$, or $M_1 = 1 + \lambda$ into the above nonlinear system and get a new one very similar to Eq. (2):

$$(n+3)\frac{d}{dt}a_n + (n+1)a_n = \sum_{m+p=n-2} a_m a_p,$$

with the change that n begins to -1 instead of 0. Let us write formally $a_n(t) = \lambda^{n+2}\delta_n + b_n(t)$, $\delta_{-1} = 1$, $b_{-1} \equiv 0$, δ_n being constants, and substitute it into the nonlinear differential system. Then we get two distinct systems. The first one is time independent $(n+1)\delta_n = \Sigma\, \delta_m \delta_{n-2-m}$ and gives directly the solution $\delta_n = 1$ whereas the second one is time dependent:

$$(n+3)\frac{d}{dt}b_n + (n+1)b_n = \sum_{m+p=n-2} (b_m \lambda^{p+2} + b_p \lambda^{m+2} + b_m b_p), \quad n = -1, 0, 1, \cdots.$$

The solution can be written

$$b_n(t) = \exp\left[-\left(\frac{n+1}{n+3}\right)t\right]\left[b_n(0) + \int_0^t \exp\left(\frac{n+1}{n+3}\right)t' \times\left[\sum_{m+p=n-2}(b_m(t')\lambda^{p+2} + b_p(t')\lambda^{m+2}\right]\right] + b_m(t')b_p(t')\,dt'],$$

where the arbitrary constants are defined at $t = 0$. Here also the solutions $(b_n(t))$ can be obtained recursively. Starting with $b_0 = b_0(0)e^{-t/3}$ we determine $b_1, b_2, \ldots$ It can be shown by induction that if $b_0(0) \neq 0$ then the $b_n(t)$ decrease at least like $\exp(-t/3)$. If $b_0(0) = 0$, we get $b_1 = b_1(0)\exp(-t/2)$ and we still show by induction that b_n decreases at least like $\exp(-t/2)$ and so on. In conclusion, $a_n(t) - \lambda^{n+2}$ tends to zero when $t \to \infty$.

Next we substitute into the Laguerre expansion (3b) (where the summation now begins at $n = -1$) neglecting the contribution due to $(b_n(t))$ and investigate the asymptotic behavior when $t \to \infty$. Taking into account the generating functional of the Laguerre polynomials[5] we get $F e^x \to \Sigma\,(-\lambda)^n L_n(x) = (1+\lambda)^{-1}\exp(\lambda x/1+\lambda)$ if $|\lambda| < 1$. For the power series expansion (3b) we get $\Sigma\,(-x)^q (q!)^{-1} \Sigma\,(-\lambda)^n c_q^n = (1+\lambda)^{-1}\Sigma\,(\lambda x/1+\lambda)^q (q!)^{-1}$ if $|\lambda| < 1$, leading of course to the same sum. In conclusion, if $M_0 = 1, M_1(0) \neq 1$, then $0 < M_1(t) \equiv M_1(0) < 2$ and $F \to M_1^{-1}\exp(-x/M_1)$ when $t \to \infty$. We note that in order to build F from the (a_n) we have to consider some restrictions ensuring the convergence of the sums in Eqs. (3a) and (3b). In the remainder of the paper we always take $M_0 \equiv M_1 \equiv 1$.

4. SOME GENERAL PROPERTIES OF THE TJON–WU DISTRIBUTION *F(x,t)*

It is outside the scope of the present paper to establish the existence and properties of $F(x, t)$ from arbitrary initial conditions at $t = 0$. The first property that we investigate is the positivity of $F(x, t)$. We shall study under what conditions the positivity of $F(x, t)$ at $t = 0$ can propagate to positive t values. Both the Laguerre polynomial and the power expansions being not very convenient for the positivity property, we shall use directly the integrodifferential equation (10). As a second property we shall establish a set of upper and lower bounds for the Laguerre orthogonal functions which are useful for the study of $F(x, t)$.

The third property that we consider is the existence of $F(x, t)$ from conditions at $F(x, 0)$ and we can use different approaches.

(i) F is the inverse Laplace transform of G. G is given by the expansion (4″) and built with the $(a_n(t))$. We can study the analytical properties in the ζ plane and the asymptotic growth of G. This requires bounds on $(|a_n(t)|)$ from given $a_n(0)$. Due to the great number of different possibilities for introducing the arbitrary constants in the system (2) such an approach is not easy if we expect simple cases like the pure solutions which is studied directly in Sec. 5.

(ii) Consider F given by the Laguerre expansion (3a). In the case of Laguerre polynomials the region of convergence is a parabola around the $x > 0$ axis with focus at $x = 0$ and F must satisfy certain growth conditions (Szego[5]).

(iii) Consider the power series given by Eq. (3b) and try to obtain sufficient conditions at $t = 0$ ensuring both the absolute convergence of the series at any $t \geqslant 0$ and the existence of sums $F(x, t)$ which are entire function in the x plane for all $t \geqslant 0$. Let us define the following with the modulus $|a_n(t)|$:

$$N_q(t) = \sum (n+2)(n+1)\cdots(n+3-q)|a_n(t)|,$$

$$N_0(t) = \sum |a_n(t)| \tag{13a}$$

and we get absolute upper bounds for the power series (3b):

$$|e^x F(x,t)| < 1 + \sum_q \frac{|x|^q}{(q!)^2} N_q(t). \tag{13b}$$

$N_q(t)$ is an absolute bound for the qth derivative of $H(u, t)$ at $u = -1$ or for the qth derivative of $e^x F(x, t)$ at $x = 0$. We want to get sufficient conditions at $t = 0$ such that $N_q(t)$ leads to entire x functions for the l.h.s. of Eq. (3b). Assume for instance that under well-defined conditions $N_q(t) \leqslant (\text{const})^q q!$, where the constant is t independent; then the l.h.s. of Eq. (14) is less than $\exp|x|$ const for any $t \geqslant 0$. In such a case we have absolute convergence for the sum (3b) and we can apply the Fubini theorem in order to justify the inversion of summations in Eq. (3a) and (3b). In order to get bounds on $N_q(t)$ we start with Eq. (2′) and differentiate $(q-1)$ times with respect to u. Equating to zero the coefficient of u^{3+n-q} we get a nonlinear differential system like Eq. (2) that we integrate from 0 to t [as Eq. (12)]:

$$\begin{aligned}(n+2)\cdots(n+3-q)\ a_n(t) &= e^{-t}[(n+2)\cdots(n+3-q)\ a_n(0) \\ &\quad + q(n+2)\cdots(n+4-q)\ a_n(0)] \\ &\quad - q(n+2)\cdots(n+4-q)\ a_n(t) + e^{-t} \\ &\quad \times \int_0^t e^{t'}\Big[2(n+2)\cdots(n+4-q)\ a_n(t') + \sum_{p=0}^{q-1} C_{q-1}^p \\ &\quad \times \sum_{m+m'=n-2} a_m(m+2)\cdots(m+3-p) \\ &\quad \times a_{m'}(m'+2)\cdots(m'+3-q+1+p)\Big]dt'. \end{aligned} \tag{14a}$$

We take the modulus of both sides, bound the r.h.s. by the sum of the modulus of the different terms, sum over n, and find

$$\begin{aligned}N_q(t) &< e^{-t}[N_q(0) + qN_{q-1}(0)] + qN_q(t) + e^{-t} \\ &\quad \times \int_0^t e^{t'}\Big[2N_{q-1}(t') \\ &\quad + \sum_{p=0}^{q-1} C_{q-1}^p\, N_q(t')N_{q-1-p}(t')\Big]dt'.\end{aligned}$$

The possibility of obtaining bounds for N_q, recursively, from the set $\{N_q(0)\}, p \leqslant q$ and $N_0(t)$ is clear.

(iv) It is also interesting to obtain sufficient conditions at $t=0$ such that $F\,e^{x/2}$ [which from Eq. (3) is expanded in a series of Laguerre orthogonal functions $e^{-x/2}L_n(x)$] is square integrable for all $t \geqslant 0$:

$$\int_0^\infty e^x F^2(x,t)dx = \sum a_n^2(t) = \widetilde{N}(t) \tag{15}$$

or equivalently to find conditions such that $\widetilde{N}(t \geqslant 0) < \infty$. This could be the starting point for a study of an expansion in L^2 space. However, let us remark that $\widetilde{N}(0)$ is not directly related to initial conditions at $t=0$ as is $N_0(0)$. In the companion paper[4] we introduce an Hilbert space of solutions so that the solution stays in this space at ulterior time if it is present at $t=0$.

There exists a class of solutions of Eq. (2) interesting because the sign of $a_n(t)$ does not change with t and is known with n. In these cases the positivity properties of the a_n play a crucial role for their study and for instance bounds on $N_q(t)$ are easily obtained.

A. Some results concerning the positivity property of $F(x,t)$ for $t \geqslant 0$

Because the positivity properties that we shall obtain are independent of the choice of the initial positive value t_0, for simplicity we choose $t_0 = 0$.

We integrate Eq. (10) to obtain an equation where $F(x,0)$ appears explicitly:

$$e^t F(x,t) = F(x,0) + A(x,t),$$

$$A(x,t) = \int_0^t e^{t'}\Big(\int_x^\infty \frac{dx'}{x'} \int_0^{x'} F(x'-x'',t')F(x'',t')dx''\Big)dt' \tag{16}$$

and consider the difference of F at x_1 and x_2:

$$e^t(F(x_1,t) - F(x_2,t)) = F(x_1,0) - F(x_2,0) + B(x_1,x_2,t),$$

$$B(x_1,x_2,t) = \int_0^t e^{t''}\int_{x_1}^{x_2}\frac{dx'}{x'} \times \int_0^{x'} F(x'-x'',t')F(x'',t')dx''dt''. \tag{17}$$

We assume $F(x,0) > 0$, $F(x,0) \underset{x\to+\infty}{\longrightarrow} 0$, and for $x > \bar{x}$ fixed, $F(x,0)$ is strictly decreasing, $F(x_1,0) > F(x_2,0)$ if $\bar{x} < x_1 < x_2$. Further we assume at this stage that there exists a finite interval Δt as small as we want such that $F(x,t) > 0$ for $0 \leqslant t < \Delta t$. Finally, we assume that $F(x,t)$ are smooth continuous functions going to zero when $x \to \infty$ in order that the solutions of the integral equations(10)–(16) exist.

First we want to show that the positivity property of F and the decreasing property $(x > \bar{x})$ at infinity of F propagate forward in time. At $t = \Delta t$ these properties are obvious because in A and B $(x_2 > x_1)$ we integrate with positive functions. Now $t > \Delta t$ and x finite $e^t \cdot F$ is a sum of two terms $F(x,0) > 0$ independent of t and $A(x,t)$ continuous in t and such that $F(x,\Delta t) > 0$. Consequently, for each finite x there exists a finite interval $\Delta \bar{t}$ such that $F(x, \Delta t + \Delta \bar{t}) > 0$. Similarly, in Eq. (17), $F(x_1,0) - F(x_2,0) > 0$ and $B(x_1,x_2,t)$, a continuous function of t, is such that $B(x_1,x_2,\Delta t) > 0$. For any couple of $x_1 > x_2$ values there exists a finite interval Δt where $F(x_1, \Delta t + \Delta \bar{t}) - F(x_2, \Delta t + \Delta \bar{t}) > 0$. Unfortunately, these intervals are (x_1, x_2) dependent; however, their existence contradicts the existence of a negative part of F for x very large and t sufficiently close to Δt because in such a case the above inequality between x_1 and x_2 must be reversed in order to satisfy $F < 0$, $\underset{x\to\infty}{\longrightarrow} 0$. On the other hand, for x finite, $\Delta \bar{t} > 0$ is also x dependent; however, its existence shows that for t sufficiently close to Δt a zero cannot appear at x finite.

Secondly, if for some $t_0 > 0$ and x_0 finite, $F(x_0,t_0) < 0$, then the negative property of F propagates backward in time and this property must have appeared first at $x = \infty$. On the one hand, in this case $A(x_0,t_0) < 0$ is larger in modulus than $F(x_0,0)$; on the other hands, there exists $\Delta \bar{t}$ such that $A(x_0, \Delta t + \Delta \bar{t}) > 0$ and consequently from the continuity property there exists t_1 strictly less than t_0 such that $A(x_0,t_1) = 0$. From Eq. (16) $F(x,t_1)$ must have some negative part and let us call x_1 the smallest x value such that $F(x_1,t_1) < 0$. With the same argument at above we define $t_2(x_2) < t_1(x_1) < t_0(x_0)$, where $F(x_2,t_2) < 0$ for the smallest x value and we go on. The strictly positive decreasing sequence $t_i(x_i)$ must have a limit $t_{\lim}$. The x values corresponding to $t_{\lim}$ cannot be finite otherwise $F(x,0) > 0$ is finite and the argument comparing $F(x,0)$ and $A(x,t_{\lim})$ works again and so the only possibility is that these x values are going to infinity. If we now let t decrease from $t_{\lim}$ up to zero, either we find $F(x,t) > 0$ or other negative parts coming from earlier time at x equals infinity. Let us now consider the smallest $t_{\lim}$. It cannot belong to the interval Δt. Now we go forward in time and use our previous argument. For $t < t_{\lim}$, $F(x,t) > 0$ and so $F(x,t_{\lim}) > 0$. For t sufficiently close to $t_{\lim}$, $F(x,t)$ cannot be negative for x finite, and $F(x,t)$ is decreasing for x large in such a way that a negative part of F cannot appear.

Concerning the existence of a small Δt domain where $F(x,t)>0$, we want to give plausibility arguments from positivity at $t=0$ using both continuity property and a first order expansion around $t=0$. From $F(x,0)>0$, $\lim_{t\to 0} A(x,t)=0$, and the continuity t property of $A(x,t)$ we see that for x finite, there exists a finite t interval (x dependent) such that $F(x,0)+A(x,t)>0$ so that $F(x,t)>0$. If we expand Eq. (16) up to order t^2 around $t=0$, we get

$$e^tF(x,t)\simeq F(x,0)+t\int_x^\infty \frac{dx'}{x'}\int_0^{x'} F(x'-x'',0)F(x'',0)dx'' + O(t^2) \qquad (16')$$

and we see that the two first order terms are positive if $F(x,0)>0$. From the assumed decreasing property of $F(x,0)$ for $x>\bar{x}$, and $\lim_{t\to 0} B(x_1,x_2,t)=0$, using the continuity property in t we see that there exists a finite t interval (x_1,x_2 dependent) such that $e^t(F(x_1,t)-F(x_2,t))>0$ when $x_1>x_2$. If we expand Eq. (17) around $t=0$, we get

$$e^t(F(x_1,t)-F(x_2,t)) \simeq F(x_1,0)-F(x_2,0)+t\int_{x_1}^{x_2}\frac{dx'}{x'} \times\int_0^{x'} F(x'-x'',0)F(x'',0)dx''+O(t^2) \qquad (17')$$

and we see that the two first order terms give $F(x_1,t)-F(x_2,t)>0$ for $\bar{x}<x_1<x_2$ or that $F(x,t)$ is also decreasing if $F(x,0)$ is decreasing. This excludes the possibility of a negative tail at large x [a negative tail and the condition $F(x,t)\to 0$ requires that $F(x,t)$ must be increasing when x is sufficiently large and goes to infinity]. As a final remark and always at the level of a plausibility argument we observe that if we let $F(x,0)$ be zero for a finite number of x_i values (x_i finite), then from Eq. (16′) the remaining first term proportional to t is positive and so for these x_i values and $t>0$ sufficiently small, $F(x,t)>0$.

B. A set of inequalities for the Laguerre orthogonal functions and application to the determination of a set of lower and upper bounds for $F(x,t)$

The study is done in Appendix A. Let us define $l_n(x) = e^{-x/2}L_n(x)$ and consider their Taylor series around $x=0$. We note that the even derivatives at $x=0$ are positive whereas the odd derivatives are negative. If we consider this Taylor expansion for $x>0$ and retain a finite number of terms, then we get lower and upper bounds

$$l_n(x)<\sum_{q=0}^{2p}\frac{x^q}{q!}\frac{\partial^q}{\partial x^q}(l_n(x))_{x=0},$$

$$l_n(x)>\sum_{q=0}^{2p-1}\frac{x^q}{q!}\frac{\partial^q}{\partial x^q}(l_n(x))_{x=0} - \frac{x^{2p}}{(2p)!}\frac{\partial^{2p}}{\partial x^{2p}}(l_n(x))_{x=0},$$

$$l_n(x)>\sum_{q=0}^{2p+1}\frac{x^q}{q!}\frac{\partial^q}{\partial x^q}(l_n(x))_{x=0},$$

$$l_n(x)<\sum_{q=0}^{2q}\frac{x^q}{q!}\frac{\partial^q}{\partial x^q}(l_n(x))_{x=0} - \frac{x^{2p+1}}{(2p+1)!}\frac{\partial^{2p+1}}{\partial x^{2p+1}}(l_n(x))_{x=0}.$$

If [see Eq. (4d)] we consider the case where the sign of $a_n(t)$ does not change with t and where either a_n or $(-1)^n a_n$ or c_k or $(-1)^k c_k$ does not change sign, then these inequalities are useful. If $(-1)^n a_n$ has always the same sign, then from Eq. (15) we get corresponding inequalities for the sum $e^{x/2}(F-1) = \Sigma(-1)^n a_n l_{n+2}(x)$ where the derivatives for l_n are replaced by the derivatives of the sum at $x=0$. If now $(-1)^a a_n$ change sign, we separate the sum in two parts: one negative and another positive and we supply the above inequalities for the two parts. If we have calculated some first terms of the expansion (3a), we can apply these results for the remaining part of the sum.

C. Convergence of the power series (3b), existence of $\widetilde{N}(t)$, and sufficient conditions on the sets $\{N_q(0)\}$ in order to have entire x functions for $e^xF(x,t)$

We assume that at $t=0$ we have absolute convergence in Eq. (3b) with the set $N_q(0)$. Equivalently, the r.h.s. of Eq. (13b) has an infinite radius of convergence at $t=0$ and we want to establish sufficient conditions in order to have the same property at $t\neq 0$. First we establish with the help of Eq. (12) bounds on $N_0(t)$ and $\widetilde{N}(t)$ from conditions on $N_0(0)$. Secondly, we find bounds on $\widetilde{N}(t)$ from conditions on $N(0)$. Thirdly, we recursively obtain bounds on $N_q(t)$ from conditions on $N_0(t)$ and $N_p(0)$, $p\leqslant q$, with the help of Eq. (14b).

1. *Bounds on* $N_0(t)=\sum_{n=n_0}^{\infty}|a_n(t)|$ *and* $\widetilde{N}=\sum_{n=n_0}^{\infty}a_n^2(t)$ *from conditions on* $N_0(0)$

We put $\gamma_n=(n+1)/(n+3)$ in Eq. (12), n_0 being either 0 or 1 or 2,...; we remark that the first nonlinear contribution appears for $n=2n_0+2$ and we have

$$|a_n(t)|<|a_n(0)|\exp(-\gamma_n t),\ \ n\leqslant 2n_0+1,$$

$$|a_n(t)|<\exp(-\gamma_n t)\Big[|a_n(0)|+(n+3)^{-1}\int_0^t\exp(\gamma_n t')\sum_{m+p=p-2}|a_m(t')||a_p(t')|dt'\Big],\ \ n\geqslant 2(n_0+1). \qquad (17a)$$

We remark that $\exp(-\gamma t)\int_0^t(\exp\gamma t')|f(t')|dt'$ is a decreasing γ function and $n+3\geqslant 2n_0+5$ in the second inequality; it follows

$$|a_n|<\exp(-\gamma_{n_0}t)\Big[|a_n(0)|+(2n_0+5)^{-1}\int_0^t\exp(\gamma_{n_0}t)\sum|a_m||a_p|dt'\Big], \qquad (17b)$$

where $\gamma_{n_0}=(n_0+1)/(n_0+3)$. Summing over n we get a nonlinear integral inequality:

$$N_0(t)\exp\gamma_{n_0} t \leqslant M_0(t),$$

$$M_0(t) = N_0(0) + (2n_0+5)^{-1}\int_0^t N_0^2(t')\exp(\gamma_{n_0} t')\,dt'. \tag{17c}$$

We remark that

$$-\frac{d}{dt}M_0^{-1}(t) = (2n_0+5)^{-1}M_0^{-2}(\exp\gamma_{n_0} t)N_0^2 < (2n_0+5)\exp(-\gamma_{n_0} t);$$

integrating both sides we get

$$N_0(0)(M_0(t))^{-1} > (1 - N_0(0)(\gamma_{n_0}(2n_0+5))^{-1}) + N_0(0)(\gamma_{n_0}(2n_0+5))^{-1}\exp(-\gamma_{n_0} t);$$

if $N_0(0) < \gamma_{n_0}(2n_0+5)$, we can substitute into the r.h.s. of Eq. (17c) and finally we find

$$N_0(t) \leqslant \frac{[(n_0+1)/(n_0+3)](2n_0+5)N_0(0)}{N_0(0) + ((n_0+1)(2n_0+5)/n_0+3 - N_0(0))\exp[((n_0+1)/(n_0+3))t]},$$

if

$$N_0(0) \leqslant \left(\frac{n_0+1}{n_0+3}\right)(2n_0+5). \tag{17d}$$

Thus, if $n_0 = 0, 1, 2, \ldots$, we must have $N_0(0) \leqslant 5/3, 7/2, 22/5, \cdots$. In all these cases we have $N_0(t) \leqslant N_0(0)$; further, the inequality $N(t) < N_0^2(t)$ is always true. It follows that if $N_0(0) \leqslant [(n_0+1)/(n_0+3)](2n_0+5)$, then $\widetilde{N}(t)$ is less than the square of the r.h.s. of the inequality (17b) and $\widetilde{N}(t) < \{[(n_0+1)/(n_0+3)](2n_0+5)\}^2$. If the inequality for $N_0(0)$ is strict, then $N_0(t) \to 0$ and $\widetilde{N}_0(t) \to 0$ when $t \to \infty$.

2. Bounds on $\widetilde{N}(t) = \sum_{n=n_0}^{\infty} a_n^2(t)$ from conditions on $\widetilde{N}(0)$

We start with the system (2), multiply by a_n, and integrate from 0 to t:

$$a_n^2(t) = a_n^2(0)\exp(-2\gamma_n t) < a_n^2(0)\exp(-2\gamma_{n_0} t),\quad n \leqslant 2n_0+1,$$

$$a_n^2(t) = \exp(-2\gamma_n t)\left[a_n^2(0) + (n+3)^{-1}\int_0^t (\exp 2\gamma_n t')\Big(2a_n(t')\sum_{m+p=n-2} a_m(t')a_p(t')\Big)dt'\right],\quad n \geqslant 2n_0+2. \tag{17a'}$$

Using both the Schwartz inequality $|\Sigma\, a_m a_p| < \Sigma\, a_m^2 < \widetilde{N}$ and majorations similar to the previous case, we find

$$a_n^2(t) < \exp(-2\gamma_{n_0} t)\left[a_n^2(0) + 2(n+3)^{-1}\right]\int_0^t (\exp 2\gamma_0 t')|a_n(t')|\widetilde{N}(t')dt',\quad n \geqslant 2n_0+2. \tag{17b'}$$

Summing over n and using the Schwarz inequality for $\Sigma\, |a_n|(n+3)^{-1}$, we find a nonlinear integral inequality

$$\widetilde{N}(t)e^{2\gamma_{n_0} t} \leqslant \widetilde{M}(t) = \widetilde{N}(0) + 2c_{n_0}\int_0^t (\exp 2\gamma_{n_0} t')\widetilde{N}^{3/2}(t')dt',$$

$$c_{n_0}^2 = \sum_{p=2n_0+5}^{\infty} P^{-2}. \tag{17c'}$$

We remark that

$$\widetilde{M}^{-3/2}\frac{d\widetilde{M}}{dt} = 2c_{n_0}(\exp 2\gamma_{n_0} t)\widetilde{N}^{3/2}\widetilde{M}^{-3/2} \leqslant 2c_{n_0}\exp(-\gamma_{n_0} t);$$

integrating both sides we get

$$\widetilde{M}^{-1/2}(t) > \widetilde{N}^{-1/2}(0) - (c_{n_0}/\gamma_{n_0}) + (c_{n_0}/\gamma_{n_0})\exp(-\gamma_{n_0} t);$$

if $\widetilde{N}^{1/2}(0) < \gamma_{n_0}/c_{n_0}$, we can substitute into the r.h.s. of Eq. (17c') and finally we find

$$\widetilde{N}(t) \leqslant \frac{([(n_0+1)/(n_0+3)](1/c_{n_0}))^2\widetilde{N}(0)}{\{\widetilde{N}(0)^{1/2} + ((n_0+1)/[(n_0+3)c_{n_0}] - \widetilde{N}(0)^{1/2})\exp[((n_0+1)/(n_0+3))t]\}^2}$$

if

$$\widetilde{N}(0) \leqslant \left(\frac{n_0+1}{n_0+3}\right)^2\left(\sum_{2n_0+5} P^{-2}\right)^{-1}. \tag{17d'}$$

Thus, if $n_0 = 0, 1, 2, \ldots$, we must have $\widetilde{N}_0(0) \leqslant 0.502, 1.628, 3.063, \cdots$. In all these cases we have $\widetilde{N}(t) \leqslant \widetilde{N}(0)$ and if the inequality is strict then $\widetilde{N}(t) \to 0_{t\to\infty}$.

3. Bounds on $N_q(t)$, $q\geqslant 1$.

The study is done in Appendix B and we find sufficient conditions at $t=0$ in order that the power series (3b) be absolutely convergent for $t\geqslant 0$. We find the following theorem from Eq. (14b):

Theorem: If $N_0(t)\leqslant N_0(0)$ and if $N_q(0)\leqslant q!\, N_0(0)(4+N_0(0))^{q-1}(N_0(0)+2/q)$ for $q\geqslant 1$, then we have $N_q(t)<q!\,N_0(0)(4+N_0(0))^q$.

We have also a system of inequalities (14b) such that we can explicitly construct upper bounds on $N_q(t)$ from the bounds (17d) on $N_0(t)$.

E. Particular properties for solutions satisfying assumptions on the sign of the $a_n(t)$

We assume that for n fixed and $t\geqslant 0$, the sign of $a_n(t)$ does not change. Further, we restrict our study to three classes: (i) in the first class a_n has the same sign for all n: class I(i), $a_n(t)<0$ and class II(i), $a_n(t)>0$; (ii) in the second class $(-1)^n a_n$ has always the same sign: class I(ii), $(-1)^{n+1}a_n>0$ and class II(ii), $(-1)^n a_n>0$; (iii) in the third class the only $a_n=c_k\neq 0$ are for $n=P-1+k(P+1)$, P integer, and $k=0,1,2,\cdots$ and $(-1)^k c_k$ has always the same sign: I(iii) $(-1)^k c_k>0$, class II(iii), $(-1)^k c_k>0$.

In all the considered classes the solutions (a_n) can have an infinite number of arbitrary constants introduced either with $\bar{a}_n$ or $a_n(0)$. As an illustration let us show how we can construct such examples from the knowledge of the signs of $\bar{a}_n$ or $a_n(0)$ and from the recursive properties of Eq. (12).

First we consider solutions constructed from the $\bar{a}_n$ and remark from Eq. (12) that $a_n(t)$ has a well defined sign for $t\geqslant 0$ if $\bar{a}_n\ddot{a}_n(t)<0$: I(i), if $\bar{a}_n<0$, we get $a_n(t)\leqslant 0$ and it follows from Eq. (1′) that $M_2>M_3>\cdots>M_q(t)>\cdots$; I(ii), if $(-1)^{n+1}\bar{a}_n>0$, we get $(-1)^{n+1}a_n(t)>0$; I(iii), the only $\bar{a}_n\neq 0$ are for $n=P-1+k(P+1)$ where P is a fixed integer and $k=0,1,2,\cdots$. If we define $\bar{a}_n=\bar{c}_k$ and $a_n=c_k$, we get that if $(-1)^k\bar{c}_k\geqslant 0$ then $(-1)^k c_k(t)\geqslant 0$.

Secondly, we consider the solutions constructed from the $a_n(0)$ and remark from Eq. (2) that $a_n(t)$ has a well defined sign for $t\geqslant 0$ if $a_n(0)\dot{a}_n(t)>0$: II(i), if $a_n(0)\geqslant 0$, we get $a_n(t)\geqslant 0$ and from Eq. (1′), $M_n(t)>0$; II(ii), if $(-1)^n a_n(0)\geqslant 0$, we get $(-1)^n a_n(t)>0$; II(iii), the only $a_n(0)=a_k(0)\neq 0$ are for $n=P-1+k(P+1)$, where P is a fixed integer; if $(-1)^{k+1}c_k(0)\geqslant 0$, we get $(-1)^{k+1}c_k(t)\geqslant 0$.

In all these cases the study is done in Appendix B where we find the following:

In the cases I(i), (ii), (iii) we have $N_q(t)\leqslant N_q(0)\times\exp(-t/3)\leqslant N_q(0)$. From our assumption about the set $N_q(0)$ it follows that Eq. (3b) is absolutely convergent for any $t\geqslant 0$ and define entire x functions for $e^x F(x,t)$. For instance, it is sufficient that $N_q(0)/q!(N_0(0))^q\leqslant$ any polynomial in N_0, independently of the index q.

In the II(i), (ii), (iii) cases we have the following theorem:

Theorem: If $N_0(t)\leqslant N_0(0)$ and $N_q(0)\leqslant q!\, N_0(0)(2+N_0(0))^q$, then for any $t\geqslant 0$ we have $N_q(t)<q!\, N_0(0)(2+N_0(0))^q$. It follows that we have for $t\geqslant 0$ the absolute convergenece in Eq. (3b) and Fe^x is an entire x function. Due to the assumed restrictions about the sign of $a_n(t)$ we can, in a way different of the previous one [Eq. (17d)], derive bounds on $N_0(t)$ [see Eq. (B11)] which unfortunately do not improve the previous ones. Finally, we note that in all these cases we find a system of inequalities [see Eq. (B7)] such that we can explicitely determine upper bounds on $N_q(t)$ from known bounds on $N_0(t)$.

E. Results concerning the existence of $\widetilde{N}(t)$ in the case of a violation of mass and energy conservation laws

In Appendix B. 3, it is shown that $a^2_{-2}(t)+a^2_{-1}(t)+\Sigma_0\, a_n^2(t)$ is bounded for $t>0$ (or even $t\to\infty$) if either $\Sigma_{-2}\, a_n^2(0)\leqslant 6\pi^{-2}$ or $\Sigma_{-2}\,|a_n(0)|\leqslant 1$.

5. PURE SOLUTIONS

We consider the class of solutions $(a_n(t))$ of Eq. (12) where we introduce, as arbitrary constants, only the $\bar{a}_n=\lim_{t\to\infty}\{\exp[(n+1)/(n+3)]t\}a_n(t)$ and in fact a finite number of such $\bar{a}_n$. While these solutions are easily determined recursively, they fail to satisfy $F(x,t)>0$ (with the exception of the Krook–Wu particular solution).

A. Fundamental solutions: $\bar{a}_{P-1}\neq 0$, P integer >1

The set $\{\bar{a}_n\}$ has only one element $\bar{a}_{P-1}\neq 0$. Assuming

$$a_n=\delta_n\exp\left(-t\frac{P}{P+2}\right)\left(\frac{n+2}{P+1}\right),\quad n\geqslant P-1$$

and substituting into Eq. (2) or (12), we find that all the δ_n can be determined recursively from $\delta_{P-1}=\bar{a}_{P-1}$ through the relation

$$\left[-(n+3)\,\frac{P}{P+2}\left(\frac{n+2}{P+1}\right)+(n+1)\right]\delta_n=\sum_{m+m'=n-2}\delta_m\delta_{m'}.$$

Many δ_n being zero, it is more convenient to consider

$$a_{P-1+k(P+1)}=d_k\exp\left(-t\frac{P(k+1)}{P+2}\right),\quad k=0,1,2,\cdots,$$

$$-k\left(k+\frac{P^2-2}{P(P+1)}\right)d_k=\frac{P+2}{P(P+1)}\sum_{m+m'=k-1}d_m d_{m'}. \tag{18b}$$

For $k=0$, the l.h.s. of Eq. (18b) is zero and we put $d_0=\bar{a}_{P-1}$. For $k=1$ the l.h.s. is $\neq 0$, the r.h.s. is proportional to d_0^2, and we obtain d_1. For $k=2$ we obtain d_2 from d_0 and d_1, and so on. All the d_k can be calculated from d_0. Let us notice the following scaling property: If we define $d_k=\bar{d}_k(\bar{a}_{P-1})^{k+1}$, $\bar{d}_0=1$, then the $\bar{d}_k$ satisfy the same relation (18b). Another way to characterize these solutions is to look at Eq. (5). Substituting Eq. (18a) and (18b) into $H(u,t)$, we find

$$G_P=\sum_k d_k(\omega_P)^k,\quad H(u,t)=\omega_P G_P,$$

$$\omega_P = \exp\Big[-\Big(\frac{P}{P+2} t - (P+1)\log u \Big) \Big]. \qquad (18c)$$

These solutions depend on only one variable combination of $\log u$ and t which from Eq. (5) are solutions of the nonlinear differential equation

$$\Big\{ \omega_P \frac{\partial^2}{\partial \omega_P^2} + \Big(2 - \frac{P+2}{P(P+1)} \Big) \frac{\partial}{\partial \omega_P} + \frac{P+2}{P(P+1)} G_P \Big\} G_P = 0.$$

These pure solutions correspond also to particular constraints for the moments $M_n(t)$. Because $a_0(t) \equiv a_1(t) \equiv \cdots \equiv a_{P-2}(t) \equiv 0$ for these solutions, then from Eq. (1) we get $M_2(t) \equiv M_3(t) \equiv \cdots \equiv M_P(t) \equiv 1$.

Investigating the properties of the solutions of the system (18a) and (18b) we easily get for large t

$$|a_n(t)| < \frac{a^n}{n} \quad 0 < a = c_1 e^{-c_2 t} \ll 1.$$

c_1 and c_2 being constants $t > t_0$. (18d)

It follows that in the u plane the corresponding solutions $H(u,t)$ have a finite radius of convergence and, as we shall see, in the x plane, the $F(x,t)$ are entire functions. In fact, from the power expansion (3b) and the bound (18d) we find

$$|e^x F(x,t)| < 1 + \text{const} \sum \Big| \frac{xa}{1-a} \Big|^q \frac{1}{q!} = 1 + \text{const} \exp\Big(\frac{|x|a}{1-a} \Big). \qquad (18e)$$

In the following we label these solutions $\bar{a}_{P-1} \neq 0$ as $P/(P+2)$ pure solutions, emphazing the fact that the first $a_n \neq 0$ decreases like $\exp[-tP/(P+2)]$ while the other $a_n \neq 0$ decrease like powers of this time dependent term.

B. Mixing of two pure solutions $P_0/(P_0+2)$, $P_1/(P_1+2)$, $P_0 < P_1$ with only $\bar{a}_{P_0-1} \neq 0$, $\bar{a}_{P_1-1} \neq 0$

If we put

$$a_n(t) = \sum_r d_n^{(r)} \exp\Big[-t \Big(\frac{P_0(n+2)}{(P_0+1)(P_0+2)} \Big) + \theta r \Big],$$

$$\theta = \frac{(P_0 - P_1)(P_0 P_1 - 2)}{(P_0+1)(P_0+2)(P_1+2)} \qquad (19a)$$

and substitute into Eq. (12) or (2), we find that the $\{d_n^{(r)}\}$ can be recursively determined

$$\Big[-(n+3)\Big(\frac{P_0(n+2)}{(P_0+2)(P_0+1)} + \theta r \Big) + (n+1) \Big] d_n^{(r)} = \sum_{\substack{p+q=n-2 \\ s+t=r}} d_q^{(s)} d_q^{(t)}. \qquad (19b)$$

The study is done in Appendix C. If $\theta \neq 0$, then the number of different time dependences in $a_n(t)$ cannot stay finite when $n \to \infty$.

C. Particular mixing of $P_0 = 1$ and $P_1 = 2$

From Eq. (19a) we see that $\theta = 0$ if $P_1 P_2 = 2$ which means $P_0 = 1$ and $P_1 = 2$ and the series for a_n reduces to one term (like the pure solution). For this particular mixing $P_0/(P_0+2) = \frac{1}{3}$ and $P_1/(P_1+2) = \frac{1}{2}$ we rewrite Eqs. (19a) and (19b) as

$$a_n(t) = d_n \exp\Big(-\frac{t}{6}(2+n) \Big),$$

$$n(n-1) d_n = -6 \sum_{M+N=n-2} d_M d_N. \qquad (20a)$$

We introduce two arbitrary constants $d_0 = \bar{a}_0$ and $d_1 = \bar{a}_1$ and all the d_n are determined recursively. If $\bar{a}_0 = -a_0^2$, $\bar{a}_1 = 2a_0^3$ (a_0 being a constant), then $d_n = (-1)^{n+1} a_0^{n+2}(n+1)$ and we recover the particular Krook–Wu solution which thus appears as the mixing of the pure 1/3, 1/2 solutions with particular relations between $\bar{a}_0$ and $\bar{a}_1$. Equation (20a) suggests that there exist for $H(u,t)$ solutions with only one variable in the u plane. We define

$$H(u,t) = -\omega^2 G(\omega)\ \omega = \exp[-(t/6 - \log u)] \qquad (20b)$$

and substituting into Eq. (5) we get

$$\frac{\partial G}{\partial \omega} = \pm 2\Big(G^3 + \frac{d_1^2}{4} + d_0^3 \Big)^{1/2}, \quad G = -\sum d_n \omega^n. \qquad (20c)$$

If $d_1^2 + 4d_0^3 = 0$, we recover the particular Krook–Wu solution $G[a\omega + 1]^{-2}$. Otherwise we have the meromorphic Weierstrass elliptic function with coefficients determined by Eq. (20a). In this case also the series (20c) has a finite radius of convergence. Furthermore, whatever values we consider for $\bar{a}_0$ and $\bar{a}_1$, investigating the recurrence relation (20a), one can show that Eq. (18d) holds, the series (3b) is absolutely convergent, and $F(x,t)$ defines an entire function in the x plane.

D. Mixing of q arbitrary pure solutions $\bar{a}_{P_0-1} \neq 0$, $\bar{a}_{P_1-1} \neq 0, \ldots, \bar{a}_{P_{q}-1} \neq 0$, $P_0 < P_1 \cdots < P_{q-1}$

We define

$$a_n(t) = \sum_{\substack{r_j \\ j=1,\ldots,q-1}} d_n^{r_1 r_2 \cdots r_{q-1}} \times \exp\Big\{ -t \Big[\frac{P_0}{P_0+2}\Big(\frac{n+2}{P_0+1} \Big) + \sum_{j=1}^{q-1} \theta_j r_j \Big] \Big\}, \qquad (21)$$

$$\theta_i = \frac{(P_0 - P_i)(P_i P_0 - 2)}{(P_0+1)(P_0+2)(P_i+2)},$$

and substituting into Eq. (2) we find that the $d_n^{r_1,\ldots,r_q}$ can be recursively determined (see Appendix C). Only for $P_0 = 1$, $P_i = 2$ can we have $\theta_i = 0$.

E. Study of the positivity property of $F(x,t)$ for the pure solutions

We consider the pure solutions ($\bar{a}_{P_n-1} \neq 0$) and the mixing of a finite number of such solutions. We seek whether there exists t_0 large but finite such that for $t \geqslant t_0$, $e^x F(x,t)$ remains positive (if the positivity is obtained at t_0, then we can apply the results of Sec. 4 C). The difficulty arises when t and x are both large. At fixed x, $\lim_{t\to\infty} F(x,t) e^x = 1$ but this result cannot guarantee the positivity in all the possible asymptotic directions of the (x,t) plane. The positivity property must remain true when we link t and x, considering for instance an $x(t)$ dependence such that $x(t) \to \infty$ when $t \to \infty$. Along this asymptotic direction $x(t)$ we want to define a dominant part when t is large and a criteria necessary for the positivity property (but not sufficient). We define a scaling variable ω equal to x multiplied by some decreasing time function such that $e^x F(x,t) \equiv \mathscr{F}(\omega, t)$. After that we take the limit $t \to \infty$ at fixed ω. If $\mathscr{F}(\omega, t = \infty) > 0$, then by continuity we can hope to find t_0 such that $\mathscr{F}(\omega, t \geqslant t_0) > 0$. If $\mathscr{F}(\omega, t = \infty)$ is not positive, then $e^x F(x,t)$ cannot stay positive for large t. In this way the problem is simpler because we are faced with only one variable. In order to understand clearly we begin with the Krook–Wu particular solution which can be rewritten $\omega = xa$, $a = \bar{a}_0 \exp(-t/6)$, $x \in [0, \infty]$:

$$\mathscr{F}(\omega,t) = e^x F(x,t) = \frac{(1 - 2a + \omega(1-a)^{-1})}{(1-a)^2} \exp[-(\omega/1-a)],$$

$$\mathscr{F}(\omega, t = \infty) = (1+\omega)e^{-\omega}.$$

We get two cases following the sign of $\bar{a}_0$: (i) $\bar{a}_0 > 0$, then $\omega > 0$ and $\mathscr{F}(\omega, \infty) > 0$; reintroducing $t \neq \infty$ or $a \neq 0$, we get $\mathscr{F}(\omega, t) > 0$ for $2a < 1$. (ii) $\bar{a}_0 < 0$; then $\omega < 0$ and $\mathscr{F}(\omega, \infty)$ has always a fixed zero at $\omega = -1$ or $x = (-\bar{a}_0)^{-1}\exp(t/6)$. Reintroducing $t \neq 0$ we see that the zero does not disappear and its location tends to this value at large t. This example is an illustration of the fact that while at fixed x, $\lim_{t\to\infty} e^x F \to 1$, then $F(x,t)$ can violate positivity.

1. Pure solutions ($P/(P+2)$)

We define $\omega = ax \geqslant 0$, $a = |\bar{a}_{P-1}| \times \exp\{-tP/[(P+1)(P+2)]\}$, and the Laguerre expansion (3a) can be written

$$\mathscr{F}(\omega,t) = 1 + \sum_k d_k \sum_{q=0}^{\lambda(k)} \frac{\omega^{\lambda(k)-q}(-a)^q}{(\lambda(k)-q)!} C_{\lambda(k)}^{\lambda(k)-q}, \tag{22a}$$

$$\lambda(k) = (P+1)(k+1),$$

where d_0 is positive or negative depending upon the sign of $\bar{a}_{P-1}$. If we rewrite Eq. (22a) as a power ω expansion, from the bound (18a) we get absolute convergence, The coefficient of $\omega^{\lambda(k)}$ is a constant plus terms going to zero when $t \to \infty$. So we take the limit $t \to \infty$ at fixed ω and get for the dominant part

$$\mathscr{F}(\omega, t = \infty) = 1 + \sum d_k \frac{\omega^{\lambda(k)}}{(\lambda(k))!},$$

where the coefficients d_k are given by the recurrence relation (18b) with $d_0 = \pm 1$. (Note that for ω fixed, t and x can go to infinity in a linked way.)

If $d_0 = -1$, then $d_k < 0$, $\mathscr{F}(\omega, t = \infty)$ is strictly decreasing for $\omega > 0$ and has always one zero. If $d_0 = +1$, then d_k alternate $[(-1)^k d_k > 0]$. We have studied numerically the function (22b) for $d_0 = 1$, $P = 1, 2, \ldots, 35$. We have always found at least one zero. For these cases $\mathscr{F}(\omega, t = \infty)$ cannot always be positive and there does not exist t_0 such that $e^{-x}F(x,t) > 0$ for $t \geqslant t_0$, $\forall x[0, \infty]$.

2. Particular mixing 1/3 and 1/2 of two pure solutions

From Eq. (20a), $a_n = (a(t))^{2+n} d_n$ with $a = \exp(-t/6)$, we define $\omega = xa(t)$, $\omega \geqslant 0$ and substituting into the Laguerre expansion we find

$$\mathscr{F}(\omega,t) = 1 + \sum d_n \sum_{q=0}^{n+2} (-a)^q \frac{\omega^{n+2-q}}{(n+2-q)!} C_{n+2}^{n+2-q}. \tag{23a}$$

Using the absolute convergence property of the corresponding power ω series, we take the limit $t \to \infty$ at fixed ω and obtain

$$\mathscr{F}(\omega, t = \infty) = 1 + \sum_{q=0}^{n+2} \frac{d_n \omega^{n+2}}{(n+2)!}. \tag{24a}$$

The analysis of the positivity property is done in Appendix C. From numerical analysis we have not found any positive $\mathscr{F}(\omega, t = \infty)$ solution other than the particular Krook–Wu solution.

In conclusion for the soliton like solutions (pure $a_{P-1} \neq 0$ or mixing 1/3, 1/2) we have not found any solution displaying the positivity property, except the particular Krook–Wu one.

3. Other mixing (finite number) of pure solutions

The analysis is done in Appendix C. If we consider a mixing of a finite number of pure solutions $P_1, P_2, \ldots, P_q$, rescale the variables in Eq. (3b) introducing appropriate scaling variables ω, there we are faced with a dominant series coming from the largest pure solution $P_q/(P_q+2)$ and thus we have not found any positive case.

Fortunately, for an infinite mixing this rule does not apply and we can get positivity property for $e^x F(x,t)$ as we shall see in the next section when we consider solutions defined by the set $\{a_n(0)\}$.

F. Infinite mixing of pure solutions

If for $m \leqslant n-2$ the coefficients $d_m^{(r)}$ and $b_m^{(r)}$ of $a_m(t) = \Sigma_r d_m^{(r)} \exp(-b_n^{(r)} t)$ are known, then it is shown in Appendix C that $a_n(t)$ is still of this type and the corresponding coefficients can explicitly be determined.

6. POSITIVE SOLUTIONS

We call positive solutions those built in Eq. (12) with the arbitrary constants $a_n(0)$. Although all these solutions do not lead to $F(x,t) \geqslant 0$, among these solutions there exists a subclass leading to $F(x,t) \geqslant 0$. Following the results of Sec. 5 they can be generated by conditions at $t = 0$, i.e., $F(x, 0) > 0$.

A. The fundamental positive solutions $a_{P-1}(0) \neq 0$, $P_{\text{integer}} \geqslant 1$

The set $\{a_n(0)\}$ has only one element $a_{P-1}(0) \neq 0$. From Eq. (12) we find that only the $a_n(t) \neq 0$ are restricted to $n = P - 1 + (k+1)(P+1)$. We define $\lambda(k) = (k+1)(P+1)$, $k = 0, 1, \cdots; a_{n=-2+\lambda(k)}(t) = c_k(t)$ and substituting into Eq. (12) we find

$$c_0(t) = a_{P-1}(0) e^{-t[P/(P+2)]},$$

$$c_k(t) = \frac{1}{1+\lambda(k)} \exp\left[-\left(\frac{\lambda(k)-1}{\lambda(k)+1}\right)t\right] \times \int_0^t \exp\left[\left[\frac{\lambda(k)-1}{\lambda(k)+1}\right]t'\right]$$

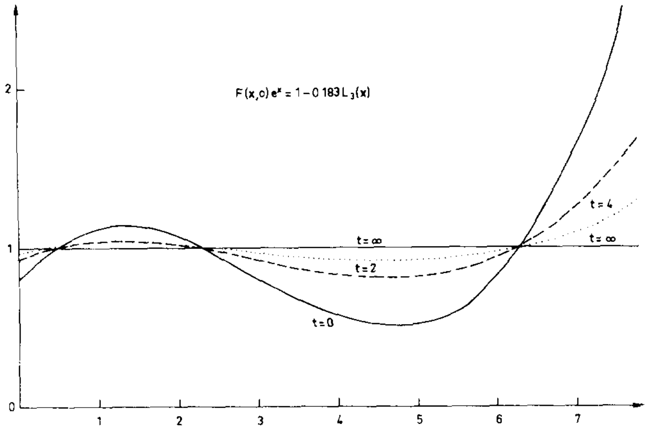

FIG. 1. Evolution in time of $f(x,t)$ for a $L_3(x)$ term at $t = 0$.

$$\times \sum_{m+m'=k-1} c_m(t')c_{m'}(t')dt' \tag{25}$$

and the c_k can be obtained recursively from $c_0(t)$. When k increases in c_k, the number of terms with different time dependences increases also. The least decreasing behavior is $\exp\{-([\lambda(k)-1]/[\lambda(k)+1])t\}$, leading to $\exp(-t)$ when k goes to infinity. In contrast, the fundamental pure solutions decrease like the terms of a geometrical series with a variable $[\exp(-\text{const } t)]$. For $k \neq 0$ we see that $c_k(t) \to 0$, when $t \to 0$ or $t \to \infty$. The Laguerre expansion (3a) becomes

$$e^x F(x,t) = 1 + \sum c_k(t)(-1)^{\lambda(k)} L_{\lambda(k)}(x),$$

$$e^x F(x,0) = 1 + a_{P-1}(0)(-1)^{P-1} L_{P+1}(x).$$

The sufficient conditions in order to get absolute convergence for the corresponding (3b) power series have been given in Sec. 4. They guarantee both that $e^x F(x,t)$ is an entire function in the x plane and the convergence of the expansions (3). We add other properties due to the sign of $a_{P-1}(0)$:

(i) If $a_{P-1}(0) > 0$, from Eq. (25) using induction we find that $a_n(t)$ or $c_k(t)$ are positive and so for the corresponding moments we have $M_n(t) > 0$. Further for $k \neq 0$ we have dC_k/dt positive at $t = 0$ and tends to 0^- when $t \to \infty$, whereas $C_0(t)$ is always decreasing. Further if P is odd, then $(-1)^{\lambda(k)} > 0$ and Eq. (26) is a sum of Laguerre polynomials with positive coefficients. It follows that for the power series associated to $F(x,t)e^{x/2}$, if we retain in this case only an odd or even number of first terms, then we get lower or upper bounds for the sum.

(ii) If $a_{P-1}(0) < 0$, we get that $C_k(t)(-1)^{k+1} > 0$ and for $k \neq 0$ the derivative $(-1)^{k+1}(d/dt)C_k(t)$ is positive at $t = 0$ and tends to 0^- when $t \to \infty$.

Positivity of $F(x,0)$: When x is large the dominant term is $a_{P-1}(0)[x^{P+1}/(P+1)!]$ and we must consider $a_{P-1}(0) > 0$. This restriction is not sufficient because $L_n(x)$ has n positive zeros and in general $F(x,0)$ will have negative parts. However, when $a_{P-1}(0)$ is zero, $e^x F(x,0)$ reduces to 1 so that for finite P we can always find $a_{P-1}(0)$ sufficiently small in order that $Fe^{+x} \geqslant 0$. For instance for $P = 1$ we get $0 < a_0(0) \leqslant 1$; for $P = 2$, we get $0 < a_1(0) < 1/(1+\sqrt{3})$, and so on. Because the oscillations of $L_{P+1}(x)$ become larger when P increases we have to retain smaller $a_{P-1}(0)$. In Fig. 1 we plot the ratio $f(x,t) = F(x,t)/F(x,\infty)$ in the two cases $e^x F(x,0)$ equals $1 + L_2(x)$ and $1 - 0.183 L_3(x)$ for different t values. In Fig. 2 we see that the second zero of $f(x,t) - 1$, when t increases from 0, begins to move to the left towards smaller x values. This effect leads to a small x interval (for not too large t values) where $f(x,t)$ is slightly larger than $f(x,0)$ or $f(x,\infty)$. In Fig. 1 the zeros of $f(x,t) - 1$ move very slowly and we do not observe the preceding effect.

We have also numerically considered cases where $F(x,0)$ can be negative. For $a_0(0) > 1$, $F(x,0)$ is not positive; however, then there exists $t_0 > 0$ such that $F(x,t) > 0$ for $t \geqslant t_0$ [for instance, for $a_0(0) = 1,5$ we have found $t_0 = 1.05$]. For other fundamental positive P solutions we have also verified that when $a_{P-1}(0) > 0$ is larger than the value for which $F(x,0) > 0$, then there exists t_0 that $F(x,t) > 0$ for $t \geqslant t_0$.

B. Mixing of different fundamental positive solutions

If we compute directly Eq. (12) for small m values, we verify that a_m can be written

$$a_m(t) = \sum_r d_m^{(r)} \exp(-b_m^{(r)} t), \tag{27}$$

where the least decreasing term is $\exp\{-[(m+1)/(m+3)]t\}$ and the sum over r contains a finite (increasing with m) number of terms. Assuming for $m = 0, 1, \ldots, n-2$ that a_m is of the Eq. (27) type with known $(d_m^{(r)})$ and (b_m^r) it is shown in Appendix C [by substitution into Eq. (12)] that $a_n(t)$ is also of this type and the corresponding coefficients can be effectively determined.

First we consider a *finite mixing* of fundamental positive solutions. We have numerically calculated the simple mixing corresponding to $e^x F(x,0) = a_0(0)L_2 + a_1(0)L_3$ and we have not found features different from the previous fundamental solutions. For instance, for $a_0(0) = 0.9$, $a_1(0) = 0.1$, then for $t \neq 0$, $e^x F(x,t)$ develops a behavior similar to Fig. 1. Let us notice that for a finite mixing and $t = 0$, when x is large the dominant behavior is provided by the highest Laguerre polynomial and the corresponding $a_n(0)$ must be positive in order to get positivity. It follows that the corresponding initial conditions are such that $e^x F(x,0) \to +\infty$ when $x \to \infty$. If we want to relax this technical constraint and

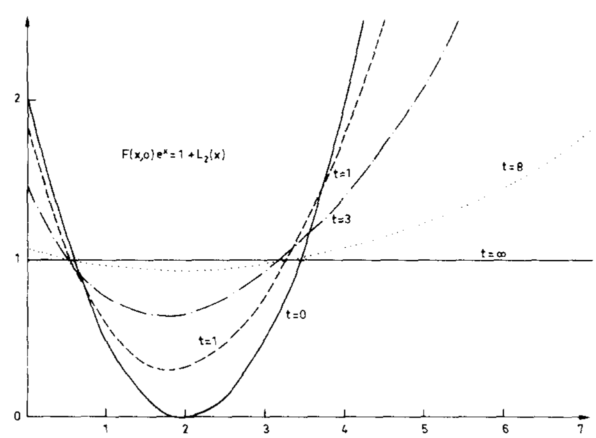

FIG. 2. Evolution of $f(x,t)$ for a $L_2(x)$ term at $t = 0$.

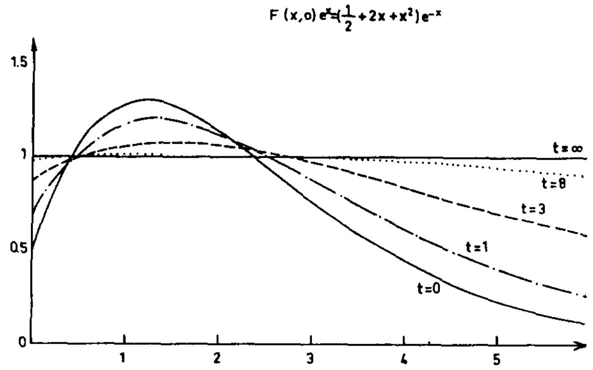

FIG. 3. Evolution of $f(x,t)$ for a product of an exponential by a polynomial at $t = 0$.

include cases where $e^xF \to 0$ when $x \to \infty$, we muct necessarily consider an infinite mixing.

Secondly, we consider an *infinite mixing* of fundamental positive solutions $(a_n(0))$ such that $e^xF(x,0)$ is both positive and decreasing when $x \to \infty$. It is not very easy to characterize the sufficient conditions on the elements of the set $a_n(0)$ ensuring these properties for the sums of the Laguerre polynomials expansions. We follow another simpler method where these sums can be written down in closed form. We start with the generating functional of the Laguerre polynomials

$$1 + \sum_1^\infty z^n L_n(x) = (1-z)^{-1}\exp\left(\frac{xz}{z-1}\right), \quad |z|<1, \quad (28)$$

where z is a parameter, and get from it simple examples where in the Laguerre polynomial expansion the coefficient of $L_1(x)$ is zero (conservation law for M_1).

(i) We consider linear combinations of Eq. (28) and of derivatives with respect to z. For instance, if we take into account derivatives of the first and of the second order, we get

$$\begin{aligned}1 + \sum_2 z_1^n(n-1)\left(\frac{z_2 n}{2}-1\right)L_n(x) &= (1-z_1)^{-2}\left[1-2z_1+\frac{xz_1}{1-z_1}+\frac{z_2z_1^2}{1-z_1}\right.\\ &\left.\quad\times\left(1-\frac{2x}{1-z_1}+\frac{x^2}{2(1-z_1)^2}\right)\right]\exp\left(\frac{xz_1}{z_1-1}\right),\end{aligned}$$

which are essentially decreasing exponential functions $(0<z_1<1)$ multiplied by polynomials of the second order in x. We can go on, take into account higher order derivatives of Eq. (28), and get an exponential multiplied by polynomials of arbitrary order in x.

(ii) Another simple family can be obtained from a linear combination of Eq. (28) for two different z values:

$$\begin{aligned}1 - z_1z_2\sum_2 L_n(x)\left[\sum_{P=0}^{n-2} z_1^P z_2^{n-P-2}\right] &= (z_2-z_1)^{-1}\left[\frac{z_2}{1-z_1}\exp\left(\frac{xz_1}{z_1-1}\right)\right.\\ &\left.\quad-\frac{z_1}{1-z_2}\exp\left(\frac{xz_2}{z_2-1}\right)\right], \quad 0<z_i<1. \quad (30)\end{aligned}$$

In these examples Eqs. (29) and (30) represent $e^xF(x,0)$ so that the arbitrary parameters z_1, z_2 must be restricted in such a way that the l.h.s. of these equations corresponds to positive functions for $x \geqslant 0$.

If in Eq. (29) we put $z_2 = 0$, the polynomial in z at the l.h.s. is of first order and we recognize the particular Krook–Wu example e^xF at $t=0$. For this simple case we know explicitly the $a_n(t)$ dependence (Sec. 5) while in the other cases we have to use the formalism previously developed in order to construct the $a_n(t)$ from $a_n(0)$.

If in Eq. (29) either $z_1 = 1$ or $z_2 = \frac{2}{3}$, we see that the coefficient of either $L_2(x)$ or $L_3(x)$ is zero. When the corresponding l.h.s. of Eq. (29) is given as input; set $(a_n(0))$ in Eq. (12) and then either $a_0(t)$ or $a_1(t)$ remains zero for $t>0$.

It is also possible to build examples with $a_0(t) \equiv a_1(t) \equiv 0$ starting at $t=0$ with Laguerre expansions without $L_2(x)$ and $L_3(x)$ components. For instance, we can take, as an initial condition at $t=0$, an appropriate linear combination of Eq. (28) and of the first, the second, and the third derivatives with respect to z. In conclusion, if in Eq (29) or (30) we restrict z_1 and z_2 to values such that the r.h.s. are $e^xF(x,0) > 0$, put the coefficients of the l.h.s. as initial values in Eq. (12) and then generate positive solutions $F(x,t)$ which are mixing of an infinite number of fundamental positive solutions.

In Fig. 3 we plot $f(x,t) = F(x,t)/F(x,\infty)$ corresponding to $z_1 = \frac{1}{2}$ and $z_2 = \frac{1}{4}$ in Eq. (29). We see that the second zero of $f-1$ moves to the right and there is a small energy range where $f(x,t)$ is slightly bigger than its values at $t=0$ and $t=\infty$.

In Fig. 4 we plot $f(x,t)$ corresponding to $z_1 = \frac{2}{5}, z_2 = \frac{3}{5}$ in Eq. (30) and we observe the same small effect in the neighborhood of the second zero of $f-1$. Let us notice that the existence of zeros of $f-1$ is a consequence of the conservation law for the moment $M_0(t) \equiv 1$.

C. Do there exist general structures suggested by the particular Krook–Wu solution?

If we do not consider too particular $F(x,t)$ soultions, then in general they will correspond to an infinite mixing of either pure solutions or of positive solutions. For the corresponding set $a_n(t)$, the least decreasing t behavior are provided with $\bar{a}_0\, e^{-t/3}$ and $\bar{a}_1\, e^{-t/2}$ and we know that the par-

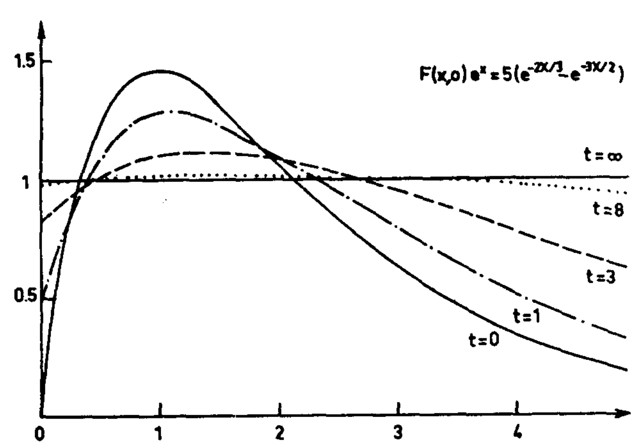

FIG. 4. Evolution of $f(x,t)$ for a difference of two exponentials at $t=0$.

ticular Krook–Wu solution contains both these smallest n terms. We recall that it is a particular $\bar{a}_0$, $\bar{a}_1$ mixing and on the one hand we have not found other finite mixing of pure solutions embodying positivity. On the other hand, if we consider the positive solutions basis, we recall that this particular solution corresponds to initial values $(a_n(0))$ given by Eq. (29) $(z_2 = 0, 0 < z_1 < 1)$ and in particular it contains, like many other acceptable solutions, nonnull coefficients of L_2 and L_3. However, there also exist families of acceptable solutions (possitivity, decreasing behavior...) such that either the coefficient of L_2 or L_3 or of both are identically zero for $t \geqslant 0$. In the last possibility, for instance, both terms $\exp(-t/3)$ and $\exp(-t/2)$ are not present and the least decreasing terms are provided by $\exp\{-[(P+1)/(P+3)]\}$ with $P \geqslant 1$. The same situation arises with the fundamental positive solutions. So it is clear that there exist positive, well-behaved, $F(x,t)$ solutions with characters different from the particular Krook–Wu solutions. [This remains true if we consider a larger sense where we retain only the features due to either $a_0(t) \not\equiv 0$ or $a_1(t) \not\equiv 0$.]

Now this discussion can be interpreted in terms of the moments $M_n(t)$: $a_0 \equiv 0$ is equivalent to $M_2(t) \equiv 1$; $a_1 \equiv 0$, $a_0 \neq 0$ correspond to $M_3(t) - 3M_2(t) + 2 \equiv 0$ and $a_0 \equiv a_1 \equiv 0$ to $M_2(t) \equiv M_3(t) \equiv 1$. It follows that if we exclude for the moments $M_n(t)$ the possibility to satisfy these particular constraints, then we conclude that the set $a_n(t)$ of coefficients of the Laguerre expansion of $e^{x}F(x,t)$ will always contain for the least t decreasing behavior either $\exp(-t/3)$ or $\exp(-t/2)$ and that these features are already present in the Krook–Wu particular solution.

7. CONCLUSION

In this paper and in the companion one[4], from the complementary point of view, we have established methods in order to build the solutions $F(x,t)$ of the Tjon–Wu model of the Boltzmann equation. These solutions can be expanded either in power or in Laguerre polynomials[6] of the energy variables so that the only nontrivial dependence is provided by their time dependent coefficients. From our results we extract some salient features.

The time behavior of the coefficients of the Laguerre polynomials is of the kind $\exp(-\text{const}\ t)$ where the constants can only take discrete values. This result can be established either directly from the Boltzmann integrodifferential equation[4] or from the structure of the nonlinear equation satisfied by the generating functional of the moments M_n. For these moments $M_n(t)$ (linear combination of the Laguerre coefficients expansion) this discretization of the time behavior subsists and this was previously found by Krook–Wu.[1] If the lowest nontrivial moments M_2 and M_3 do not satisfy particular constraints, then the lowest time decreasing behaviors of the Laguerre coefficients expansion are $\exp(-t/3)$ and $\exp(-t/2)$ and these dependence are also present (in a well-defined combination) in the particular solution of Krook and Wu.[1,3,6,7]

We give strong arguments and conditions in order that smooth distribution function solutions positive at $t = 0$ remain positive for $t > 0$. We have considered two different bases for our solutions corresponding to the possibility of characterizing the coefficients of the Laguerre expansions by arbitrary constants either at $t = 0$ or infinity. In both cases we give explicit methods in order to construct the solutions; however, the first choice seems more convenient for the positivity property because we control directly this property at $t = 0$.

For the convergence of the expansion in Laguerre polynomials, in the companion paper,[4] a Hilbert space is constructed so that the solution stays in this space at ulterior time if it is present at $t = 0$. Here, for the power series expansion, we give sufficient conditions at $t = 0$ such that the expansion is still valid at $t > 0$ and even when t goes to infinity. In this way we obtain sufficient initial conditions, at $t = 0$, such that at $t > 0$ (or even $t \to \infty$) the solutions are entire functions of the energy variable. Further, for the existence at $t > 0$ (or even $t \to \infty$) of $\Sigma\, a_n^2(t)$ which represents the sum of the square of the coefficients of the Laguerre polynomials, we find sufficient conditions at $t = 0$.

Finally, we note that we have found some examples where the distribution function $F(x,t)$ is, at finite time and for some energy range, slightly larger than the corresponding ones both at initial time and at equilibrium.

ACKNOWLEDGMENTS

We thank Professor T.T. Wu for introductory discussions on this problem. We thank Professors R. Balian and A. Morel for discussions and for their interest in the work. All the computer programs of the numerical analysis have been done by Claudine Verneyre. We thank her for her great help.

APPENDIX A

We recall[8] that previously were obtained lower and upper bound for the classical orthogonal polynomials which possess a dominant "forward peak." We extend these properties for the Laguerre orthogonal functions (and some generalization of it).

We consider $f(x)$ having a convergent Taylor expansion such that f and all its derivatives on some interval $[0, x_0]$ $(x_0 > 0$, x_0 can be infinite) have values bounded in modulus by the corresponding quantities at $x = 0$:

$$\left|\frac{\partial^q}{\partial x^q} f(x)\right| < \left|\left(\frac{\partial^q}{\partial x^q} f(x)\right)_{x=0}\right|, \quad q = 0,1,2,\cdots, \quad x \in [0, x_0]. \tag{A1}$$

We consider for $x > 0$ and $(\partial^p/\partial x^p) f(x)|_{x=0} \neq 0$:

$$\int_0^x dx_1 \int_0^{x_1} dx_2 \ldots \int_0^{x_p} dx_p \frac{\partial^p}{\partial (x_p)^p} f(x_p) = f(x) - \sum_{q=0}^{p-1} \frac{x^q}{q!}\left(\frac{\partial^q}{\partial x^q} f(x)\right)_{x=0}. \tag{A2}$$

Taking into account Eq. (A1) we get

$$f(x) - \sum_{q=0}^{p-1} \frac{x^q}{q!} \left.\frac{\partial^q f(x)}{\partial x^q}\right|_{x=0} \times \begin{cases} < \dfrac{x^p}{p!}\left|\left(\dfrac{\partial^p}{\partial x^p} f(x)\right)_{x=0}\right|, \\ > -\dfrac{x^p}{p!}\left|\left(\dfrac{\partial^p}{\partial x^p} f(x)\right)_{x=0}\right|, \end{cases} \tag{A3}$$

$(\partial^p/\partial x^p)f(x)$ being either positive or negative; it follows that always one of these inequalities corresponds to an upper or a lower bound of f obtained from the Taylor expansion cutoff after a finite number of terms.

We apply to the Laguerre orthogonal functions $e^{-x/2}L_n(x)$ or more generally to $l_{n,\alpha}(x) = e^{-x/2}L_n^{(\alpha)}(x)$, where $L_n^{(\alpha)}$ are the standard[9] generalized Laguerre polynomials and $\alpha = 0, 1, 2, \cdots$, positive or null integers. We recall

$$\frac{\partial^p}{\partial x^p} L_n^{(\alpha)}(x)_{x=0} = \frac{(-1)^p(n+\alpha)!}{(n-p)!(\alpha+p)!}$$

and the inequalities for $x \geqslant 0$:

$$\left| e^{-x/2} \frac{\partial^p}{\partial x^p} L_n^{(\alpha)}(x) \right| \leqslant \frac{(n+\alpha)!}{(n-p)!(\alpha+p)!} = \left| \left(\frac{\partial^p}{\partial x^p} L_n^{(\alpha)}(x) \right)_{x=0} \right|, \tag{A4}$$

which are easily obtained from $(d/dx)L_n^{(\alpha)} = -L_{n-1}^{(\alpha+1)}$ and $|e^{-x/2}L_n^{(\beta)}(x)| \leqslant (n+\beta)/n!\beta!$. At $x=0$, the even (odd) derivatives of $l_{n,\alpha}(x)$ are positive (negative):

$$\left.\frac{\partial^p}{\partial x^p} l_{n,\alpha}(x)\right|_{x=0} = (-1)^p \sum_{q=0}^{p} \left(\frac{1}{2}\right)^{p-q} C_p^q \frac{(n+\alpha)!}{(n-q)!(\alpha+q)!},$$

$$\frac{\partial^p}{\partial x^p} l_{n,\alpha}(x) = \sum_{q=0}^{p} \left(-\frac{1}{2}\right)^{p-q} C_p^q e^{-x/2} \frac{\partial^q}{\partial x^q} L_n^{(\alpha)}(x), \tag{A5}$$

$$\left| \frac{\partial^p}{\partial x^p} l_{n,\alpha}(x) \right| \leqslant \sum_q \left(\frac{1}{2}\right)^{p-q} C_p^q \frac{(n+\alpha)!}{(n-q)!(\alpha+q)!} = \left| \left(\frac{\partial^p}{\partial x^p} L_n^{(\alpha)}(x) \right)_{x=0} \right|.$$

with $l_{n,\alpha}(x)$ satisfying Eq. (A1), and the Taylor series verify the inequalities (A3); if we take an even (or odd) number of terms, we get lower (upper) bounds for $x > 0$.

APPENDIX B

B_1: *We get bounds on $N_q(t)$ from conditions at $t=0$ and with Eq.* (14b)

$$N_q(t) < e^{-t}[N_q(0) + qN_{q-1}(0)] + qN_{q-1}(t) + e^{-t}\int_0^t e^{t'}\Big(2N_{q-1}(t') + \sum_q (t')\Big)\, dt',$$

$$\sum_q (t) = \sum_{q=0}^{q-1} C_{q-1}^p N_p(t) N_{q-1-p}(t),$$

$$\sum_1 = N_0^2(t), \quad \sum_2 = 2N_0(t)N_1(t). \tag{B1}$$

Theorem: If we assume

$$N_0(t) \leqslant N_0(0),$$
$$N_q(0) \leqslant q!(4+N_0(0))^{q-1}N_0(0)[N_0(0)+2/q], \quad q \geqslant 1, \tag{B2}$$

then we have for any t

$$N_q(t) \leqslant q!(4+N_0(0))^q N_0(0). \tag{B3}$$

The proof is obtained by induction. From Eqs. (B1) and (B2) we get the following for $q=1$:

$N_1(t) < e^{-t}[3N_0(0) + N_0^2(0)] + N_0(0) + (1-e^{-t})\times(2N_0(0) + t\,N_0^2(0))$ and Eq. (B3) holds.

For $q=2$ we find $N_2(t) < e^{-t}2N_0(0)(N_0^2(0) + 6N_0(0) + 6) + 2N_0(0)(4+N_0(0)) + (1-e^t)\times(1+N_0(0))2N_0(0)(N_0(0)+4) < 2N_0(0)[N_0^2(0) + 7N_0(0) + 10] < 2N_0(0)(4+N_0(0))^2$ and Eqs. (B3) holds. Let us assume that Eq. (B3) holds for $q = 1,2,\ldots,q-1$; we want to show that it holds for $q \geqslant 2$. From $B(1-2-3)$ we get $N_q(t) < X_1 + \; + X_2 e^{-t}\; X_3(1-e^{-t})$, where $qN_{q-1}(t) < X_1 = q!\,N_0(0)(4+e\;N_0(0))^{q-1}$, $N_q(0) + qN_{q-1}(0) < X_2 = q!N_0(0)(4+N_0(0))^{q-2}(N_0^2(0) + N_0(0)(5+(2/q)) + (8/q) + 2(q-1))$, $2N_{q-1}(t) + \sum C_{q-1}^p N_p(t) N_{q-1-p}(t) < X_3 = q!(4+N_0(0)^{q-2}N_0(0)(N_0^2(0) + N_0(0)([4+(2/q)] + (8/q))$. We find $X_2 + X_3 < q!N_0(0)(4+N(0)^{q-2}(N_2^0(0) + N_0(0)([5+(2/q)] + (8/q) + (2/(q-1))$ and finally $N_q(t) < q!N_0(0)(4+N_0(0)^{q-2}(N_0^2(0) + N_0(0)([6+(2/q)] + (8/q) + 2/(q-1) + 4)$ less than the bound (B3) for $q \geqslant 2$.

B_2: *We get bounds on $N_q(t)$ at $t>0$ from conditions at $t=0$ in the different cases* I(i), (ii), (iii); II(i), (ii), (iii) *defined in Sec.* 4 D.

We start with Eq. (2.2′):

$$\sum u^{2+n}[(n+1)(\dot{a}_n + a_n) + 2\dot{a}_n] = \left[\sum u^{n+2}a_n\right]^2, \quad \dot{a}_n = \frac{da_n}{dt}, \tag{B4}$$

$$\dot{a}_n + a_n = (n+3)^{-1}\left[\sum_m a_m a_{n-m-2} + 2a_n\right]. \tag{B5}$$

1. Bounds on $N_q(t)$ in the II case

Case II(i): $|a_n| = a_n$. We derive $(q-1)$ times Eq. (B4) and take $u=1$:

$$\dot{N}_q + N_q + q\dot{N}_{q-1} + (q-2)N_{q-1} = \sum_{q=0}^{p-1} C_{q-1}^p N_q N_{q-1-p}. \tag{B6}$$

From Eq. (B5) we get $\dot{a}_n + a_n > 0$, $\dot{N}_q + N_q > 0$ and from Eq. (B6) a set of inequalities that we can integrate:

$$N_q(t) < e^{-t}\left[N_q(0) + \int_0^t e^{t'}(2N_{q-1}(t')) + \sum_{p=0}^{q-1} C_{q-1}^p N_p(t')N_{q-1-p}(t')dt'\right]. \tag{B7}$$

Case II(ii): $|a_n| = (-1)^n a_n$. We derive $(-q-1)$ times Eq. (B4), take $u=-1$, and get the relations (B6). From Eq. (B5) we get $(-1)^n\dot{a}_n + (-1)^n a_n > 0$ or $\dot{N}_q + N_q > 0$ and we have the same inequalities (B7).

Case II(iii): $a_n = c_k$, for $n = -2 + \lambda(k)$, $\lambda(k) = (k+1)(P+1)$ and $(-1)_{c_k}^{k+1} > 0$. Here we have

$$\sum_k u^{\lambda(k)}[(\lambda(k)-1)(\dot{c}_k + c_k) + 2\dot{c}_k] = \left[\sum u_{c_k}^{\lambda(k)}\right]^2, \tag{B4′}$$

$$\dot{c}_k + c_k = (n+3)^{-1}\left[\sum c_m c_{n-m-1} + 2c_k\right]. \tag{B5′}$$

Equation (B4′) for $u^{P+1} = \exp i\pi$ leads to Eq. (B6). We differentiate $(q-1)$ times Eq. (B4′), multiply by u^{q-1}, take $u^{P+1} = \exp i\pi$, and get the other relations of Eq. (B6). From Eq. (B5′) we get $(-1)^{k+1}(\dot{c}_k + c_k) > 0$ or $\dot{N}_q + N_q > 0$ and we get the inequalities (B7).

Theorem: In the II(i), (ii), (iii) cases if the following assumptions are satisfied:

$$N_0(t) < N_0(0),\ N_q(0) < q!N_0(0)(2 + N_0(0))^q, \tag{B8}$$

then we have for any t

$$N_q(t) < q!N_0(0)(2 + N_0(0))^q. \tag{B9}$$

The proof is obtained by induction. From the first inequality in Eq. (B7) we get

$$N_1(t) < N_1(0)e^{-t} + N_0(0)(2 + N_0(0))(1 - e^{-t}) < N_0(0)(2 + N_0(0))$$

and so Eq. (B9) is true for $q = 1$. Let us assume Eq. (B9) true for $q = 1, 2, \ldots, q-1$; we want to show that it holds for q. From Eqs. (B7) and (B8) we have

$$N_q(t) - N_q(0)e^{-t} < (1 - e^{-t})N_0(0)(2 + N_0(0))^{q-1}(q-1)! \times \left[2 + \sum_{p=0}^{q-1} N_0(0)\right] < (1 - e^{-t})q!(2 + N_0(0))^q N_0(0)$$

and the result (B9) follows if we take into account Eq. (B8) for $N_q(0)$.

2. Bounds on $N_0(t)$

In the II(i) case $a_n > 0$ we get from Eq. (B5)

$$3\dot{a}_n + a_n \leqslant \tfrac{3}{5} \sum a_m a_{n-m-2}, \tag{B10}$$

$$3\dot{N}_0 + N_0 < \tfrac{3}{5} N_0^2\ N_0(0)/N_0(t) > \tfrac{3}{5} N_0(0) + (1 - \tfrac{3}{5} N_0(0))\exp \tfrac{t}{3},$$

$$N_0(t) < \tfrac{5}{3} N_0(0)\left[N_0(0) + (\tfrac{5}{3} - N_0(0))e^{t/3}\right]^{-1} \leqslant N_0, \quad \text{if } N_0(0) \leqslant \tfrac{5}{3}. \tag{B11}$$

In the II(ii) case we put $b_n = (-1)^n a_n > 0$, get Eq. (B10) with the b_n instead of the a_n, and Eq. (B11) hold.

3. Bounds on $N_q(t)$ in the I case

Let us define $b_n = -a_n > 0$ in the I(i) case, $b_n = (-1)^{n+1}a_n > 0$ in the I(ii) case, and $e_k = (-1)^k c_k$ in the I(iii) case where $a_n = c_k$ for $n = \lambda(k) - 2$, $\lambda = (k+1) \times (P+1)$. We get $N_q = \Sigma\,(n+2)\cdots(n+3-q)b_n$ in the two first cases and $N_q = \Sigma_k\, e_k(\lambda(k))(\lambda(k) - 1) \cdots(\lambda(k) + 1 - q)$ in the third case. From Eqs. (B5) and (B5′) we get $3\dot{b}_n + b_n < 0$, $3\dot{e}_k + e_k < 0$ or $3\dot{N}_q + N_q < 0$ or $\dot{N}_q < 0$, and finally $N_q(t) < N_q(0)\exp(-t/3)$.

B_3: *Bounds on* $\widetilde{N}(t) = \Sigma_{-2}\, a_n^2(t)$. In Eq. (2″) we put $b_p = a_{n+2}$, $p \geqslant 0$:

$$\dot{b}_p + b_p = (p+1)^{-1} \sum_{m=0}^{p} b_m b_{p-m}. \tag{B12}$$

Firstly, we multiply Eq. (B12) by b_p, integrate from 0 to t, and sum over p:

$$\widetilde{N}(t)e^{2t} = \widetilde{N}(0) + \int_0^t e^{2t'} \sum_p b_p(p+1)^{-1} \sum b_m b_{p-m}\, dt'.$$

Using the Schwarz inequality

$$\left|\sum b_m b_{p-m}\right| \leqslant \widetilde{N}(t),\quad \left|\sum b_p(p+1)^{-1}\right| < (\widetilde{N}(t))^{1/2}\pi/\sqrt{6},$$

we get

$$\widetilde{N}(t)e^{2t} \leqslant \widetilde{N}(0) + 2\pi/\sqrt{6} \int_0^t e^{2t'}\widetilde{N}^{3/2}(t')dt' = \widetilde{M}(t) \tag{B13}$$

or

$$-\frac{d}{dt}(\widetilde{M}^{-1/2}) = \pi/\sqrt{6}e^{2t}\widetilde{N}^{3/2}\widetilde{M}^{-3/2} \leqslant \pi/\sqrt{6}e^{-t}$$

or

$$\widetilde{M}^{-1/2} \geqslant \left(\widetilde{N}(0)^{-1/2} - \frac{\pi}{\sqrt{6}}\right) + \frac{\pi}{\sqrt{6}} e^{-t},$$

and finally

$$\widetilde{N}(t) \leqslant \widetilde{N}(0)[(1 - \pi/\sqrt{6}\widetilde{N}^{1/2}(0))e^t + \pi/\sqrt{6}\widetilde{N}^{1/2}(0)]^{-2}, \quad \text{if } \widetilde{N}(0) \leqslant 6\pi^{-2}. \tag{B14}$$

Secondly, we integrate directly Eq. (B12), sum over p, define $N(t) = \Sigma_{-2}\,|a_n(t)|$, and get

$$N(t)e^t \leqslant N(0) + \sum_p \int_0^t e^{t'}(p+1)^{-1} \sum |b_m||b_{p-m}|dt' \leqslant M(t) = N(0) + \int_0^t e^{t'}N^2(t')dt'$$

or

$$-(d/dt)M^{-1} = M^{-2}e^t N^2 \leqslant e^{-t}$$

or

$$M(t)^{-1} \geqslant N(0)^{-1} - 1 + e^{-t}$$

and finally

$$N(t) \leqslant N(0)[(1 - N(0))e^t + N(0)]^{-1}, \quad \text{if } N(0) \leqslant 1. \tag{B15}$$

Noticing $\widetilde{N}(t) \leqslant N^2(t)$ we conclude that $\widetilde{N}(t) < \infty$ if either $\widetilde{N}(0) < 6\pi^{-2}$ or $N(0) \leqslant 1$. [The last bound cannot be improved because if $a_{n-2}(0) > 1$ we know that $a_{n-2}(t)$ diverges for some t value.]

APPENDIX C

1. Mixing of two pure solutions: $P_0 < P_1$, $P_0 P_1 \neq 2$ $\bar{a}_{P_0-1} \neq 0$, $\bar{a}_{P_1-1} \neq 0$

(i) The number of different time dependences cannot stay finite when $n \to \infty$. We recall the time dependence for pure solutions

$$a_{p-1+k(p+1)} = d_k \exp\left(-t\frac{p(k+1)}{p+2}\right) \tag{C1}$$

and the term a_q in the mixing with $q = P_0 - 1 + P_1(P_0 + 1) = P_1 - 1 + P_0(P_1 + 1)$ has two time dependences $\exp(-\alpha_i t)$, $\alpha_i = P_i/(P_i + 1)(P_i + 2)$, $i = 0, 1$. ($\alpha_0 \neq \alpha_1$ if

$P_0P_1 \neq 2$.) Always in the mixing the term a_n, $n = 2q+2$ [for which a_q^2 enter into the l.h.s. of Eq. (2)] has at least three time dependences $\exp(-2\alpha_0 t)$, $\exp(-2\alpha_1 t)$, $\exp[-(\alpha_0+\alpha_1)t]$ and the term a_n, $n = 4q+6$ has at least five time dependences $\exp(-4\alpha_0 t)$, $\exp(-4\alpha_1 t)$, $\exp[-2(\alpha_0+\alpha_1)t]$, $\exp[-(3\alpha_0+\alpha_1)t]$, $\exp[-(3\alpha_1+\alpha_0)t]$, and so on when $n\to\infty$.

(ii) The coefficients of the mixing can be recursively determined. We recall

$$a_n(t) = \sum_r d_n^{(r)} \exp\left[-t\left(\frac{p_0(n+2)}{(p_0+2)(p_0+1)} + \theta t\right)\right],$$

$$\theta = \frac{(p_0-p_1)(p_0p_1-2)}{(p_0+1)(p_0+2)(p_1+2)} \leqslant 0, \tag{C2}$$

$$\left[-(n+3)\left(\frac{p_0(n+2)}{(p_0+2)(p_0+1)} + \theta t\right) + (n+1)\right] d_n^{(t)} = \sum_{s,q} d_n^{(s)} d_{n-q-2}^{(t-s)}. \tag{C3}$$

$r = 0: \bar{a}_{P-1} = d^0_{P_0-1}$ and d^0_n through Eq. (C3); $t = 1$: $\bar{a}_{P_1-1} = d^{(1)}_{P_1-1}$ and $d^{(1)}_n$ through Eq. (C3), $s = 0,1$.

We get $n < P_1 - 1: d_n^{(t)} = 0$ if $t \neq 0$, $d_n^{(0)} \neq 0$ if $n = P_0 - 1 + k(P_0+1)$, $k = 0, 1,\cdots$; the first $d_n^{(1)} \neq 0$ for $n = P_1 - 1$, $n < 2P_1$, $d_n^{(t)} = 0$ if $t \neq 0$ and $t \neq 1$, and the first $d_n^{(2)} \neq 0$ for $n = 2P_1$. $n < P_1 - 1 + (t-1)(P_1+1): d_n^{(t')} = 0$ if $t' \neq 0, 1,\ldots,t-1$, the first $d_n^{(t)} \neq 0$ for $n = P_1 - 1 + (t-1)(P_1+1)\cdots$, and so on. So $\{d_n^{(0)}\}$ are the coefficients of the pure P_0/P_0+2, $\theta < 0$, the least decreasing terms ($d_n^{(t)} \neq 0$ $t \neq 0$) are either pure P_1 solutions terms or mixed terms.

2. Mixing of q arbitrary solutions: $P_0 < P_1 < \cdots < P_{q-1}$, $\bar{a}_{P_0-1} \neq 0, \bar{a}_{P_1-1} \neq 0 \ldots \bar{a}_{P_{q-1}-1} \neq 0$

We have

$$a_n(t) = \sum_{t_i, t = 1,\ldots,q-1} d_n^{(t_1,t_2,\ldots,t_{q-1})} \times \exp\left\{-t\left[\frac{P_0}{P_0+2}\left(\frac{n+2}{P_0+1}\right) + \sum_{j=1}^{q-1} \theta_j t_j\right]\right\}. \tag{C4}$$

The $d_n^{(t_1,\ldots,t_{q-1})}$ can be recursively determined

$$\left\{-(n+3)\left[\frac{P_0}{P_0+2}\left(\frac{n+2}{P_0+1}\right) + \sum_{j=1}^{q-1}\theta_j t_j\right] + n+1\right\} d_n^{(t_1,\ldots,t_{q-1})} = \sum_{q,s_1,\ldots,s_{q-1}} d_q^{(s_1,s_2,\ldots,s_{q-1})} d_{n-q-2}^{(t_1-s_1,t_2-s_2,\ldots,t_{q-1}-s_{q-1})}. \tag{C5}$$

The arbitrary constants are $d^{(0,0,\ldots,0)}_{P_0-1} = \bar{a}_{P_0-1}$, $d^{(1,0,0,\ldots,0)}_{P_1-1} = \bar{a}_{P_1-1},\ldots,d^{(0,0,\ldots,1)}_{P_{q-1}-1} = \bar{a}_{P_{q-1}-1}$; for these values the l.h.s. of Eq. (C5) is zero and leads to $(P_i+2) \times (P_0+2)(P_0+1)\theta_i = 0$ if $P_0P_i = 2$.

3. Positivity property (mixing $\frac{1}{2},\frac{1}{3}$)

We have

$$\mathscr{F}(\omega, t = \infty) = 1 + \sum d_n \omega^{n+2}((n+2)!)^{-1}, \quad \omega > 0, \tag{C6}$$

with d_n given in Eq. (20a). We put $d_n = (-1)^{n+1}\delta_n(n+1)$, $\mathscr{F} = 1 + \Sigma(1-)^{n+1}\omega^{n+2}(n+1)\times\delta_n((n+2)!)^{-1}$ such that for $\delta_0 = \delta_1 = 1$; then $\delta_n = 1$ (Krook-Wu solution) and $\mathscr{F} = (1+\omega)e^{-\omega} > 0$. We get in the δ_0, δ_1 plane: (i) $\delta_0 > 0$, $\delta_1 > 0$; if $\delta_1 = \lambda\delta_0^{3/2}$, $\lambda > 0$, then $\delta_n = (\text{const.})_n(\delta_0^{1/2})^{n+2}$ and the scaling $\omega \to (\omega\delta_0^{1/2})$ does not modify the positivity; if $\lambda = 1$ we have the positivity of the Krook-Wu solution whereas for $\lambda \neq 1$, from numerical analysis, we have not found any $\mathscr{F}$ without zeros. (ii) $\delta_0 > 0$, $\delta_1 < 0$; then $\delta_{2P} > 0$, $\delta_{2P+1} < 0$, and $\mathscr{F}$ strictly decreasing has one zero. (iii) $\delta_0 < 0$; from numerical analysis we have not found any $\mathscr{F}$ without zeros.

4. Positivity property for a mixing of P_0, P_1, ($P_0P_1 \neq 2$) pure solutions, $P_0 < P_1$

From Eq. (C2), $\theta < 0$, and for n fixed the least decreasing term in a_n corresponds to the largest t value. We rescale the series $e^x F(x,t) = \mathscr{F}(\omega,t)$ with a new variable $\omega = x\exp(-ct)$ such that $\lim_{t\to\infty}\mathscr{F}(\omega,t)$ is finite and nontrivial. For each power $\gamma_P\omega^P$ we want $\gamma_P(t) \to$ const when $t\to\infty$ ($\gamma_P\to\infty$ is forbidden). For $a_n(t)L_{n+2}(x)$ we want that it goes to $(\text{const})_n\omega^{n+2}$ and a scaling compatible with the other nth terms. We get three cases following the origin of the least decreasing term in $a_n(t)$: (i) It is a pure P_1 solution term (coefficient depending only on $\bar{a}_{P_1-1}$); then $\omega_{P_1} = x\exp\{-t[P_1/(P_1+1)(P_1+2)]\}$. (ii) It is a pure P_0 solution term and $\omega_{P_0} = x\exp\{-t[P_0/(P_0+1)(P_0+2)]\}$.

However, $\omega_{P_0} = \omega_{P_1}\exp(-\gamma t) \underset{t\to\infty}{\longrightarrow} 0$ at fixed ω_{P_1} because $\gamma > 0$:

$$\gamma(P_0+1)(P_1+1)(P_0+2)(P_1+2) = (P_1-P_0)(P_1P_0-2) > 0.$$

(iii) It is a mixing of P_0 and P_1. Rescaling the corresponding a_nL_{n+2} with a new variable ω, we get $\omega = \omega_{P_1}$ multiplied by a factor $\to 0$ when $t\to\infty$. Finally, $e^xF = \mathscr{F}(\omega_{P_1},t)$ and when $t\to\infty$, $\mathscr{F}\to 1 + \sum d_{k_1}\omega_{P_1}^{(k+1)(P_1+1)}[(k+1)(P_1+1)]!$, which is the dominant term of the pure P_1 solution.

If now we consider a mixing of a finite number of pure solutions $P_1, P_2,\ldots,P_q$ and rescale the variable $x\to\omega$, then we are faced with a dominant series, coming from the largest pure solution P_q.

5. Infinite mixing of pure or positive solutions

Let us assume that the $a_m(t)$, $m = 0, 1,\ldots,n-2$ are of the type

$$a_m(t) = \sum_t d_m^{(t)} e^{-b_m^{(t)} t}, \tag{C9}$$

where the sum over t is finite and the $d_m^{(t)}, b_m^{(t)}$ are known. Substituting into Eq. (12) we find if the introduced arbitrary nth constant is $a_n(0)$,

$$a_n(t) = \left(a_n(0) + \sum_{m,t_1,t_2} \frac{d_m^{t_1} d_{n-2-m}^{t_2}}{[(b_m^{t_1} + b_{n-2-m}^{t_2})(n+3) - (n+1)]}\right) \times \exp\left[-\left(\frac{n+1}{n+3}\right)t\right] - \sum_{m,t_1,t_2} \frac{d_m^{t_1} d_{n-m-2}^{t_2} \exp[(-b_m^{t_1} + b_{n-m-2}^{t_2})t]}{[(b_m^{t_1} + b_{n-m-2}^{t_2})(n+3) - (n+1)]}. \tag{C10}$$

Because m, t_1, t_2 run over a finite number of values, then Eq. (C10) can be rewritten like Eq. (C9) and the constants entering into a_n can be determined from those coming from a_m, $m = 0,\ldots,n-2$.

If the introduced nth constant is $\bar{a}_n$, then $a_n(t)$ is still given by Eq. (C10) with

$$\bar{a}_n = a_n(0) + \sum_{m,t_1,t_2} \frac{d_m^{t_1} d_{n-2-m}^{t_2}}{[(b_m^{t_1} + b_{n-m-2}^{t_2})(n+3) - (n+1)]}. \tag{C11}$$

On the other hand, for the first m values for which we have not both $\bar{a}_m = a_m(0) = 0$ and where the nonlinear part of Eq. (12) is present we verify easily the representation (C9).

[1]Max Krook and Tai Tsun Wu, Phys. Rev. Lett. **36**, 1107 (1976); Phys. Fluids **20**, 1589 (1977).

[2]John Tjon and Tai Tsun Wu, Phys. Rev. A **19**, 883 (1979).

[3]A.V. Bobylev, Sov. Phys. Dokl. **20**, 320, 322 (1976).

[4]M. Barnsley and H. Cornille, to appear in Proc. R. Soc. London.

[5]G. Szego, *Orthogonal Polynomials* (American Mathematical Society, Providence, 1939).

[6]See also, M.H. Ernst, Phys. Lett. A. **63**, 390 (1979). The authors wish to thank M.H. Ernst for bringing his work to our notice.

[7]G. Tenti and W.H. Hui, J. Math. Phys. **19**, 774 (1978).

[8]H. Cornille, Phys. Rev. **14**, 1693 (1976).

[9]*Higher Transcendental Functions* (Bateman Manuscript Project), edited by A. Erdélyi (McGraw-Hill, New York, 1953), Vol. 2, Chap. X.

Hierarchical equations of evolution of an anharmonic system[a]

R. Esposito

Istituto Matematico, Università di Napoli, Napoli, Italy

M. Pulvirenti

Istituto Matematico, Università di Roma, Rome, Italy

(Received 1 May 1979; accepted for publication 24 May 1979)

We investigate the evolution of the states of a system of infinitely many anharmonic oscillators via a hierarchy of equations similar to the BBGKY one. We prove an existence theorem for the solutions in the L_1 sense for a large class of initial data.

I. INTRODUCTION

The BBGKY is a tool to study the time evolution of the states of a system of infinitely many classical particles.[1] It is an infinite chain of evolution equations connecting the time evolved n-bodies correlation functions with higher order time evolved ones.

One would expect that high order truncations of this chain do simulate the "true" behavior of the system. It means, from a mathematical point of view, that the solution of the infinite hierarchy actually exists and moreover it may be approximated by solutions of finite number of bodies problems. Unfortunately there are very few results in this direction, each of them giving the existence of weak solutions for the hierarchy associated to a one and two dimensional systems in situations near equilibrium.[2] More recent announced results can be found in Ref. 3. Stationary solutions have also been studied in Ref. 4.

In this paper we study a hierarchy of equations similar to the BBGKY equations, describing the time evolution of a lattice of interacting anharmonic oscillators and we give an existence theorem for the solutions of these equations. The initial data are chosen in quite a large class. This theorem is obtained by means of a limit procedure on the solutions of the truncated equations, by explicit estimates, avoiding nonconstructive arguments. More precisely we check the convergence of the time evolved restricted probability distributions in the L_1 sense. This result of course implies the existence of weak solutions (i.e., solutions for the time evolved expectation values of local observables) that can be obtained in a much more direct way.

The plan of the paper is the following: In Sec. 3 we give the main results. Section 2 is devoted to definitions and preliminarys. Finally in Sec. 4 are the proofs.

2. NOTATIONS, DEFINITIONS AND PRELIMINARY RESULTS

We consider a system of unbounded oscillators on a ν-dimensional cubic lattice $\mathbb{Z}^\nu$. The single oscillator phase space is assumed to be $\mathbb{R}^1\times\mathbb{R}^1$. The *phase space* of the system is $\mathscr{X}=\{(q_i,p_i)\,|\,i\in\mathbb{Z}^\nu, p_i,q_i\in\mathbb{R}^1\}$. For every $\Lambda\subset\mathbb{Z}^\nu$, $\mathscr{X}_\Lambda=\{(q_i,p_i)\,|\,i\in\Lambda,\ p_i,q_i\in\mathbb{R}^1\}$ is the phase space associated to the region Λ. $x,y,z,\cdots$ and $x_\Lambda, y_\Lambda, z_\Lambda,\cdots$ denote respectively points of $\mathscr{X}$ and of $\mathscr{X}_\Lambda$. $\mathscr{X}$ and the $\mathscr{X}_\Lambda$'s are equipped with the usual product topology. If $\Lambda'\supset\Lambda$ the map $\mathscr{X}_{\Lambda'}\ni x_{\Lambda'}\to(x_{\Lambda'})_\Lambda\in x_\Lambda$ is defined by ignoring all the coordinates of $x_{\Lambda'}$ outside of Λ.

A *state* P of the system is a family of positive Borel functions $P_\Lambda:\mathscr{X}_\Lambda\to\mathbb{R}$, $\Lambda\subset\mathbb{Z}^\nu$, Λ finite, with the following properties:

$$\text{(i)}\ \int P_\Lambda(x_\Lambda)dx_\Lambda=1\,;$$

$$\text{(ii)}\ \int P_\Lambda(x_\Lambda)dx_{\Lambda\setminus\Omega}=P_\Omega((x_\Lambda)_\Omega),\qquad \Lambda\supset\Omega.$$

where dx_Λ denotes the Lebesgue measure on $\mathscr{X}_\Lambda$. As is well known, such a family of functions define uniquely a Borel probability measure, still denoted by P, on $\mathscr{X}$.

The time evolution of the system is described by the following family of Hamiltonians:

$$H_\Lambda(x_\Lambda)=\sum_{i\in\Lambda}\Big(p_i^2/2m+\lambda q_i^4+Kq_i^2-J\sum_{j\in U_i\cap\Lambda}q_jq_i\Big),\tag{2.1}$$

where $\lambda, J>0, k\in\mathbb{R}$, $m>0$ is the mass of a single oscillator.

$$U_i=\{\,j\in\mathbb{Z}^\nu\,|\,|i-j|=1\},\quad |i-j|=\sup_{1\leqslant\alpha\leqslant\nu}|i_\alpha-j_\alpha|.$$

Let $\varphi:\mathbb{R}\to[1,+\infty)$ the following function:

$$\varphi(k)=\max(\log^2k,1),$$

and

$$\mathscr{L}(x_\Lambda)\doteq\sup_{i\in\Lambda}\frac{\mathscr{L}_i(x_\Lambda)}{\varphi(|i|)},\quad \Lambda\subseteq\mathbb{Z}^\nu,\tag{2.2}$$

where

$$\mathscr{L}_i(x_\Lambda)=p_i^2/2m+\lambda q_i^2+kq_i^4+1.\tag{2.3}$$

We put

$$\mathscr{X}_0=\{x\in\mathscr{X}\,|\,\mathscr{L}(x)<+\infty\}.\tag{2.4}$$

For every $\Lambda\subset\mathbb{Z}^\nu$, Λ bounded, $t\in\mathbb{R}$ we denote $S_t^\Lambda x_\Lambda$ the solution of the motion equations with initial point $x_\Lambda\in\mathscr{X}_\Lambda$, governed by the Hamiltonian H_Λ. In the sequel we shall denote

$$S_t^\Lambda x_\Gamma=S_t^\Lambda(x_\Gamma)_\Lambda\cup x_{\Gamma\setminus\Lambda},\ \Gamma\supset\Lambda\,,$$

and $S_t^n=S_t^{\Lambda_n}$ where $\Lambda_n=[-n,n]^\nu$. We denote

$$\Lambda_x=\sup_{n\leqslant x}\Lambda_n,\forall x\in\mathbb{R}.$$

[a]Partially supported by G.N.F.M. of the Italian C.N.R.

All the dynamical properties we need in this paper are summarized in the following proposition:

Proposition 2.1:

(i) *For all* $x\in\mathscr{X}_0$ *the following inequality holds*:

$$\mathscr{L}(S_t^n x)\leqslant e^{a|t|}\,\mathscr{L}(x), \tag{2.5}$$

for some $a\in\mathbb{R}^+$. The sequence $S_t^n x$ *is a Cauchy sequence in* $\mathscr{X}$ *and defines* a *one parameter group of transformations in* $\mathscr{X}_0$:

$$S_t x = \lim_{n\to\infty} S_t^n x, \quad x\in\mathscr{X}_0. \tag{2.6}$$

Moreover there exist positive continuous functions $c_i(t)\, i=1,2,\cdots$ *such that*:

(ii) *If* $m>n>l>k\geqslant 0$ *then*

$$\sup_{i\in\Lambda_l} \max\Big[|q_i(S_t^{\Lambda_m\setminus\Lambda_k}x)-q_i(S_t^{\Lambda_n\setminus\Lambda_k}x)|,\; |p_i(S_t^{\Lambda_m\setminus\Lambda_k}x)-p_i(S_t^{\Lambda_n\setminus\Lambda_k}x)|\Big] \leqslant \frac{[c_1(t)\varphi(n)\mathscr{L}(x)]^{n-l+1}}{(n-l)!}, \tag{2.7}$$

$$\sup_{i\in\Lambda_n\setminus\Lambda_l} \max\Big[|q_i(S_t^n x)-q_i(S_t^{\Lambda_n\setminus\Lambda_k}x)|,\; |p_i(S_t^n x)-p_i(S_t^{\Lambda_n\setminus\Lambda_k}x)|\Big] \leqslant \frac{[c_1(t)\varphi(n)\mathscr{L}(x)]^{l-k+1}}{(l-k)!} \tag{2.8}$$

where $q_i(x)$, $p_i(x)$ *are the coordinate and momentum of the oscillator at the site i*.

(iii) Let $\Lambda\subset\mathbb{Z}^\nu$, $x\in\mathscr{X}_0$ *and*

$$D_{j,i}^{\Lambda}(x,t) = \max\left(\left|\frac{\partial q_j(S_t^\Lambda x)}{\partial q_i}\right|, \left|\frac{\partial q_j(S_t^\Lambda x)}{\partial p_i}\right|, \left|\frac{\partial p_j(S_t^\Lambda x)}{\partial q_i}\right|, \left|\frac{\partial p_j(S_t^\Lambda x)}{\partial p_i}\right|\right). \tag{2.9}$$

Then:

$$D_{j,i}^{\Lambda}(x,t)\leqslant \sum_{m\geqslant|i-j|} \frac{(c_2(t))^m}{m!}[\mathscr{L}(x)\varphi(i+m)]^{m/2}. \tag{2.10}$$

(iv) *If* $x\in\mathscr{X}_0$ *and* $m>n>k$,

$$\sup_{i,j\in\Lambda_k}\left|\frac{\partial y_j(S_t^n x)}{\partial x_i}-\frac{\partial y_j(S_t^m x)}{\partial x_i}\right| \leqslant \frac{[c_3(t)\varphi(n)\mathscr{L}(x)]^{n-k}\exp(c_4(t)\mathscr{L}(x)^\delta)}{(n-k)!}, \tag{2.11}$$

where y_j *and* x_i *are respectively* p_j *or* q_j *and* p_i *or* q_i, *and* $\delta<1$.

The proof of Proposition 2.1 is a consequence of the basic estimate (2.5) obtained in Ref. 5. The estimates (2.7) and (2.8) are straightforward and sketched in Ref. 6. The estimates (2.10) and (2.11) are consequences of the inequality (2.5) and of an integral inequality coming from the definition of the derivatives. The proof of such estimates is essentially contained in Ref. 7.

3. FORMULATIONS OF THE PROBLEM AND RESULTS

As in the case of particles system the time evolved probability distributions defining a state P will satisfy a hierarchy of equations. In fact, defining: ($\Lambda,\Omega\subset\mathbb{Z}^\nu$; Λ,Ω bounded)

$$P_\Lambda^\Omega(t,x_\Lambda) = P_\Lambda(S_t^\Omega x_\Lambda), \text{ if } \Lambda\supseteq\Omega,$$

$$P_\Lambda^\Omega(t,x_\Lambda) = \int P_\Omega(S_t^\Omega x_\Omega)dx_{\Omega\setminus\Lambda}, \quad \text{if } \Lambda\subset\Omega \tag{3.1}$$

and putting $P_j^n(t,x_{\Lambda_j}) = P_{\Lambda_j}^{\Lambda_n}(t,x_{\Lambda_j})$, one obtains, at least formally from the Liouville equation

$$\frac{\partial}{\partial t}P_j^n(t,x_{\Lambda_j}) = (L_j P_j^n)(t,x_{\Lambda_j}) + (C_{j,j+1}P_{j+1}^n)(t,x_{\Lambda_j}), \quad n>j, \tag{3.2}$$

where

$$(L_j P_j^n)(t,x_{\Lambda_j}) = \{P_j^n, H_{\Lambda_j}\}(t,x_{\Lambda_j}),$$

$$(C_{j,j+1}P_{j+1}^n)(t,x_{\Lambda_j}) = \int dx_{\Lambda_{j+1}\setminus\Lambda_j}\{P_{j+1}^n(t,x_{\Lambda_{j+1}}),\; U(x_{\Lambda_j}|x_{\Lambda_{j+1}\setminus\Lambda_j})\}, \tag{3.3}$$

and $\{\cdot,\cdot\}$ is the usual Poisson bracket and

$$U(x_{\Lambda_j}|x_{\Lambda_{j+1}\setminus\Lambda_j}) = J \sum_{\substack{i,k:\\ |i-k|=1\\ i\in\Lambda_j, k\in\Lambda_{j+1}\setminus\Lambda_j}} q_i q_k. \tag{3.4}$$

The Proposition 2.1 suggests that for all j the sequence $\{P_j^n(t,\cdot)\}$ converges as $n\to\infty$ in some sense to a limit $\{P_j(t,\cdot)\}$ which satisfies an infinite hierarchy formally similar to (3.2). We underline that Eq. (3.2) are nothing else than the Liouville equation if $n<+\infty$. If $n=+\infty$ then Eq. (3.2) becomes a way of describing the time evolution of an infinite system in terms of its statistical properties. In this paper we develop the above point of view by showing the convergence of the sequence $\{P_j^n(t,\cdot)\}$ in the $L_1(\mathscr{X}_{\Lambda_j})$ sense and that the limit $\{P_j(t,\cdot)\}$ satisfies (in the L_1 sense) Eq. (3.2) under very general assumptions on the initial state. More precisely we introduce that class $\mathscr{E}$ of states defined by the following properties:

(i) $P(x_\Lambda)>0$ for all $x_\Lambda\in\mathscr{X}_\Lambda$ and $\Lambda\subset\mathbb{Z}^\nu$; moreover there exist $R\in\mathbb{N}$ such that if $\bar\Lambda\equiv\{i\in\Lambda \mid |i-j|\leqslant R\}$ and $\Omega\supset\bar\Lambda$ then

$$P(x_\Lambda|x_{\Omega\setminus\Lambda}) = P(x_\Lambda|(x_\Omega)_{\partial\Lambda}), \tag{3.5}$$

where $\partial\Lambda=\bar\Lambda\setminus\Lambda$,

$$P(x_\Lambda|\bar x_{\Lambda'}) = P_{\Lambda\cup\Lambda'}(x_\Lambda\cup\bar x_{\Lambda'})/P_{\Lambda'}(\bar x_{\Lambda'}),$$

and $\Lambda'\cap\Lambda=\phi$.

(ii) There exist a $b>0$ such that

$$\int(\exp[b\mathscr{L}(x_\Lambda)])\,P_\Lambda(x_\Lambda)dx_\Lambda\leqslant\mathscr{C}, \tag{3.6}$$

where $\mathscr{C}>0$ and is independent of Λ.

(iii) The P_Λ's are assumed to be at least twice differentiable and to satisfy the following inequalities:

$$\left|\frac{\partial P_\Lambda(x_\Lambda)}{\partial x_i}\right|\leqslant P_\Lambda(x_\Lambda)\Big[\eta\mathscr{L}(x)\sup_{j\in\Lambda}\varphi(|j|)\Big]^\xi, \tag{3.7}$$

$$\left|\frac{\partial^2 P_\Lambda(x_\Lambda)}{\partial x_i \partial x_j}\right| \leqslant P_\Lambda(x_\Lambda)\Big[\bar{\eta}\mathscr{L}(x) \sup_{k\in\Lambda} \varphi(|k|)\Big]^{\bar{\xi}}, \tag{3.8}$$

for some $\xi, \bar{\xi}, \eta, \bar{\eta} \in \mathbb{R}^+$.

Remark 1: If $P\in\mathscr{C}$ then $P(\mathscr{X}_0)=1$ because inequality (3.6) implies that:

$$P(\{x|\mathscr{L}(x)\geqslant\alpha\})\leqslant Ae^{-b\alpha}, \tag{3.9}$$

for some $A, b\in\mathbb{R}^+$.

Remark 2: From (3.7) and (3.8) analogous estimates for conditional probabilities follow.

Remark 3: We notice that the class $\mathscr{C}$ contains all the Gibbs states generated by a family of smooth enough Hamiltonians $\{F_\Lambda\}$ that are short range (to satisfy *i*), superstable, and lower regular in order to satisfy the Ruelle's superstable estimate (see Ref. 8) that implies (ii) and finally such that their first and second derivatives are bounded by powers of $\mathscr{L}$.

The main result that we obtain is the following:

Theorem 3.1: *Let* $P\in\mathscr{C}$; *then for all* $j\in\mathbb{N}$:

$$\lim_{\substack{n\to\infty\\ m>n}} ||P_j^n(t,\cdot) - P_j^m(t,\cdot)||_1 = 0, \tag{3.10}$$

$$\lim_{\substack{n\to\infty\\ m>n}} ||C_{j,j+1}(P_{j+1}^n(t,\cdot) - P_{j+1}^m(t,\cdot))||_1 = 0, \tag{3.11}$$

$$\lim_{\substack{n\to\infty\\ m>n}} ||\{H_{ij}, P_j^n(t,\cdot) - P_j^m(t,\cdot)\}||_1 = 0, \tag{3.12}$$

uniformly in t on compact sets.

Theorem 3.1 will be proven in Sec. 4. We note also that the techniqes employed for the proof are based on explicit estimates that allow one, at least in principle, to compute the velocity of the above convergence. This is interesting from a physical point of view because one needs to evaluate the error arising in truncating the hierarchy to some finite order. As a consequence of Theorm 3.1 we have

Thereom 4.2: *For all* $j\in\mathbb{N}$, *let* $P_j(t,\cdot)$ *defined by*

$$L_1 - \lim_{n\to\infty} P_j^n(t,\cdot) = P_j(t,\cdot). \tag{3.13}$$

Then $P_j(t,\cdot)$ *is* L_1 *differentiable and, denoting by* $\bar{L}_j$ *and* $\bar{C}_{j,j+1}$ the closure of the operators L_j and $C_{j,j+1}$ (*such closures do exist*), *then*

$$\frac{d}{dt}P_j(t,\cdot) = \bar{L}_j p_j(t,\cdot) + \bar{C}_{j,j+1}P_{j+1}(t,\cdot) \tag{3.14}$$

Proof: It is easy to see that the operators L_j and $C_{j,j+1}$ defined respectively on $C_0^\infty(x_{\Lambda_j})$ and $C_0^\infty(x_{\Lambda_{j+1}})$ are closable and hence the rhs of (3.14) makes sense. Furthermore it cannot fail to coincide with the strong derivative $d/dt\,(P_j(t,\cdot))$ that it is shown to exist by standard arguments and by Theorems 3.1.

4. PROOFS

We need some definitions. Let $\chi_\alpha\in C_0^\infty(\mathbb{R})$, with the following properties:

(i) $\chi_\alpha(r)=1$, if $r\in[-\alpha+1, \alpha-1]$;

(ii) $\chi_\alpha(r)=0$, if $r\notin[-\alpha,\alpha]$; (4.1)

(iii) $|\chi'_\alpha(r)|<M$,

for some $M\in\mathbb{R}$ not depending on α.

We put $\chi_\alpha(x_{\Lambda_m})\equiv\chi_\alpha(\mathscr{L}(x_{\Lambda_m}))$ and $\chi_\alpha(x)=\chi_\alpha(\mathscr{L}(x))$. We finally define:

$$S_{n,m}(t) = \int dx_\Omega \Big|\int dx_{\Lambda_m\setminus\Omega}\chi_\alpha(x_{\Lambda_m}) \times (P_m^n(t,x_{\Lambda_m}) - P_m^m(t,x_{\Lambda_m}))\Big| \tag{4.2}$$

and

$$T_{n,m}(t) = \sup_{i\in\Omega}\int dx_\Omega \Big|\int dx_{\Lambda_m\setminus\Omega}\chi_\alpha(x_{\Lambda_m}) \times\mu_i(x_{\Lambda_m})\left[\frac{\partial}{\partial x_i}P_m^n(t,x_{\Lambda_m}) - \frac{\partial}{\partial x_i}P_m^m(t,x_{\Lambda_m})\right]\Big|, \tag{4.3}$$

where x_i is p_i or q_i, $i\in\Omega$ and $\mu_i(x_{\Lambda_m}) = -J\Sigma_{j\in U_i}q_j + \lambda q_i^3$ or $\mu_i(x_{\Lambda_m})=p_i$.

The strategy of the proof of Theorem 3.1 is based in two steps: First we prove Proposition 4.1 below. Then, with the aid of the property (3.6) we remove the cutoff χ_α and prove Theorem 3.1

Proposition 4.1:

(i) $\lim_{\substack{n\to\infty\\ m>n}} S_{n,m}(t)=0$;

(ii) $\lim_{\substack{n\to\infty\\ m>n}} T_{n,m}(t)=0$; (4.4)

and both the limits hold uniformly in t on compact sets.

Proof: The proof of Proposition 4.1 consist in estimating various terms by using Proposition 2.1. We shall estimate only typical terms and we briefly sketch how to obtain the other bounds. We start by proving (4.4) (ii):

$$T_{n,m}(t)\leqslant T_{n,m}^1(t) + T_{n,m}^2(t),$$

where

$$T_{n,m}^1 = \int dx_\Omega\Big|\int dx_{\Lambda_m\setminus\Omega}\chi_\alpha(x_{\Lambda_m})\frac{\partial h_m^n(t,x_{\Lambda_m})}{\partial x_i}\mu_i(x_{\Lambda_m}) \times(P_m^n(t,x_{\Lambda_m}) - P_m^m(t,x_{\Lambda_m}))\Big| \tag{4.5}$$

$$T_{n,m}^2(t) = \int dx_\Omega\Big|\int dx_{\Lambda_m\setminus\Omega}\,\chi_\alpha(x_{\Lambda_m})P_m^m(t,x_{\Lambda_m}) \times\mu_i(x_{\Lambda_m})\Big(\frac{\partial h_m^n(t,x_{\Lambda_m})}{\partial x_i} - \frac{\partial h_m^m(t,x_{\Lambda_m})}{\partial x_i}\Big)\Big|. \tag{4.6}$$

Here h with some indices and of some argument denotes the logarithm of P with the same indices and of the same arguments. In the sequel the symbol $\partial h/\partial x_i|y_\Lambda$ will denote the derivative of $\log P_\Lambda$ with respect to x_i at the point y_Λ. In order to estimate (4.6) we put

$$Z_{n,m} = \left|\frac{\partial h_m^n(t,x_{\Lambda_m})}{\partial x_i} - \frac{\partial h_m^m(t,x_{\Lambda_m})}{\partial x_i}\right|. \tag{4.7}$$

Then:

$$Z_{n,m} \leqslant Z_{n,m}(p) + Z_{n,m}(q), \tag{4.8}$$

where

$$Z_{n,m}(p) = \Big| \sum_{j \in \Lambda_m} \Big(\frac{\partial h_m}{\partial p_j}\Big|_{S_t^n x_{\Lambda_m}} \frac{\partial p_j(S_t^n x_{\Lambda_m})}{\partial x_i} - \frac{\partial h_m}{\partial p_j}\Big|_{S_t^m x_{\Lambda_m}} \frac{\partial p_j(S_t^m x_{\Lambda_m})}{\partial x_i} \Big) \Big|, \tag{4.9}$$

and $Z_{n,m}(q)$ is analogously defined by replacing p_j by q_j. Finally

$$Z_{n,m}(p) \leqslant Z^1_{n,m}(p) + Z^2_{n,m}(p), \tag{4.10}$$

where

$$Z^1_{n,m}(p) = \Big| \sum_{j \in \Lambda_m} \frac{\partial h_m}{\partial p_j}\Big|_{S_t^m x_{\Lambda_m}} \times \Big(\frac{\partial p_j(S_t^n x_{\Lambda_m})}{\partial x_i} - \frac{\partial p_j(S_t^m x_{\Lambda_m})}{\partial x_i} \Big) \Big|, \tag{4.11}$$

$$Z^2_{n,m}(p) = \Big| \sum_{j \in \Lambda_m} \Big(\frac{\partial h_m}{\partial p_j}\Big|_{S_t^m x_{\Lambda_m}} - \frac{\partial h_m}{\partial p_j}\Big|_{S_t^n x_{\Lambda_m}} \Big) \times \frac{\partial p_j}{\partial x_i}(S_t^n x_{\Lambda_m}). \tag{4.12}$$

Then

$$Z^1_{n,m}(p) \leqslant \sum_{j \in \Lambda_{n/2}} \Big| \frac{\partial h_m}{\partial p_j}\Big|_{S_t^m x_{\Lambda_m}} \times \Big(\frac{\partial p_j}{\partial x_i}(S_t^n x_{\Lambda_m}) - \frac{\partial p_j(S_t^m x_{\Lambda_m})}{\partial x_i} \Big) \Big| + \sum_{j \in \Lambda_m \setminus \Lambda_{n/2}} \Big| \frac{\partial h_m}{\partial p_j}\Big|_{S_t^m x_{\Lambda_m}} \times \Big(\frac{\partial p_j(S_t^n x_{\Lambda_m})}{\partial x_i} - \frac{\partial p_j(S_t^m x_{\Lambda_m})}{\partial x_i} \Big) \Big|, \tag{4.13}$$

defining $h(\cdot|\cdot)$ in a natural way as $\log P(\cdot|\cdot)$, the short range property (3.5) gives:

$$h_\Lambda(y_\Lambda) = h(y_\Omega | (y_\Lambda)_{\partial\Omega}) + h(y_{\Lambda \setminus \Omega})$$

if $\Lambda \supset \Omega$. Then, if $j \in \Lambda_m$ and $\Omega = \Lambda_j$,

$$\frac{\partial h_m}{\partial p_j}(y_{\Lambda_m}) = \frac{\partial h}{\partial p_j}\big(y_{\Lambda_j} \big| (y_{\Lambda_m})_{\partial\Lambda_j}\big). \tag{4.14}$$

So by Remark 2 we have, in virtue of (2.11), (2.10), and (2.5)

$$\sup_{x_\Lambda : \mathscr{L}(x_\Lambda) \leqslant \alpha} Z^1_{n,m}(p) \leqslant \frac{n^\nu (\eta\alpha\varphi(n))^\xi [b_1 c_5(t) \varphi(n)\alpha]^{n/2+1} \exp[c_4(t)\alpha^\delta]}{\mathscr{I}[n/2]!} + \sum_{\substack{j: \\ |j| > n/2}} (\eta\alpha\varphi(|j| + R))^\xi \sum_{k > |j| - r} \frac{[c_6(t)]^k}{k!} \times [\alpha\varphi(k) b_2]^{k/2}, \tag{4.15}$$

for some constants $\{b_i\} i = 1,2,\cdots$ and with $\Omega \subset \Lambda_r$. (Here $\mathscr{I}[x]$ denotes the integer part of $x \in \mathbb{R}$.)

$$Z^2_{n,m}(p) \leqslant \sum_{j \in \Lambda_{n/2}} \Big| \frac{\partial p_j(S_t^n x_{\Lambda_m})}{\partial x_i} \times \Big(\frac{\partial h(y_{\Lambda_{n/2}} | y_{\partial\Lambda_{n/2}})}{\partial p_j}\Big|_{S_t^m x_{\Lambda_m}} - \frac{\partial h(y_{\Lambda_{n/2}} | y_{\partial\Lambda_{n/2}})}{\partial p_j}\Big|_{S_t^n x_{\Lambda_m}} \Big) \Big| + \sum_{j \in \Lambda_m \setminus \Lambda_{n/2}} \Big| \frac{\partial p_j(S_t^n x_{\Lambda_m})}{\partial x_i} \times \Big(\frac{\partial h(y_{\Lambda_m \setminus \Lambda_{n/2}})}{\partial p_j}\Big|_{S_t^m x_{\Lambda_m}} - \frac{\partial h(y_{\Lambda_m \setminus \Lambda_{n/2}})}{\partial p_j}\Big|_{S_t^n x_{\Lambda_m}} \Big) \Big|. \tag{4.16}$$

$Z^2_{n,m}(p)$ may be estimated in the following way: The first term in (4.16) is convergent by the locality of dynamics [estimate (2.7)] that can be applied after controlling the second derivatives of h via (3.8), Remark 2, and (2.5), and $\partial p_j / \partial x_i$ via the bound (2.10); the second one has a convergence factor $\partial p_j / \partial x_i$ in virtue of the bound (2.10); the remaining term is estimated by (3.7), (4.14) and (2.5). $Z_{n,m}(q)$ is estimated in complete analogy.

Hence

$$\sup_{x_{\Lambda_m} : \mathscr{L}(x_{\Lambda_m}) \leqslant \alpha} \mu_i(x_{\Lambda_m}) Z_{n,m} \underset{\substack{n \to \infty \\ m > n}}{\longrightarrow} 0$$

uniformly on t on compact sets because

$$|\mu_i(x_{\Lambda_m})| \leqslant \alpha\varphi(|i| + 1)$$

and so, to prove (ii) it remains to estimate $T^1_{n,m}$. To estimate $T^1_{n,m}$ we first observe that, if $\mathscr{L}(x_{\Lambda_m}) < \alpha$, then

$$\Big| \frac{\partial h^n_m(t, x_{\Lambda_m})}{\partial x_i} \Big| \leqslant \Big| \sum_j \Big(\frac{\partial h}{\partial p_j}\Big|_{S_t^n x_{\Lambda_m}} \frac{\partial p_j}{\partial x_i}(S_t^n x_{\Lambda_m}) + \frac{\partial h}{\partial q_j}\Big|_{S_t^n x_{\Lambda_m}} \frac{\partial q_j}{\partial x_i} \Big) \Big| \leqslant \sum_j D^n_{j,i}(t,x) \alpha^\xi \varphi(|j|)^\xi \leqslant f_1(t,\alpha), \tag{4.18}$$

here and in the sequel $f_i(t,\alpha)$, $i = 1,2\cdots$ are continuous functions on t and α. Putting

$$\gamma^\alpha_{n,m}(x_{\Lambda_m}) = \chi_\alpha(x_{\Lambda_m}) \frac{\partial h^n_m(t, x_{\Lambda_m})}{\partial x_i} \mu_i(x_{\Lambda_m}), \tag{4.19}$$

then

$$T^1_{n,m}(t) = \int dx_\Omega \Big| \int dx_{\Lambda_m \setminus \Omega}\, \gamma^\alpha_{\Lambda,m}(x_{\Lambda_m}) \times P^c((S_t^n x_{\Lambda_m})_{n/2}) P((S_t^n x_{\Lambda_m})_{\Lambda_m \setminus \Lambda_{n/2}}) - P^c((S_t^m x_{\Lambda_m})_{n/2}) P((S_t^n x_{\Lambda_m})_{\Lambda_m \setminus \Lambda_{n/2}}), \tag{4.20}$$

where

$$P^c((y_{\Lambda_m})_k) = P((y)_{\Lambda_k} | y_{\partial\Lambda_k}) \quad m > k.$$

We have

$$T^1_{n,m}(t) \leqslant W_1 + \bar{W}_2, \tag{4.21}$$

where

$$W_1 = \int dx_{\Lambda_m}\, \gamma^\alpha_{n,m}(x_{\Lambda_m}) P((S_t^m x_{\Lambda_m})_{\Lambda_m \setminus \Lambda_{n/2}}) \times \big| P^c((S_t^m x_{\Lambda_m})_{n/2}) - P^c((S_t^n x_{\Lambda_m})_{n/2}) \big|, \tag{4.22}$$

$$\overline{W}_2 = \int dx_\Omega \left| \int dx_{\Lambda_m\setminus\Omega} [\gamma^\alpha_{n,m}(x_{\Lambda_m}) P^c((S^n_t x_{\Lambda_m})_{\Lambda_{n/2}})] \times [P(S^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) - P((S^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}})] \right|, \quad (4.23)$$

$$\overline{W}_2 \leqslant W_2 + W_3 + \overline{W}_4, \quad (4.24)$$

where denoting $\tilde{S}^n_t = S^{\Lambda_k\setminus\Omega}_t$ for $\Lambda_k \supset \Omega$ we define

$$W_2 = \int dx_{\Lambda_m} \gamma^\alpha_{n,m}(x_{\Lambda_m}) P^c((S^n_t x_{\Lambda_m})_{n/2}) \times \left| P((S^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) - P((\widetilde{S}^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) \right|, \quad (4.25)$$

$$W_3 = \int dx_{\Lambda_m} \gamma^\alpha_{n,m}(x_{\Lambda_m}) P^c((S^n_t x_{\Lambda_m})_{n/2}) \times \left| P((S^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) - P((\widetilde{S}^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) \right|, \quad (4.26)$$

$$\begin{aligned}\overline{W}_4 &= \int dx_\Omega \left| \int dx_{\Lambda_m\setminus\Omega} \gamma^\alpha_{n,m}(x_{\Lambda_m}) P^c((S^n_t x_{\Lambda_m})_{n/2}) \times P((\widetilde{S}^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) - P((\widetilde{S}^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) \right|, \\ &= \int dx_\Omega \left| \int dx_{\Lambda_m\setminus\Omega} P(x_{\Lambda_m\setminus\Lambda_{n/2}}) \times [\gamma^\alpha_{n,m}(\widetilde{S}^m_{-t} x_{\Lambda_m}) P^c((S^n_t \widetilde{S}^m_{-t} x_{\Lambda_m})_{n/2}) - \gamma^\alpha_{n,m}(\widetilde{S}^n_{-t} x_{\Lambda_m}) P^c((S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{n/2})] \right|, \end{aligned} \quad (4.27)$$

where we have used the Liouville theorem. Then

$$\overline{W}_4 \leqslant W_4 + \overline{W}_5, \quad (4.28)$$

where

$$W_4 = \int dx_{\Lambda_m} P(x_{\Lambda_m\setminus\Lambda_{n/2}}) \gamma^\alpha_{n,m}(\widetilde{S}^m_{-t} x_{\Lambda_m}) \times \left| P^c((S^n_t \widetilde{S}^m_{-t} x_{\Lambda_m})_{n/2}) - P^c((S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{n/2}) \right|, \quad (4.29)$$

$$\begin{aligned}\overline{W}_5 &= \int dx_\Omega \left| \int dx_{\Lambda_m\setminus\Omega} x_{\Lambda_m\setminus\Omega} P(x_{\Lambda_m\setminus\Lambda_{n/2}}) \times P^c((S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{n/2}) \times [\gamma^\alpha_{n,m}(\widetilde{S}^m_{-t} x_{\Lambda_m}) - \gamma^\alpha_{n,m}(\widetilde{S}^n_{-t} x_{\Lambda_m})] \right| \quad (4.30) \\ &= \int dx_\Omega \left| \int dx_{\Lambda_m\setminus\Omega} P((\widetilde{S}^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) \times P^c((S^n_t x_{\Lambda_m})_{n/2}) [\gamma^\alpha_{n,m}(\widetilde{S}^m_{-t} \widetilde{S}^n_t x_{\Lambda_m}) - \gamma^\alpha_{n,m}(x_{\Lambda_m})] \right|. \quad (4.31)\end{aligned}$$

Finally we have

$$\overline{W}_5 \leqslant W_5 + W_6, \quad (4.32)$$

where

$$W_5 = \int dx_{\Lambda_m} [\gamma^\alpha_{n,m}(\widetilde{S}^m_{-t} \widetilde{S}^n_t x_{\Lambda_m}) - \gamma^\alpha_{n,m}(x_{\Lambda_m})] P^c((S^n_t x_{\Lambda_m})_{n/2}) \times [P((\widetilde{S}^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) - P((S^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}})], \quad (4.33)$$

$$W_6 = \int dx_{\Lambda_m} P((S^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) \times P^c((S^n_t x_{\Lambda_m})_{n/2}) [\gamma^\alpha_{n,m}(\widetilde{S}^m_{-t} \widetilde{S}^n_{-t} x_{\Lambda_m}) - \gamma^\alpha_{n,m}(x_{\Lambda_m})]. \quad (4.34)$$

So $T^2_{n,m}(t) \leqslant \Sigma^6_{i=1} W_i$ and we estimate such sum term by term

$$W_1 \leqslant f_2(t,\alpha) \int dx_{\Lambda_m} P(S^m_{-t} x_{\Lambda_m}) \chi_\alpha(x_{\Lambda_m}) \times \left| 1 - \frac{P^c((S^n_t x_{\Lambda_m})_{n/2})}{P^c((S^m_t x_{\Lambda_m})_{n/2})} \right| \quad (4.35)$$

$$\leqslant f_2(t,\alpha) \sup_{\substack{x_{\Lambda_m}: \\ \mathscr{L}(x_{\Lambda_m})<\alpha}} \left| 1 - \exp\{ h((S^n_t x_{\Lambda_m})_{\Lambda_{n/2}} | (S^n_t x_{\Lambda_m})_{\partial\Lambda_{n/2}}) - h((S^m_t x_{\Lambda_m})_{\Lambda_{n/2}}) | ((S^m_t x_{\Lambda_m})_{\Lambda_{n/2}}) \} \right|.$$

So, combining (2.7) remark and (3.7) we obtain

$$\{h - h\} \leqslant \left[\left(\frac{n}{2} + R \right)^\nu \varphi \left(\frac{n}{2} \right)^\xi \right] \times \frac{[f_3(t,\alpha)\varphi(n)]^{n/2 - R + 1}}{\mathscr{I}[n/2 - R]!}, \quad (4.36)$$

$$W_2 \leqslant f_4(t,\alpha) \int dx_{\Lambda_m} \chi_\alpha(x_{\Lambda_m}) \frac{P^c((S^n_t x_{\Lambda_m})_{n/2})}{P^c((S^m_t x_{\Lambda_m})_{n/2})} \times P((S^m_t x_{\Lambda_m})) \left| 1 - \frac{P((\widetilde{S}^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}})}{P((S^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}})} \right|. \quad (4.37)$$

The above argument used in the estimation of W_1 shows that

$$\chi_\alpha(x_{\Lambda_m}) \frac{P^c((S^n_t x_{\Lambda_m})_{n/2})}{P^c((S^m_t x_{\Lambda_m})_{n/2})}$$

is uniformly bounded in n and m and t belonging to a compact, while

$$\sup_{\substack{x_{\Lambda_m}: \\ \mathscr{L}(x_{\Lambda_m})\leqslant\alpha}} \left| 1 - \exp\{ h((\widetilde{S}^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) - h((S^m_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) \} \right| \quad (4.38)$$

may be estimated by the use of (2.8), (4.14), Remark 2, and (3.7) [see also the bound (4.40) below]. W_3 is estimated with the same arguments that worked for W_2.

$$W_4 \leqslant f_5(t,\alpha) \int dx_{\Lambda_m} \chi_\alpha(\widetilde{S}^m_{-t} x_{\Lambda_m}) P(S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m}), \times \left| 1 - \frac{P^c((S^n_t \widetilde{S}^m_{-t} x_{\Lambda_m})_{n/2})}{P^c((S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{n/2})} \right| \times \left| \frac{P(x_{\Lambda_m\setminus\Lambda_{n/2}})}{P((S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}})} \right|. \quad (4.39)$$

W_4 may be evaluated by estimating ($K \subset \mathbb{R}'$, K compact)

$$\sup_{t\in K} \sup_{\substack{x: \\ \mathscr{L}(\widetilde{S}^m_{-t} x)<\alpha}} \left| h((S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}}) - h(x_{\Lambda_m\setminus\Lambda_{n/2}}) \right|, \quad (4.40)$$

$$\sup_{t\in K} \sup_{\substack{x: \\ \mathscr{L}(\widetilde{S}^m_{-t} x)\leqslant\alpha}} \left| h((S^n_t \widetilde{S}^m_{-t} x_{\Lambda_m})_{\Lambda_{n/2}} | (S^n_t \widetilde{S}^m_{-t} x_{\Lambda_m})_{\partial\Lambda_{n/2}}) - h((S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{\Lambda_{n/2}}) | (S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m})_{\partial\Lambda_{n/2}}) \right|. \quad (4.41)$$

To bound (4.40) we observe that if $\mathscr{L}(\widetilde{S}^m_{-t} x) \leqslant \alpha$, $j \in \Lambda_m \setminus \Lambda_{n/2}$

$$\begin{aligned}|q_j(S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m}) - q_j(x_{\Lambda_m})| &= |q_j(S^n_t \widetilde{S}^n_{-t} x_{\Lambda_m}) - q_j(S^n_t S^n_{-t} x_{\Lambda_m})| \\ &\leqslant \sum_i D^{\Lambda_n}_{j,i}(t, \xi_i(t)) \{ |p_i(\widetilde{S}^n_{-t} x_{\Lambda_m})\end{aligned}$$

$$- p_i(S^m_{-t}x_{\Lambda_m})| \vee |q_i(\widetilde{S}^n_{-t}x_{\Lambda_m}) - q_i(S^n_{-t}x_{\Lambda_m})|\}, \tag{4.42}$$

where $\xi_i(t)$ is some point in $\mathscr{X}_{\Lambda_m}$ for which $\mathscr{L}(\xi_i(t)) \leqslant f_6(t,\alpha)$. We estimate (4.40) via the bound (4.42) which is infinitesimal. In fact we divide the Σ_i in two sums. One in which the i's are "far" from j and so $D^{\Lambda_n}_{i,j}$ is small. The other one in which the i's are "near" to j and so far enough from Ω in such a way that the two dynamics $\widetilde{S}$ and S are almost the same and so the term in the brackets in (4.42) is small. More precisely by (4.14), (3.7) and remark 2:

$$\begin{aligned}(4.40) \leqslant & \sum_{j\in\Lambda_m\setminus\Lambda_{n/2}} (\eta\alpha e^{a|t|}\varphi(|j|+R)^\xi \\ & \times \sum_i D^{\Lambda_n}_{j,i}(t,\xi_i(t))\{|p_i(\widetilde{S}^n_{-t}x_{\Lambda_m}) \\ & - p_i(S^n_{-t}x_{\Lambda_m})| \vee |q_i(\widetilde{S}^n_{-t}x_{\Lambda_m}) \\ & - q_i\widetilde{S}^n_{-t}x_{\Lambda_m})|. \end{aligned}\tag{4.43}$$

Then

$$\begin{aligned}(4.40) \quad \leqslant & \sum_{j\in\Lambda_m\setminus\Lambda_{n/2}} \Big(\sum_{i\in\Lambda_{n/4}} + \sum_{i\notin\Lambda_{n/4}}\Big)(\eta\alpha e^{a|t|}\varphi(|j|+R))^\xi \\ & \times(D^{\Lambda_n}_{j,i}(t,\xi_i(t))\{\cdots\vee\cdots\}.\end{aligned}$$

The first double sums is bounded by a term converging more than exponentially as $n\to\infty$ in virtue of (2.10) and (2.5). The second double sum is also more than exponentially convergent by (2.8), which gives the convergence factor and (2.10) to bound the derivatives.

The same idea, applied to (4.41) [replacing (2.8)] gives also that (4.41)$\to 0$ as $h\to\infty$, $m>n$.

$$\begin{aligned}W_5 \leqslant f_6(t,\alpha)\int dx_{\Lambda_m}|\chi_\alpha(x_{\Lambda_m}) - \chi_\alpha(\widetilde{S}^m_{-t}\widetilde{S}^n_{-t}x_{\Lambda_m})| \\ \times P(S^n_t x_{\Lambda_m})\left|1 - \frac{P((\widetilde{S}^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}})}{P((S^n_t x_{\Lambda_m})_{\Lambda_m\setminus\Lambda_{n/2}})}\right|,\end{aligned}\tag{4.44}$$

so W_5 may be estimated as W_2.

Finally

$$\begin{aligned}W_6 \leqslant & \int dx_{\Lambda_m}P(S^n_t x_{\Lambda_m})|\gamma^\alpha_{n,m}\widetilde{S}^m_{-t}\widetilde{S}^n_t x_{\Lambda_m}) - \gamma^\alpha_{n,m}(x_{\Lambda_m})| \\ \leqslant & \int dx_{\Lambda_m}P(S^n_t x_{\Lambda_m})\mu_i\frac{\partial h^n_m}{\partial x_i}(t,x_{\Lambda_m}) \\ & \times|\chi_\alpha(\widetilde{S}^m_{-t}\widetilde{S}^n_t x_{\Lambda_m}) - \chi_\alpha(x_{\Lambda_m})| \\ & + \int dx_{\Lambda_m}P(\widetilde{S}^n_{-t}x_{\Lambda_m})\chi_\alpha(\widetilde{S}^m_{-t}\widetilde{S}^n_t x_{\Lambda_m}) \\ & \times\left|\mu_i\cdot\frac{\partial h^n_m}{\partial x_i}(t,x_{\Lambda_m}) - \mu_i\cdot\frac{\partial h^n_m}{\partial x_i}(t,\widetilde{S}^m_{-t}\widetilde{S}^n_t x_{\Lambda_m})\right|.\end{aligned}\tag{4.45}$$

The first term in virtue of (4.18) is bounded by

$$\begin{aligned}f_1(t,\alpha)\int P(dx)|\chi_\alpha(\widetilde{S}^m_{-t}\widetilde{S}^n_t S^n_{-t}x) - \chi_\alpha(S^n_{-t}x)| \\ \leqslant f_8(t,\alpha)\int P(dx)|\mathscr{L}(\widetilde{S}^m_{-t}\widetilde{S}^n_t S^n_{-t}x) - \mathscr{L}(\widetilde{S}^n_{-t}x)|.\end{aligned}\tag{4.46}$$

To estimate (4.46) we observe that defining

$$\overline{\mathscr{L}}(x) = \sup_{i\in\mathbb{Z}^\nu}\frac{\mathscr{L}_i(x)}{\bar\varphi(|i|)}, \tag{4.47}$$

where $\forall k\in\mathbb{R}$, $\bar\varphi(k) = \max(\log k, 1)$ one easily obtains

$$\overline{\mathscr{L}}(S^\Lambda_t x) \leqslant e^{\bar a t}\overline{\mathscr{L}}(x),\ \Lambda\subset\mathbb{Z}^\nu. \tag{4.48}$$

The set $\overline{\mathscr{X}} = \{x|\overline{\mathscr{L}}(x) < +\infty\}$ satisfies $P(\overline{\mathscr{X}}) = 1$.

Then, because of the fact that if $x\in\overline{\mathscr{X}}$ then $\widetilde{S}^m_{-t}\widetilde{S}^n_t S^n_{-t}x$ and $S^n_{-t}x$ are in $\overline{\mathscr{X}}$ in virtue of (4.48), it follows

$$\begin{aligned}(4.46) \leqslant & \int P(dx)\sup_{i\in\Lambda_{n/2}}\left|\frac{\mathscr{L}_i(\widetilde{S}^m_{-t}\widetilde{S}^n_t S^n_{-t}x)}{\varphi(|i|)}\right. \\ & \left. - \frac{\mathscr{L}_i(S^n_{-t}x)}{\varphi(|i|)}\right| + \int P(dx)\sup_{i\notin\Lambda_{n/2}}\left| \right. \\ & \left.\times\frac{\mathscr{L}_i(\widetilde{S}^m_{-t}\widetilde{S}^n_t S^n_{-t}x)}{\bar\varphi(|i|)} - \frac{\mathscr{L}_i(S^n_{-t}x)}{\bar\varphi(|i|)}\right|\frac{1}{\log n}\end{aligned}$$

for n large enough.

The first term is estimated by locality of dynamics [see, e.g., estimated (4.43)] and the last term is bounded by

$$2e^{\bar a t}\int P(dx)\overline{\mathscr{L}}(x)\cdot\frac{1}{\log h}.$$

The last term in (4.45) is estimated as $T^2_{m,n}$ [see (4.6) and following arguments].

Part (i) of (4.4) has already been proven by estimating T^n_m and so the proof is completed.

Proof of Theorem 3.1:

$$\begin{aligned}||P^n_j(t,\cdot) - P^m_j(t,\cdot)||_1 \leqslant & \int dx_{\Lambda_j}\left|\int dx_{\Lambda_m\setminus\Lambda_j}\chi_\alpha(x_{\Lambda_m})\right. \\ & \times[P(S^m_t x_{\Lambda_m})] - [P(S^n_t x_{\Lambda_m})]| \\ & + \int dx_{\Lambda_m}|(1-\chi_\alpha(x_{\Lambda_m})| \\ & \times P(S^m_t x_{\Lambda_m}) \\ & + \int dx_{\Lambda_m}|(1-\chi_\alpha(x_{\Lambda_m})|P(S^n_t x_{\Lambda_m}).\end{aligned}\tag{4.49}$$

So, to prove (3.10) [by the use of Proposition 4.1 part (i)] it is enough to prove that

$$\int dx_{\Lambda_m}(1-\chi_\alpha(x_{\Lambda_m}))P(S^h_t x_{\Lambda_m}), \tag{4.50}$$

goes to 0 for $\alpha\to\infty$ uniformly in m and h. Here $h=m$ or $h=n$. We have

$$\begin{aligned}(4.50) \leqslant & \int P(dx)(1-\chi_\alpha(S^h_t x)) \\ & \leqslant P(\{x\in\mathscr{X}|\mathscr{L}(S^h_t x) > \alpha\}) \\ & \leqslant P(\{x\in\mathscr{X}|\mathscr{L}(x) > (\exp - a|t|)\alpha\}) \\ & \leqslant A\exp(-b'(t)\cdot\alpha),\end{aligned}\tag{4.51}$$

where b' is a continuous function. The last step follows from (3.9). The same procedure used above and Proposition 4.1, part (ii) allows one to prove (3.12) once one proves that

$$\int dx_{\Lambda_m}(1-\chi_\alpha(x_{\Lambda_m}))\left|\frac{\partial P}{\partial_{x_i}}(S^h_t x)\mu_i(x)\right|, \tag{4.52}$$

goes to 0 for $\alpha\to\infty$ uniformly in m and n.

$$(4.52) \leqslant \sum_{i\in\Lambda_m}\int dx_{\Lambda_m}P(S^h_t x_{\Lambda_m})|\mu_i(x_{\Lambda_m})|$$

$$\times \sum_{x=p,q} \left| \left(\frac{\partial h_m}{\partial x_l} \right|_{S_t^h x_{\Lambda_m}} D_{l,i}^{\Lambda_h}(x_{\Lambda_m},t) \right|$$
$$\times \left| (1-\chi_\alpha(x_{\Lambda_m})) \right| \leqslant \int P(dx) \mathscr{L}(S_{-t}^h x) \varphi(|i|)$$
$$\times |1-\chi_\alpha(S_{-t}^h x_{\Lambda_m})| \sum_{l\in\Lambda_m} D_{l,i}^{\Lambda_h}(S_{-t}^h x)(\eta\varphi(|l|$$
$$\tag{4.53}$$
$$+R)\mathscr{L}(S_t^h x_{\Lambda_m}))^\xi \leqslant \sum_{l\in\Lambda_m} \sum_{k\geqslant|i-l|} \frac{c_5(t)^k}{k!}$$
$$\times \varphi(|i+k|)^{k/2+\xi+1} \int P(dx) \left|1-\chi_\alpha(S_t^h x_{\Lambda_m})\right|$$
$$\times \mathscr{L}^{k/2+\xi+1}(S_t^h x) \leqslant \sum_l \sum_{k\geqslant|i-l|} \frac{c_5(t)^k}{k!}$$
$$\times \varphi(|i+k|)^{k/2+\xi+1} \sum_{s\geqslant\alpha(t)} s^{k/2+\xi+1} e^{-bs},$$
where $c_5(t)$ is a continuous function and $\alpha(t) = e^{-a|t|}\alpha$. Since $\Sigma_s \Sigma_l \Sigma_k < +\infty$ the thesis is proven. The statement (3.11) follows trivially by the above arguments.

Remark: The above arguments work also for the "true" BBGKY hierarchy of continuous particle systems, if linear estimates for global quantities [See, e.g., (2.5)] would be known. Unfortunately we have at most polynomial estimates.[9]

ACKNOWLEDGMENTS

We are very grateful to Yu. Suhov for stimulating discussions and comments.

[1]A modern review on the argument may be found in: O.E. Lanford, III, *Time evolution of large classical system*, Vol. 38 of *Lecture Notes in Physics* (Springer, Berlin, New York, 1975).

[2]G. Gallavotti, O.E. Lanford, III, J.L. Lebowitz, J. Math. Phys. **11**, 2898 (1970); Ya. G. Sinai, Yu. M. Suhov, Theor. Math. Phys. **19**, 344 (1974).

[3]Yu. M. Suhov, Sov. Math. (Doklady) **5**, 244 (1979).

[4]B.M. Gurevich, Yu. M. Suhov, Commun. Math. Phys. **49**, 63 (1976); **54**, 81 (1977); **56**, 225 (1977).

[5]O.E. Lanford, III J.L. Lebowitz, E.H. Lieb, J. Stat. Phys. .**16**, 453 (1977).

[6]C. Marchioro, A. Pellegrinotti, M. Pulvirenti, Yu. M. Suhov, Commun. Math. Phys. **66**, 131 (1979).

[7]C. Marchioro, A. Pellegrinotti, M. Pulvirenti, and L. Triolo, J. Stat. Phys. **19**, 499 (1978).

[8]D. Ruelle, Commun. Math. Phys. **50**, 189 (1976).

[9]R. Dobrushin and J. Fritz, Commun. Math. Phys. **55**, 275 (1977); **57**, 67 (1977).

Two approaches to nonstationary relativistic thermodynamics

F. Bampi and A. Morro
Istituto di Matematica, Università di Genova, 16132 Genova, Italy

(Received 20 November 1979; accepted for publication 21 January 1980)

A nonstationary thermodynamic theory in general relativity is set up through a hidden variable approach. Such a theory, which is especially suited to the description of the behavior of heat conducting viscous fluids, considers evolution equations involving the gradient of the hidden variables too. The expression adopted for the second law differs from the usual Clausius–Duhem inequality because of the presence of an entropy extra-flux. Besides other results it is shown that, via a suitable choice of the free energy function, the theory so obtained reduces just to Müller–Israel's one.

1. INTRODUCTION

In recent years topics concerning viscosity mechanisms in cosmology have become more and more important particularly in connection with the attempt to account for the anomalously high entropy per baryon in the contemporary universe.[1] Unfortunately, the customary description of viscosity, namely Navier–Stokes' law both in the classical and in the relativistic context, suffers from an unpleasant feature in that it leads to parabolic differential equations whereby disturbances would propagate at infinite speed in contradiction with the principle of causality. From the physical standpoint this fact is hardly surprising because the conventional theory is applicable only to phenomena which are slowly varying on space and time scales characterized by the mean free path and mean collision time. Accordingly, the theory is inadequate to wave propagation problems as well as to many phenomena in high-energy astrophysics involving steep gradients or rapid variations.

Undoubtedly any acceptable theory of nonstationary thermodynamics must cure this deficiency besides the analogous one occurring in Fourier's theory of heat conduction. In this sense, after the paper by Cattaneo,[2] many proposals concerning the theory of heat conduction appeared in the literature both in the classical and in the relativistic framework.[3] On the other hand, as far as we are aware, only very few endeavors to eliminate the paradox of infinite speed in viscous materials have been presented; among them we cite Ref. 4.

Lately, Israel[5] has elaborated a relativistic theory of nonstationary thermodynamics which gives a unified account of viscosity and heat conduction. The novelty of Israel's approach is the abandonment of the usual hypotheses whereby the entropy flux is simply proportional to heat flux and the entropy density is independent of heat conduction and viscosity. Precisely, Israel allows for an entropy extra-flux determined by viscous stresses and heat flux through second-order terms.[6] The irreversible thermodynamic theory so obtained is applicable to nonstationary processes and does not violate causality. We remark that Israel's work, which has a purely phenomenological character, is corroborated by subsequent papers of Israel and Stewart[7-9] concerning kinetic theory.

The recent scientific literature bears evidence of an increasing interest in the hidden variable approach to nonstationary thermodynamics.[10] The utility of this approach has been shown by the authors in connection with wave propagation in relativistic thermo-viscous fluids.[11] Motivated by the encouraging results achieved so far, in this work we set up a relativistic nonstationary thermodynamics with hidden variables. In doing this we are concerned with states which are arbitrarily away from equilibrium; the generality of our approach will allow us to obtain Israel's phenomenological equations merely through a suitable choice of the free energy function.

The plan of the paper is as follows. The general description of heat conducting viscous fluids via hidden variables is delivered in Sec. 2. Then, for the benefit of the reader desiring a quick overview and for a motivation of next developments, Sec. 3 assembles the main ideas underlying Müller–Israel's theory. So we are able, in Sec. 4, to incorporate some of these ideas in our theory and to complete it by deriving the constitutive equations of the fluid through the restrictions placed by the second law of thermodynamics. Finally, Sec. 5 exhibits a detailed comparison between Müller–Israel's theory and ours.

2. HIDDEN VARIABLES IN RELATIVISTIC THERMO-VISCOUS FLUIDS

Setting aside a formal mathematical account of the subject,[12] we say that a material with hidden variables consists of a set of response functions

$$\phi = \phi(y,\alpha),$$

and of a function f governing the evolution of the hidden variables α through the ordinary differential equation

$$\dot{\alpha} = f(y,\alpha),$$

a superposed dot denoting the proper time derivative with respect to the particle velocity u. The symbol y stands for a suitable set of real variables; for instance, when dealing with fluids y may be identified with the pair (ϑ,r) of the absolute temperature ϑ and the proper rest mass density r. If, further, the problem at hand involves irreversible effects like viscosity and heat conduction, we have to account for the dependence on the relativistic temperature gradient $\lambda_\alpha = h_\alpha{}^\beta(\vartheta_{,\beta}$

 0022-2488/80/051201-04$01.00

$+ \vartheta\dot{u}_\beta$),on the expansion $\theta = u^\alpha{}_{;\alpha}$, and on the shear $\sigma_{\alpha\beta} = h_\alpha{}^\lambda h_\beta{}^\mu u_{(\lambda;\mu)} - \frac{1}{3}\theta h_{\alpha\beta}$, h being the projector onto the three-space orthogonal to u. Yet, as shown in a previous paper,[13] compatibility with wave propagation at finite speed implies that σ, θ, and λ cannot enter the argument of ϕ though they affect the value of $\dot{\alpha}$. Accordingly a model describing a thermo-viscous fluid may be represented by a set of equations as

$$\phi = \phi(\vartheta,r,\alpha),$$
$$\dot{\alpha} = f(\vartheta,r,\sigma,\theta,\lambda,\alpha).$$

This ansatz appears to be sufficiently general and appropriate for wave propagation topics; however, to avoid inessential formal complications, this scheme has been investigated in detail by letting f be a linear function and α be the triple of a scalar Θ and of a second-order symmetric traceless tensor Σ and a vector Λ subject to $\Sigma^{\beta\gamma}u_\gamma = 0$, $\Lambda^\gamma u_\gamma = 0$.[11]

With a view to the applications we have in mind the previous model needs an improvement. Precisely, by analogy with customary theories of nonstationary thermodynamics[5,6] involving spatial gradients of the quantities under consideration, a dependence on the spatial derivatives of the hidden variables must be taken into account. Yet, in this case too, compatibility with wave propagation leads us to assert that such a dependence is admissible only in the evolution equation. On the other hand, to make an immediate comparison with Müller–Israel's theory we specialize the function f so as to obtain a proper set of linear evolution equations. On the basis of these observations, we claim that the behavior of a thermo-viscous fluid is characterized by the response functions

$$\phi = \phi(\vartheta,r,\Sigma,\Theta,\Lambda),$$

and by the evolution equations

$$\langle\dot{\Sigma}_{\alpha\beta}\rangle = (1/\tau_s)(\sigma_{\alpha\beta} - \Sigma_{\alpha\beta}) + a\langle\Lambda_{\alpha;\beta}\rangle, \tag{2.1}$$

$$\dot{\Theta} = (1/\tau_b)(\theta - \Theta) + b\Lambda^\alpha{}_{;\alpha}, \tag{2.2}$$

$$h_\alpha{}^\beta\dot{\Lambda}_\beta = (1/\tau_c)(\lambda_\alpha - \Lambda_\alpha) + h_\alpha{}^\beta(c\Theta_{,\beta} + d\Sigma_\beta{}^\mu{}_{;\mu}), \tag{2.3}$$

where $\langle A_{\alpha\beta}\rangle = \frac{1}{2}h_\alpha{}^\mu h_\beta{}^\nu(A_{\mu\nu} + A_{\nu\mu} - \frac{2}{3}h_{\mu\nu}h^{\rho\sigma}A_{\rho\sigma})$ and a, b, c, d are, as yet, indeterminate coefficients. For the sake of definiteness, the symbol ϕ may be thought of as the set (ψ,s,S,q) of the specific free energy ψ, the specific entropy s, the stress S, and the heat flux q. The form of the Eqs. (2.1)–(2.3) assigns to the parameters τ_s, τ_b, $\tau_c > 0$ the meaning of relaxation times.

It is a noteworthy property of the hidden variables Σ, Θ, Λ, considered as solutions of the evolution Eqs. (2.1)–(2.3), that the values $\Sigma(t)$, $\Theta(t)$, $\Lambda(t)$ at the proper time t are independent of the present values $\sigma(t)$, $\theta(t)$, $\lambda(t)$. To clarify the meaning of this assertion consider, for example, the Eq. (2.2); at any particle of the fluid a formal integration yields

$$\begin{aligned}\Theta(t) = \frac{1}{\tau_b}\int_{t_0}^{t} &\exp[-(t-\xi)/\tau_b] \\ &\times[\theta(\xi) + b\tau_b\Lambda^\alpha{}_{;\alpha}(\xi)]\,d\xi \\ &+ \Theta(t_0)\exp[-(t-t_0)/\tau_b].\end{aligned} \tag{2.4}$$

Now, given a set of C^1 functions σ_1, θ_1, λ_1, look at C^1 functions σ_2, θ_2, λ_2 such that $\sigma_1 = \sigma_2$, $\theta_1 = \theta_2$, $\lambda_1 = \lambda_2$ in $(t_0, t-\epsilon)$ and $\sigma_2(t)$, $\theta_2(t)$, $\lambda_2(t)$ are arbitrary. Then, by virtue of (2.4), it is evident that the choice of a small enough ϵ makes it as little as we please the change of $\Theta(t)$ induced by $\sigma_1, \theta_1, \lambda_1 \to \sigma_2, \theta_2, \lambda_2$. The same property holds for the tensor quantities Σ, Λ as well; in this instance the formal integration is to be carried out by referring to the invariant components of Σ, Λ with respect to an orthonormal spatial triad.

3. RESUMÉ OF MÜLLER–ISRAEL THEORY

Before considering the relativistic thermodynamics elaborated by Israel, let us cast an eye on the classical theory developed by Müller.[6]

The starting point for Müller's nonstationary thermodynamics is the assumption that the entropy density s depends also on the heat flux q_i and on the stress tensor $-p_{ij} = -\bar{p}_{ij} - \frac{1}{3}p_{ll}\delta_{ij}$ besides two extensive quantities like e.g., the internal energy e and the mass density ρ. Then, introducing an entropy extra-flux $N_i = -L\,p_{ll}q_i - K\bar{p}_{ij}q_j$ and restricting the attention to states which are not too remote from equilibrium, Müller achieves an expression of the entropy production $\rho(ds/dt) + (\partial/\partial x^i)(q_i/T + N_i)$ which turns out to be identically nonnegative if suitable linear phenomenological equations for q_i, $\bar{p}_{ij}$, and p_{ll} are adopted.

While writing his paper, Israel was not aware of the paper by Müller. Nevertheless, in a sense, Israel's work may be viewed as the relativistic counterpart of Müller's theory.[7] In fact, paralleling Müller's classical procedure, it is a simple matter to obtain a further motivation of Israel's phenomenological laws in the particle frame

$$\langle -S_{\alpha\beta}\rangle = -2\eta\langle\sigma_{\alpha\beta} + \beta_2\langle -S_{\alpha\beta}\rangle^{\cdot} - \bar{\alpha}_1 q_{\alpha;\beta}\rangle, \tag{3.1}$$

$$\pi = -\zeta(\theta + \beta_0\dot{\pi} - \bar{\alpha}_0 q^\alpha{}_{;\alpha}), \tag{3.2}$$

$$q^\alpha = -\kappa h^{\alpha\beta}(\lambda_\beta + \vartheta\bar{\beta}_1\dot{q}_\beta - \vartheta\bar{\alpha}_0\pi_{,\beta} - \vartheta\bar{\alpha}_1\langle -S_\beta{}^\mu\rangle_{;\mu}), \tag{3.3}$$

where π is the bulk stress and $\langle -S_{\alpha\beta}\rangle$ is the shear stress, while the symbols β's and $\bar{\alpha}$'s are those used by Israel.[5]

In spite of the formal resemblance between Israel's ansatz (3.1)–(3.3) and our Eqs. (2.1)–(2.3), the two sets of equations have a deeply different conceptual meaning. Precisely, Eqs. (3.1)–(3.3) are constitutive equations for the response functions $\langle -S\rangle$, π, and q, while Eqs. (2.1)–(2.3) govern the evolution of the hidden variables Σ, Θ, Λ. Yet, a close connection between Eqs. (3.1)–(3.3) and (2.1)–(2.3) is attained if direct relations between $\langle -S\rangle$, π, q and Σ, Θ, Λ are accessible. This point will be investigated carefully in Sec. 4 where such relations are derived through compatibility with thermodynamics. In passing we note that Eqs. (3.1)–(3.3) were proposed by Israel as purely phenomenological laws; their validity was subsequently substantiated by Israel–Stewart's works[8,9] concerning kinetic theory arguments.[14]

It is worth remarking that the use of the proper time derivative is open to an objection, namely, that such a derivative appears not to be objective. To remedy this unpleasant feature, sometimes researchers have recourse to other derivatives.[15] Notwithstanding this, the proper time derivative is frequently adopted because, to our mind, it gives rise to handier formulas in several contexts and, meanwhile, it leaves the description of the physical behavior qualitatively unaffected.

4. HIDDEN VARIABLE APPROACH AND THE SECOND LAW OF THERMODYNAMICS

This section is intended as a completion of the hidden variable approach, outlined in Sec. 2, for the purpose of comparing it with Müller–Israel's theory surveyed in Sec. 3. Following along the lines of Müller's work, we assume the existence of an entropy extra-flux N in addition to the flux q/ϑ. Accordingly, the second law of thermodynamics may be given the form of the inequality

$$-r(\dot\psi + s\dot\vartheta) + S^{\alpha\beta}\sigma_{\alpha\beta} + \tfrac{1}{3}S^{\alpha}{}_{\alpha}\theta - (1/\vartheta)q^{\alpha}\lambda_{\alpha} + \vartheta N^{\alpha}{}_{;\alpha} \geqslant 0, \tag{4.1}$$

which must hold identically at any particle. Within the present context, the most natural choice of N involving the hidden variables Σ, Θ, Λ is expressed by

$$N^{\alpha} = K\Sigma^{\alpha\beta}\Lambda_{\beta} + L\Theta\Lambda^{\alpha}, \tag{4.2}$$

where K, L are phenomenological coefficients; a possible connection of K, L with a, b, c, d will be given shortly. Substitution of (2.1)–(2.3) and (4.2) in (4.1) provides

$$\begin{aligned}
&-r(\psi_{\vartheta} + s)\dot\vartheta + [S^{\alpha\beta} - (r/\tau_s)\psi_{\Sigma_{\alpha\beta}}]\sigma_{\alpha\beta}\\
&+ [r^2\psi_r + \tfrac{1}{3}S^{\alpha}{}_{\alpha} - (r/\tau_b)\psi_{\Theta}]\theta\\
&- [(1/\vartheta)q^{\alpha} + (r/\tau_c)\psi_{\Lambda_\alpha}]\lambda_{\alpha}\\
&+ r[(1/\tau_s)\psi_{\Sigma_{\alpha\beta}}\Sigma_{\alpha\beta}\\
&+ (1/\tau_b)\psi_{\Theta}\Theta + (1/\tau_c)\psi_{\Lambda_\alpha}\Lambda_{\alpha}] + (\vartheta K\Sigma^{\alpha\beta} - ra\psi_{\Sigma_{\alpha\beta}})\Lambda_{\alpha;\beta}\\
&+ (\vartheta L\Theta - rb\psi_{\Theta})\Lambda^{\alpha}{}_{;\alpha}\\
&+ (\vartheta L\Lambda^{\alpha} - rc\psi_{\Lambda_\alpha})\Theta_{,\alpha}\\
&+ (\vartheta K\Lambda^{\alpha} - rd\psi_{\Lambda_\alpha})\Sigma_{\alpha}{}^{\beta}{}_{;\beta} \geqslant 0,
\end{aligned} \tag{4.3}$$

where the subscripts ϑ, r, Σ, Θ, Λ denote partial differentiations. Owing to the independence of the hidden variables, and hence of ψ, of the present values $\dot\vartheta, \sigma, \theta$, and λ, we conclude that (4.3) holds identically only if

$$s = -\psi_{\vartheta}, \tag{4.4}$$

$$S^{\alpha\beta} = -r^2\psi_r h^{\alpha\beta} + (r/\tau_s)\psi_{\Sigma_{\alpha\beta}} + (r/\tau_b)\psi_{\Theta}h^{\alpha\beta}, \tag{4.5}$$

$$q^{\alpha} = -(r/\tau_c)\psi_{\Lambda_\alpha}, \tag{4.6}$$

and if

$$\begin{aligned}
&r[(1/\tau_s)\psi_{\Sigma_{\alpha\beta}}\Sigma_{\alpha\beta} + (1/\tau_b)\psi_{\Theta}\Theta + (1/\tau_c)\psi_{\Lambda_\alpha}\Lambda_{\alpha}]\\
&\quad + (\vartheta K\Sigma^{\alpha\beta} - ra\psi_{\Sigma_{\alpha\beta}})\Lambda_{\alpha;\beta} + (\vartheta L\Theta - rb\psi_{\Theta})\Lambda^{\alpha}{}_{;\alpha}\\
&\quad + (\vartheta L\Lambda^{\alpha} - rc\psi_{\Lambda_\alpha})\Theta_{,\alpha} + (\vartheta K\Lambda^{\alpha} - rd\psi_{\Lambda_\alpha})\Sigma_{\alpha}{}^{\beta}{}_{;\beta} \geqslant 0.
\end{aligned} \tag{4.7}$$

Each function ψ satisfying (4.7) makes the response functions s, S, q automatically consistent with the second law of thermodynamics in the form (4.1). Among the admissible choices of ψ, we confine our attention to a free energy leading to a scheme endowed with satisfactory physical properties. Indeed, letting

$$\psi = \Psi(\vartheta, r) + (1/r)(\eta\tau_s\Sigma_{\alpha\beta}\Sigma^{\alpha\beta} + \tfrac{1}{2}\zeta\tau_b\Theta^2 + (\kappa\tau_c/2\vartheta)\Lambda_{\alpha}\Lambda^{\alpha}), \tag{4.8}$$

the inequality (4.7) is identically true if and only if $\eta\geqslant 0$, $\zeta\geqslant 0$, $\kappa\geqslant 0$, once we introduce the further relationships

$$a = \frac{\vartheta K}{2\eta\tau_s}, \quad b = \frac{\vartheta L}{\zeta\tau_b}, \quad c = \frac{\vartheta^2 L}{\kappa\tau_c}, \quad d = \frac{\vartheta^2 K}{\kappa\tau_c}, \tag{4.9}$$

between the phenomenological coefficients. Also, (4.4)–(4.6), and (4.8) deliver

$$s = -\Psi_{\vartheta} + \frac{\kappa\tau_c}{2r\vartheta^2}\Lambda_{\alpha}\Lambda^{\alpha}, \tag{4.10}$$

and

$$S_{\alpha\beta} = -ph_{\alpha\beta} + 2\eta\Sigma_{\alpha\beta} + \zeta\Theta h_{\alpha\beta}, \tag{4.11}$$

$$q_{\alpha} = -\kappa\Lambda_{\alpha}, \tag{4.12}$$

where $p = r^2\psi_r$. Equations (4.11) and (4.12) turn out to be the most natural generalizations of Fourier's and Navier–Stokes' laws. In fact, in the case of stationary uniform processes, asymptotically Eqs. (2.1)–(2.3) yield

$$\Sigma = \sigma, \quad \Theta = \theta, \quad \Lambda = \lambda, \tag{4.13}$$

whereby the hidden variables coincide with the corresponding real variables, and what is more Eqs. (4.11) and (4.12) reduce to the usual laws of viscosity and heat conduction. In this instance the hidden variables disappear at the outcome thus allowing us to describe the fluid in terms of real variables only. This feature, occurring here in a trivial manner, is true in any general problem in the sense that the hidden variables Σ, Θ, Λ are to be viewed as solutions of the evolution Eqs. (2.1)–(2.3).

5. CONCLUDING REMARKS

We are now in a position to examine the connection between Müller–Israel's theory and ours. To this purpose look at the Eqs. (2.1)–(2.3) and write the expressions for the hidden variables Σ, Θ, and Λ in terms of $\langle -S\rangle$, $\pi = -\tfrac{1}{3}S^{\alpha}{}_{\alpha} - p$, and q via the relations (4.11) and (4.12). Then it follows at once that the equations so obtained coincide exactly with Israel's equations (3.1)–(3.3) provided the following identifications

$$\beta_0 = \tau_b/\zeta, \quad \bar\beta_1 = \tau_c/\kappa\vartheta, \quad \beta_2 = \tau_s/2\eta,$$
$$\bar\alpha_0 = \vartheta L/\zeta\kappa, \quad \bar\alpha_1 = \vartheta K/2\eta\kappa,$$

are made. This precise connection enables us to obtain a further significant result in that the estimates of Israel's phenomenological coefficients, established through Israel–Stewart's kinetic approach,[9] apply to our phenomenological coefficients as well.

In conclusion, it seems that two points weigh in favor of the hidden variable tool. First, as shown above, once the evolution equations are given, the response functions of the material are ultimately derived as a consequence of the second law of thermodynamics. This makes the whole approach automatically consistent with thermodynamics. Second, the hidden variables allow a very flexible account of irreversible phenomena in that any choice of ψ, satisfying the inequality (4.7), supplies admissible theories.

ACKNOWLEDGMENTS

We would like to thank Professor M.A. Anile for bringing the articles of Israel and Stewart to our attention. This

research is supported by the "Laboratorio per la Matematica Applicata-C.N.R., Genova" through the project "Conservazione del Suolo-Dinamica dei Litorali."

[1]S. Weinberg, Astrophys. J. **168**, 175 (1971); Z. Klimek, Acta Astron. **25**, 79 (1975).

[2]C. Cattaneo, Atti Semin. Mat. Fis. Univ. Modena **3**, 83 (1948).

[3]We refer the interested reader to G.A. Maugin, J. Phys. A **7**, 465 (1974); and to the references cited therein.

[4]B.D. Coleman, J.M. Greenberg, and M.E. Gurtin, Arch. Ration. Mech. Anal. **22**, 333 (1966); M. Carrassi, and A. Morro, Nuovo Cimento B **9**, 321 (1972).

[5]W. Israel, Ann. Phys. **100**, 310 (1976).

[6]The classical counterpart of this procedure is performed in I. Müller, Z. Phys. **198**, 329 (1967); that is why the relativistic theory is referred to as Müller–Israel's theory.

[7]J.M. Stewart, Proc. R. Soc. London A **357**, 59 (1977).

[8]W. Israel and J.M. Stewart, Phys. Lett. A **58**, 213 (1976).

[9]W. Israel and J.M. Stewart, Proc. R. Soc. London A **365**, 43 (1979).

[10]See, e.g., the classical applications carried out in W. Kosiński and P. Perzyna, Arch. Mech. **24**, 629 (1972); A. Morro, Arch. Mech. **32**, 145 (1980).

[11]F. Bampi and A. Morro, Acta Phys. Pol. B **10**, 1081 (1979).

[12]Besides various applications of the hidden variable approach, a formal mathematical account is delivered, for example, in B.D. Coleman and M.E. Gurtin, J. Chem. Phys. **47**, 597 (1967); W.A. Day, Arch. Ration. Mech. Anal. **62**, 367 (1976).

[13]A. Morro, Rend. Semin. Mat. Univ. Padova (to be published).

[14]In connection with kinetic derivations of nonstationary relativistic thermodynamics we mention also the paper M. Kranyš, Nuovo Cimento B **8**, 417 (1972), where Chernikov's generalization of Grad's thirteen-moment approximation is so developed as to avoid the paradox of infinite propagation speed.

[15]See, e.g., B. Carter and H. Quintana, Proc. R. Soc. London A **331**, 57 (1972); G.A. Maugin, J. Math. Phys. **19**, 1198 (1978).

Localized solutions of a charged nonlinear spinor field in a Coulomb-like potential

Antonio F. Rañada and Juan M. Usón [a)]
Departmento de Física Teórica and GIFT, Universidad Complutense, Madrid-3, Spain

(Received 2 October 1979; accepted for publication 21 December 1979)

We study some global properties of a self-interacting Dirac field which has been used to construct a classical stable model of the hydrogen atom. It is shown that the normalization condition produces the correct values of both charge and spin and also the correct Landé factors for the states with $j = 1/2$. The linear limit is given by the branching points of the continuum of solutions.

1. INTRODUCTION

In a previous paper (Rañada, 1977),[1] to be referred to as I, a nonlinear model of the hydrogen atom has been studied. Therein the electron is represented by a Dirac field with a self-interaction of the type $(\bar{\psi}\psi)^2$. For a very large interval of values of the self-coupling constant, the effect of the nonlinearity was found to be too small to be appreciated. In spite of this and of the fact that the linear theory leads to very good results in atomic physics there are some reasons to make such a study. For instance, Weyl (1950)[2] proposed some geometrical arguments which indicate that gravitation induces nonlinear self-coupling terms in the equations of motion of spin 1/2 fields. Even if the coupling constant were very small there is an important qualitative change which must be considered carefully. On the other hand, nonlinear Dirac fields have interesting localized solutions (Soler, 1970)[3] and can be used to construct models of the nucleon (Rañada *et al.* 1974, (Rañada and Vázquez, 1976).[4,5] We would also like to point out that the present interest in nonlinear equations is leading to the discovery of completely unexpected results.

The model was used in I to make some considerations on the relation between classical and quantum physics and to explore the possibility of a nonlinear quantum mechanics. However, the results were based on numerical analysis. In this work we study some aspects of the model from a more rigorous point of view using an analytical approximation.

In Sec. 2 we report some global properties of the solutions and obtain lower bounds of their radius.

In Sec. 3 we consider the problem of the bifurcation of the nonlinear solutions at the linear value of the frequency. For this purpose we use the Ritz–Galerkin method and find complete agreement with the previously obtained numerical results.

2. CLASSICAL NONLINEAR CHARGED DIRAC FIELD IN A COULOMB POTENTIAL

We consider the atom described by the Lagrangian

$$\mathscr{L} = \frac{i}{2}\{\bar{\psi}\gamma^\mu\partial_\mu\psi - (\partial_\mu\bar{\psi})\gamma^\mu\psi\} - m\bar{\psi}\psi - e\bar{\psi}\gamma^0\psi A_0 + \lambda(\bar{\psi}\psi)^2, \tag{2.1}$$

in which the Dirac field for the electron is physically interpreted so that

$$\rho_e \equiv \rho_e(\mathbf{r},t) = e\psi^+\psi \tag{2.2}$$

represents the electric charge density at $(\mathbf{r},t)$.

A_0 is the electrostatic potential created by the nucleus and the quadratic term is the dominant part (in the weak field limit) of the one suggested first by Weyl (1950) as a dynamical consequence of the spin.

Out notation will be

$$g^{\mu\mu} = (1,-1,-1,-1),\ \hbar = 1,\ c = 1,\ e = -|e|,\quad \alpha = e^2/4\pi, \tag{2.3}$$

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix},$$

σ_k being the well-known Pauli matrices.

From (2.1) we obtain the field equation

$$i\gamma^\mu\partial_\mu\psi - m\psi - eA_0\gamma^0\psi + 2\lambda(\bar{\psi}\psi)\psi = 0, \tag{2.4}$$

as well as the corresponding equation for the adjoint spinor

$$-i\partial_\mu\bar{\psi}\gamma^\mu - m\bar{\psi} - eA_0\bar{\psi}\gamma^0 + 2\lambda(\bar{\psi}\psi)\bar{\psi} = 0. \tag{2.5}$$

Both of them produce the conservation law for the electric current

$$\partial_\mu j^\mu = \partial_\mu(\bar{\psi}\gamma^\mu\psi) = 0, \tag{2.6}$$

so that the well known conservation of the electric charge

$$Q = e\int_{\mathbf{R}^3} \psi^+\psi\, d^3\mathbf{r}, \tag{2.7}$$

for the case of localized solutions is not spoiled by the nonlinear term.

We are interested in the behavior of stationary solutions; therefore we study solutions factorized in spherical coordinates in the form

$$\psi = \psi_j l_{j3} = e^{-i\omega t}\begin{bmatrix} g(r) & \mathscr{Y}^{j3}_{j,l} \\ if(r) & \mathscr{Y}^{j3}_{j,l'} \end{bmatrix} \tag{2.8}$$

restricting ourselves to the case in which $j = \frac{1}{2}$, $\mathscr{Y}^{j3}_{j,l}$ being the spinor spherical harmonics, and l' depending on l and j in the usual way.

The change of variables

$$(g,f) = \left(\frac{2m\pi}{|\lambda|}\right)^{1/2}(G,F),\quad \rho = mr,$$
$$eA_0 = mV,\quad \text{and}\quad \Omega = \omega/m \tag{2.9}$$

[a)] Present address: Physics Department, Joseph Henry Laboratories, Princeton University, P.O. Box 708, Princeton, NJ 08540.

 0022-2488/80/051205-05$01.00

leads to the equations

$$F' + \frac{1-\kappa}{\rho} F - (1 - \Omega + V)G + \operatorname{sig}\lambda\,(G^2 - F^2)G = 0,$$
$$G' + \frac{1+\kappa}{\rho} G - (1 + \Omega - V)F + \operatorname{sig}\lambda\,(G^2 - F^2)F = 0, \tag{2.10}$$

κ being defined in the usual way (Sakurai 1973).[6]

The electrostatic potential is given by

$$V(\rho) = -\frac{\alpha}{\rho}, \quad \rho > \rho_0,$$
$$V(\rho) = -\frac{\alpha}{\rho_0}\left(\frac{3}{2} - \frac{1}{2}\left(\frac{\rho}{\rho_0}\right)^2\right), \quad \rho < \rho_0, \tag{2.11}$$

$R_0 = \rho_0/m$ being the radius of the nucleus which we regard as a sphere with a homogeneous charge density distribution.

The energy of any given solution is

$$E = \frac{2\pi}{|\lambda|m}(\Omega I_1 + \tfrac{1}{2}\operatorname{sig}\lambda I_2), \tag{2.12}$$

as shown in I, where

$$I_1 = \int_0^\infty (G^2 + F^2)\rho^2\,d\rho,$$

and

$$I_2 = \int_0^\infty (G^2 - F^2)^2\rho^2\,d\rho, \tag{2.13}$$

the corresponding norm being

$$N = \frac{2\pi}{|\lambda|m^2} I_1, \tag{2.14}$$

the usual L^2-norm.

Several families of solutions have been obtained numerically in I corresponding to the waves $1s_{1/2}$, $2s_{1/2}$, $2p_{1/2}$, $3s_{1/2}$, and $3p_{1/2}$. To prove their existence remains an open question although it is possible to obtain some global conditions which are useful to check the accuracy of the numerical solutions (Vázquez 1977).[7]

Lemma 1: The localized solutions of (2.10) are the critical points of the functional $I(F,G) = \int_0^\infty (G^2 - F^2)^2\rho^2\,d\rho$. subject to the constraint

$$\int_0^\infty \left\{(\Omega - V)(G^2 + F^2) - (G^2 - F^2) - \frac{2\kappa FG}{\rho} + GF' - FG'\right\}\rho^2\,d\rho = -R, \quad R > 0. \tag{2.15}$$

Proof: If $(\overline{G},\overline{F})$ is a critical point of the variational problem, then

$$\overline{F}' + \frac{1-\kappa}{\rho}\overline{F} - (1 - \Omega + V)\overline{G} + 2M\ \operatorname{sig}\lambda\,(\overline{G}^2 - \overline{F}^2)\overline{G} = 0,$$
$$\overline{G}' + \frac{1+\kappa}{\rho}\overline{G} - (1 + \Omega - V)\overline{F} + 2M\ \operatorname{sig}\lambda\,(\overline{G}^2 - \overline{F}^2)\overline{F} = 0, \tag{2.16}$$

where M sigλ must be a positive constant in order to satisfy the constraint. Let us remark that sigλ does affect only to the sign of the Lagrange multiplier. Thus, $(2|M|)^{1/2}\cdot(\overline{F},\overline{G})$ satisfies (2.10).

Lemma 2: The localized solutions of (2.10) satisfy the following integral conditions:

$$\kappa\int_0^\infty (F^2 + G^2)\rho\,d\rho = \int_0^\infty 2(\Omega + |V|)FG\rho^2\,d\rho, \tag{2.17}$$

$$(\tfrac{1}{2} - \kappa)\int_0^\infty (F^2 - G^2)\,d\rho = 2\int_0^\infty (\Omega + |V|)FG\rho\,d\rho, \tag{2.18}$$

$$\int_0^\infty (\Omega - 1)G^2\rho^2\,d\rho + \int_0^\infty (\Omega + 1)F^2\rho^2\,d\rho + \int_0^\infty (G^2 + F^2)|V|\rho^2\,d\rho - \frac{\operatorname{sig}\lambda}{2}\int_0^\infty (G^2 - F^2)^2\rho^2\,d\rho = 0. \tag{2.19}$$

Proof: Equations (2.17) and (2.18) can easily be obtained by multiplying the field equations (2.10), respectively, by F and G, subtracting them from each other, and afterwards integrating over $\mathbb{R}^3$ using $\rho^2\,d\rho$ and $\rho^3\,d\rho$ as integration kernels.

They give lower limits on the size of the atom:

$$\langle\rho\rangle \geqslant \langle\rho^{-1}\rangle^{-1} \geqslant \frac{1-\alpha}{|\Omega|}, \tag{2.20}$$

and

$$\langle\rho^2\rangle \geqslant \langle\rho^{-2}\rangle^{-1} \geqslant \frac{(\frac{1}{2}-\alpha)(1-\alpha)}{|\Omega|^2}, \tag{2.21}$$

with the usual definition

$$\langle\rho^k\rangle \equiv \frac{\int_0^\infty (F^2 + G^2)\rho^{k+2}\,d\rho}{\int_0^\infty (F^2 + G^2)\rho^2\,d\rho}, \tag{2.22}$$

having used the inequality

$$\langle\rho^k\rangle\cdot\langle\rho^{-k}\rangle \geqslant 1, \tag{2.23}$$

which follows immediately from Hölder's with $p = 2$, if both mean values are finite.

Thus, the mean radius and the mean square radius of the solutions are bounded below, which solves in some way the classical problem of the "shrinking" of the atom (see also Lieb, 1976).[8]

Equation (2.19) can be obtained using Vazquez's extension of a theorem by Rosen (Rosen 1969, Vázquez, 1977)[9,10]

The stationary solutions of (2.4) allow the action to be not explicity time dependent so that the following global condition has to be satisfied:

$$\left[\frac{d}{d\sigma} I\{\psi(\sigma\mathbf{r},t),A_0(\sigma\mathbf{r})\}\right]_{\sigma=1} = 0, \tag{2.24}$$

having defined

$$I = \int_{\mathbb{R}^3} \mathscr{L}\,d^3\mathbf{r}. \tag{2.25}$$

It is easy to check that

$$I_\sigma = \frac{1}{\sigma^2} I_{D1} + \frac{1}{\sigma^3} I_{D2} + \frac{1}{\sigma^3} I_{\rm INT} + \frac{1}{\sigma^3} I_{\rm NL}, \tag{2.26}$$

having defined

$$I_{D1} = \int_{\mathbb{R}^3} \frac{i}{2} (\bar\psi\gamma^k\partial_k\psi - (\partial_k)\bar\psi\gamma^k\psi)\, d^3\mathbf{r}, \tag{2.27}$$

$$I_{D2} = \int_{\mathbb{R}^3} (\omega\psi^*\psi - m\bar\psi\psi)\, d^3\mathbf{r}, \tag{2.28}$$

$$I_{\rm INT} = \int_{\mathbb{R}^3} - e\,\bar\psi\gamma^0\psi A_0\, d^3\mathbf{r}, \tag{2.29}$$

and

$$I_{\rm NL} = \int_{\mathbb{R}^3} \lambda\, (\bar\psi\psi)^2\, d^3\mathbf{r}, \tag{2.30}$$

from which (2.19) easily follows. Let us remark that this result does not depend on the choice for ρ_0 because the function $V(\rho)$ belongs to the class $C^2(\mathbb{R}^3)$.

It is now possible to get the following expression of the energy:

$$E = \frac{2\pi}{|\lambda|m} \Big[\int_0^\infty (G^2 - F^2)\rho^2\, d\rho + \int_0^\infty (G^2 + F^2)V\rho^2\, d\rho + \mathrm{sig}\lambda \int_0^\infty (G^2 - F^2)^2\rho^2\, d\rho \Big], \tag{2.31}$$

which does not depend explicitly on the frequency, thus stressing that in this theory the energy is not obtained from the frequency as in the Dirac theory.

The spin can be easily obtained by integrating the corresponding density, which follows from the spatial part of the energy momentum tensor (Jauch–Rohrlich, 1977)[11]:

$$J_k = \tfrac{1}{2}\epsilon_{ijk}\, J^{ij}, \tag{2.32}$$

J^{ij} being

$$J^{ij} = \int_{\mathbb{R}^3} (x^i T^{j0} - x^j T^{i0})\, d^3\mathbf{r}, \tag{2.33}$$

which in our case leads to the values

$$J^1 = 0 = J^2$$

and

$$J^3 = \frac{\pi}{|\lambda|m^2} \int_0^\infty (G^2 + F^2)\rho^2\, d\rho = \tfrac{1}{2}N, \tag{2.34}$$

so that the normalization condition produces the correct values for both charge (2.7) and spin.

The magnetic moment of the electron is obtained by means of the integral

$$\mathbf{M} = \frac{1}{2}\int_{\mathbb{R}^3} \mathbf{r}\times\mathbf{j}_e\, d^3\mathbf{r}, \tag{2.35}$$

$\mathbf{j}_e$ being the spatial part of the electric four current.

In our case it is easy to obtain that

$$M^1 = 0 = M^2, \tag{2.36}$$

and

$$M^3 = -\frac{2\pi}{|\lambda|m^2}\cdot\frac{2e}{3m}\int_0^\infty FG\rho^3\, d\rho \quad (ns_{1/2}\ \text{states}) \tag{2.37}$$

and

$$M^3 = \frac{2\pi}{|\lambda|m^2}\cdot\frac{2e}{3m}\int_0^\infty FG\rho^3\, d\rho \quad (np_{1/2}\ \text{states}). \tag{2.38}$$

Lemma 3: The localized solutions of (2.10) for which the condition

$$\lim_{\rho\to\infty} \rho^3(F^2 + G^2) = 0 \tag{2.39}$$

holds verify the following global condition:

$$-\frac{3}{2}\int_0^\infty (F^2 + G^2)\rho^2\, d\rho + (1-\kappa)\int_0^\infty F^2\rho^2\, d\rho + (1+\kappa)\int_0^\infty G^2\rho^2\, d\rho - 2\int_0^\infty FG\rho^3\, d\rho - 2\,\mathrm{sig}\lambda \int_0^\infty (F^2 - G^2)FG\rho^3\, d\rho = 0. \tag{2.40}$$

Proof: The condition is easily obtained by direct integration of (2.40) as in the case of Lemma 2, in this case adding both equations and using as integration kernel. All the numerical solutions obtained in I verify the condition (2.39).

Let us remark that (2.40) does not depend upon the choice of the spherically symmetric external potential $V(\rho)$, which only has to give rise to the existence of localized solutions verifying (2.39).

The condition (2.40) can be used to evaluate the magnetic moment of the electron, which is

$$M^3 = -\left[1 - \frac{4}{3}\frac{\int_0^\infty F^2\rho^2\, d\rho}{\int_0^\infty (F^2+G^2)\rho^2\, d\rho} + \frac{4}{3}\,\mathrm{sig}\lambda\, \frac{\int_0^\infty (F^2 - G^2)FG\rho^3\, d\rho}{\int_0^\infty (F^2+G^2)\rho^2\, d\rho}\right] N\mu_B \quad (s\ \text{states}), \tag{2.41}$$

$$M^3 = -\left[\frac{1}{3} - \frac{4}{3}\frac{\int_0^\infty F^2\rho^2\, d\rho}{\int_0^\infty (F^2+G^2)\rho^2\, d\rho} + \frac{4}{3}\,\mathrm{sig}\lambda\, \frac{\int_0^\infty (F^2 - G^2)FG\rho^3\, d\rho}{\int_0^\infty (F^2+G^2)\rho^2\, d\rho}\right] N\mu_B \quad (p\ \text{states}), \tag{2.42}$$

For the case of the above mentioned numerical solutions the second and third factors are always of order α^2, or α^3, so that in the first approximation the correct Landé factors $g_s = 2$ and $g_p = \frac{2}{3}$ are obtained.

3. THE BIFURCATION OF THE SOLUTIONS

It was shown in I that there is a family of solutions for each bound state of the linear theory depending continuously on the frequency. It was found by numerical analysis that when $\Omega\to\Omega_l$ (the linear value of the frequency) the solution tends to zero in such a way that the norm is proportional to $\Omega - \Omega_l$. The nonlinear solution bifurcates from zero at $\Omega = \Omega_l$.

Let us now study this problem using the Ritz–Galerkin method (e.g., Rose, 1978),[12] an iterative method which gives an approximate solution as an expansion in the basis of eigenfunctions of the linear part of the differential operator.

In the case of stationary solutions, our equation can be written as

$$i\gamma^0\gamma^k\bar\partial_k\phi - \gamma^0\phi - V\phi + \mathrm{sig}\lambda\,(\bar\phi\phi)\gamma^0\phi + \Omega\phi = 0, \tag{3.1}$$

where $\phi = (2m\pi/|\lambda|)^{1/2}\psi$ is a dimensionless spinor and $\bar\partial_k$

$= (1/m)(\partial/\partial x^k)$. For simplicity in the following we will consider a point nucleus.

The field equation can be written as

$$-L\phi + N(\phi) + \Omega\phi = 0, \tag{3.2}$$

where L and N are the linear and the nonlinear parts of the differential operator

$$L = -i\gamma^0\gamma^k\partial_k + \gamma^0 + V, \tag{3.3a}$$

$$N(\phi) = \mathrm{sig}\lambda\,(\bar\phi\phi)\gamma^0\phi. \tag{3.3b}$$

L has a well known associate orthonormal set of eigenfunctions

$$\phi_n \pi\!\!\pi L\phi_n = \Omega_n\phi_n. \tag{3.4}$$

We make the expansion

$$\phi = \sum_{i=1}^{N} a_i\phi_i, \tag{3.5}$$

substitute it in (3.2), and project on ϕ_k, $k = 1,...,N$. In this way, we obtain a nonlinear system of algebraic equations which can be solved for a_k. Let us consider the ground state and take $N = 1$ as a first approximation. It is very easy to show that

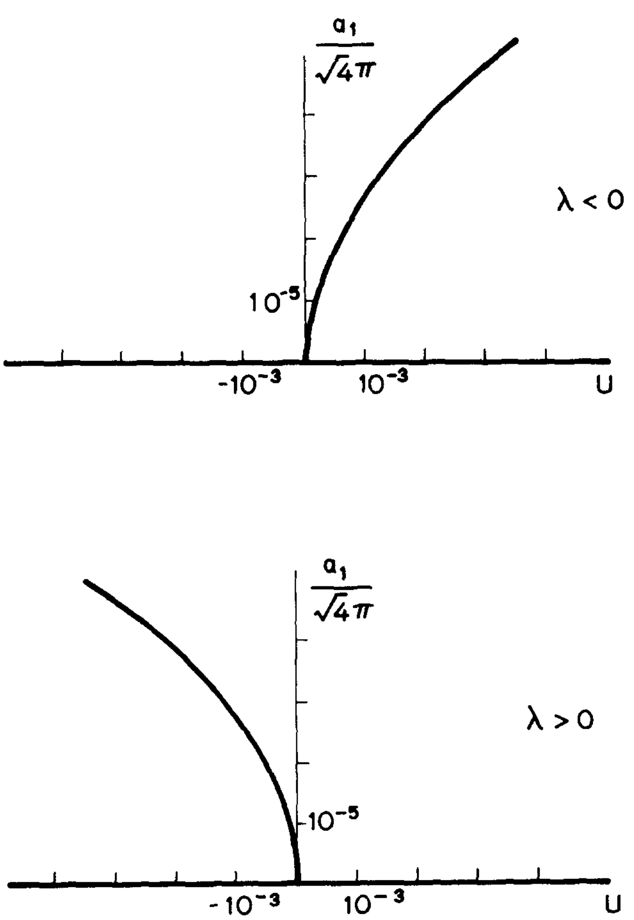

FIG. 1. Shape of the functions $a_1(U)$ for $\lambda < 0$ (a) and $\lambda > 0$ (b). The physical solution corresponds to $U = 0.5119\times 10^5\,\lambda\mathrm{m}^2$. Similar figures are obtained for the other states.

TABLE I. $a_1/\sqrt{4\pi}$ versus U near the linear limit of the ground state.

U	$a_1/\sqrt{4\pi}$ Eq. (3.8)	Numerical result
+0.001	0.278×10^{-4}	0.279×10^{-4}
+0.0005	0.197×10^{-4}	0.199×10^{-4}
+0.0003	0.152×10^{-4}	0.153×10^{-4}
+0.0001	0.881×10^{-5}	0.914×10^{-5}
+0.00005	0.623×10^{-5}	0.623×10^{-5}
Branching point		
−0.00005	0.623×10^{-5}	0.578×10^{-5}
−0.0001	0.881×10^{-5}	0.826×10^{-5}
−0.0003	0.152×10^{-4}	0.151×10^{-4}
−0.0005	0.197×10^{-4}	0.195×10^{-4}
−0.001	0.278×10^{-4}	0.278×10^{-4}

$$a_1^2 = \frac{U\alpha^6}{-\,\mathrm{sig}\lambda}\,\frac{1}{\int_{\mathbb{R}^3}(\bar\phi_1\phi_1)^2\,d^3\mathbf{p}}, \tag{3.6}$$

where U is given by

$$U = \frac{\Omega - \Omega_1}{\alpha^6}. \tag{3.7}$$

As a first consequence there is a solution for a given Ω only, if $\lambda U < 0$, a result obtained numerically in I.

It has an evident physical interpretation because $\lambda > 0$ gives rise to an attractive potential which lowers the frequency of the bound states (García and Usón, 1979).[13]

If this condition is verified, (3.6) can be written as

$$a_1 = 1.7632534\times 10^{-3}\sqrt{|U|}, \tag{3.8}$$

which coincides up to six figures with the results of I (Fig. 1, Table I). To check if $N = 1$ gives a good approximation let us consider $N = 2$. This gives two equations for a_1, a_2 from which it can be shown that if $|U| < 10^5$, then

$$\frac{a_2}{a_1} \simeq 2.28\times 10^{-9}U. \tag{3.9}$$

As in the physically interesting solutions $U < 1$, the effect of the second wave can be neglected.

The norm and energy of the solutions are then

$$N = \frac{2\pi}{|\lambda|m^2}a_1^2 = \frac{2\pi}{|\lambda|m^2}(3.109063\times 10^{-6})|U|, \tag{3.10}$$

$$E = \frac{2\pi}{|\lambda|m}\left(\Omega a_1^2 + \frac{1}{8\pi}a_1^4\int(\bar\phi_1\phi_1)^2\,d^3\mathbf{p}\right), \tag{3.11}$$

which can be written as

$$E = N\Omega m(1 + \epsilon), \tag{3.12}$$

where $\epsilon \simeq 10^{-13}|U|$. The correction can certainly be neglected if $|U| < 1$. In this case the energy has the value

$$E = N\Omega_1 m + \Delta E, \qquad \Delta E = NmU\alpha^6. \tag{3.13}$$

It is clear that for fixed λ the energy and the norm go to zero if $\Omega \to \Omega_1$. This bifurcation phenomenon is represented in Fig. 1 in two diagrams (a_1, U), one for each sign of λ.

As explained in I, the norm must adjust itself to 1. In that case, it follows from (3.10) that

$$U = -\frac{\lambda m^2}{1.9535\times 10^{-5}}. \qquad (3.14)$$

If $|\lambda|m^2 < 10^{-5}$, then $\Delta E < 2m\alpha^6$ and $U < 0.5$ and the preceding approximations are valid.

The same results are obtained from the other states.

In the preceding considerations the value of λ remained constant. The linear limit should be obtained for $\lambda \to 0$. However, (12) and (14) seem to indicate that in this limit everything blows up, the energy and the norm becoming infinite. However, if we let $\lambda \to 0$ and $\Omega \to \Omega_l$ taking into account (3.14) which implies that λ/U has to be kept constant, the Dirac solutions and energies are obtained in the limit. This is the correct way of obtaining the linear limit.

In conclusion, the numerical results of I are in perfect agreement with those obtained by means of the Ritz–Galerkin Method.

ACKNOWLEDGMENTS

We wish to thank Professor L. Vázquez for many very useful discussions. One of us (J.M.U.) thanks the Juan March Foundation for financial support. This work has been supported in part by Junta de Energía Nuclear, Madrid-3.

[1]A.F. Rañada, Int. J. Theor. Phys. **16**, 795 (1977).
[2]H. Weyl, Phys. Rev. **77**, 699 (1950).
[3]M. Soler, Phys. Rev. D **1**, 2766 (1970).
[4]A.F. Rañada, M.F. Rañada, M. Soler and L. Vázquez, Phys. Rev. D **10**, 517 (1974).
[5]A.F. Rañada and L. Vázquez, Prog. Theor. Phys. **56**, 311 (1976).
[6]J.J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, Mass., 1973).
[7]L. Vázquez, J. Phys. A **10**, 1361 (1977).
[8]E.H. Lieb, Rev. Mod. Phys. **48**, 553 (1976).
[9]G. Rosen, J. Appl. Math. **27**, 133 (1969).
[10]L. Vázquez, J. Math. Phys. **18**, 1343 (1977).
[11]J.M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Springer, Berlin, 1977).
[12]G. Rosen, Bull. Math. Biol. **40**, 853 (1978).
[13]L. García and J.M. Usón, "A simple model for Poincaré self stresses," Found. Phys. **10**, 137 (1980).

An application of the maximum principle to the study of essential self-adjointness of Dirac operators. II

J. J. Landgren
Department of Mathematics, University of Tennessee, Chattanooga, Tennessee 37402
P. A. Rejto[a)]
School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455
Martin Klaus[b)]
Department of Physics, Princeton University, Princeton, New Jersey 08540

(Received 6 December 1978; accepted for publication 23 March 1979)

In the present part II of this paper we prove the remaining part of the essential self-adjointness theorem stated in part I. This proof makes essential use of the maximum principle.

INTRODUCTION

In the present part II of this paper we prove Theorem 4.2 which was stated in part I and which was needed to prove the main Theorem 2.1. For convenience we continue the numbering of sections, however, we start anew the numbering of references.

In Sec. 6 we isolate two consequences of the maximum principle.[1] We start this isolation by formulating the positivity requirements (6.3) and (6.4) on a given potential. These requirements were introduced by Friedrichs[2,3] and Sears.[4] In Lemma 6.1 we consider a second order differential inequality corresponding to such a potential. Then we introduce a class of Cauchy data for which this differential inequality has no $\mathfrak{L}_2$-solutions. The nonexistence of $\mathfrak{L}_2$-solutions of the corresponding differential equation is the key fact in the proof of the Dunford-Schwartz version [5f] of the Friedrichs–Sears discreteness criterion. The application of the maximum principle allows us to replace this differential equation by the differential inequality of Lemma 6.1. It is a consequence of Lemma 6.1 that zero is in the resolvent set of the corresponding operator. This is the statement of Corollary 6.1. Lemma 6.2 states that if p and q are two potentials which satisfy the assumptions of Corollary 6.1, and if $p \leqslant q$, then the negatives of the corresponding kernels of the inverse satisfy the same inequality. Moreover, these kernels are positive. Lemma 6.2 is, essentially, an extended version of a remark of Faris.[6] It is a consequence of the minimum principle[1] which, in turn, is a simple consequence of the maximum principle.[1a]

In Sec. 7 we prove Theorem 4.2. We start this proof by introducing the scalar valued potentials $p_{\pm}(\kappa)$. Here, as in part I, κ denotes the reducing subspace parameter. In Theorem 7.1 we estimate the adjusted resolvent of Theorem 4.2 in terms of adjusted resolvents of the operators $L(p_{\pm}(\kappa))$. To prove Theorem 7.1 we state Lemma 7.1. It is a refinement of the well known fact that the square of the free particle Dirac operator equals the negative Laplacian plus the identity.[7,8,9a] Lemma 7.1 is a refinement inasmuch as it states the corresponding fact for the parts of these operators. It says that $L(P(0,\kappa))^2+I$ is the orthogonal sum of $L(p_+(\kappa))$ and $L(p_-(\kappa))$. A simple formal computation shows that this implies a corresponding orthogonal decomposition for the product of the adjoint of the resolvent and the resolvent of the operator $L(P(0,\kappa))$. To make this formal computation rigorous we need Corollary 6.1 and an abstract fact which is stated in Lemma 7.2. We complete the proof of Theorem 7.1 by combining these facts. Theorem 7.2 states that the norms of the adjusted resolvents of the operators $L(p_{\pm}(\kappa))$ are small for large $|\kappa|$. One might be tempted to prove Theorem 7.2 by introducing the comparison potentials: $q_{\pm}(\kappa) = p_{\pm}(\kappa) - 1$, and apply Lemma 6.2 to them. However, zero is in the continuous spectrum of $L(q_{\pm}(\kappa))$ and so these potentials violate the assumptions of Lemma 6.2. Nevertheless, using the usual construction for the Green's function, [10b] we can define a comparison kernel. In Lemma 7.3 we show that this comparison kernel majorizes negative of the resolvent kernel of the operator $L(p_{\pm}(\kappa))$. Note that by Lemma 6.2 this resolvent kernel is positive. Since the adjusted resolvent kernel is also positive it is majorized by the adjusted comparison kernel. In Lemma 7.4 we show that the adjusted comparison kernel does define a bounded operator and that the norm of this operator is small for large $|\kappa|$. This completes the proof of Theorem 7.2. Then inserting Theorem 7.2 in Theorem 7.1 proves Theorem 4.1 of part I. This, in turn, completes the proof of the main Theorem 2.1 of part I.

6. TWO CONSEQUENCES OF THE MAXIMUM PRINCIPLE FOR SECOND ORDER OPERATORS WITH POSITIVE POTENTIALS

In this section we isolate two consequences of the maximum principle for second order ordinary differential operators with positive potentials. For this purpose we need some assumptions.

First let p be a given real valued function satisfying the basic assumption:

p is piecewise continuous on $\mathscr{R}_+$ and bounded on each compact subinterval of $\mathscr{R}_+$. (6.1)

We shall also refer to such a function as a potential. With the aid of such a potential and the operators of definitions (4.1) and (4.2) of part I we define the operator $L(p)$ to be the closure of

$$L(p) = -D^2 + M(p) \quad \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+). \tag{6.2}$$

[a)]Supported by NSF Grant MCS78-02199.
[b)]Author of Appendix.

Note that the basic assumption (4.1) ensures that $L(p)$ maps $\mathfrak{C}_0^\infty(\mathscr{R}_+)$ into $\mathfrak{L}_2(\mathscr{R}_+)$ and that this operator admits a closure. Also note that replacing the real valued potential p by the matrix valued potential P, the resulting operator is different from the operator of definition (4.4). To emphasize this difference we shall always denote real valued potentials by lower case italic letters. Secondly we introduce two positivity requirements on p. We assume that

$$\inf_{\xi\in\mathscr{I}} p(\xi)>0, \textit{ for each compact subinterval } \mathscr{I} \text{ of } \mathscr{R}_+, \tag{6.3}$$

and that there is a point ξ_2 such that

$$p(\xi)\geqslant 3/4\xi^2 \quad \text{for} \quad \xi\in(0,\xi_2). \tag{6.4}$$

Lemma 6.1: *Let the potential p satisfy assumptions* (6.1), (6.3), *and* (6.4). *Suppose that the function u is such that*

u is two times piecewise continuously differentiable on $\mathscr{R}_+$, (6.5)

and

$$-u''+pu\leqslant 0. \tag{6.6}$$

Suppose further that there is a point ξ_1 *such that*

$$u(\xi_1)\geqslant 0 \quad \textit{and} \quad u'(\xi_1)<0. \tag{6.7}$$

Then

$$u(\xi)>0 \quad \textit{for} \quad \xi\in(0,\xi_1), \tag{6.8}$$

and

$$u\notin\mathfrak{L}_2(\mathscr{R}_+). \tag{6.9}$$

We prove conclusion (6.8) indirectly. Accordingly we assume that there is a point ξ_0 such that

$$\xi_0\in(0,\xi_1) \quad \textit{and} \quad u(\xi_0)\leqslant 0. \tag{6.10}$$

Assumptions (6.7) and (6.10) together show that u attains its maximum in the interior of the open interval (ξ_0,ξ_1). According to assumption (6.3), the potential p is positive; hence the maximum principle [1a] holds for the solutions of inequality (6.6). Therefore we can conclude [1a] that u is a constant on the interval $[\xi_0,\xi_1]$. This clearly contradicts assumption (6.7) on the Cauchy data. This contradiction shows that the indirect assumption (6.10) is false and completes the proof of conclusion (6.8).

To prove conclusion (6.9) first we define the comparison potential,

$$q(\xi)=\min\left(p(\xi),\frac{3}{4},\frac{1}{\xi^2}\right). \tag{6.11}$$

Then we define v to be the solution of the initial value problem

$$v(\xi_1)=u(\xi_1) \quad \text{and} \quad v'(\xi_1)=\tfrac{1}{2}u'(\xi_1), \tag{6.12}$$

and

$$-v''(\xi)+q(\xi)v(\xi)=0 \quad \text{for } \xi\in(0,\xi_1). \tag{6.13}$$

Note that according to definition (6.12) the Cauchy data of v satisfy assumption (6.7). Also note that according to definition (6.11) the potential q is positive. Since the proof of conclusion (6.8) used only the positivity of the potential and the Cauchy data of assumption (6.7), these two facts together with equation (6.13) allow us to repeat the proof of conclusion (6.8) for the function v. This positivity, in turn, together with inequality (6.6) and equation (6.13), allows us to obtain the differential inequality for the difference,

$$-(u-v)''(\xi)+p(\xi)(u-v)(\xi)<0 \quad \text{for} \quad \xi\in(0,\xi_1). \tag{6.14}$$

We see from the Cauchy data (6.12) and from assumption (6.7) that the difference $(u-v)$ again satisfies this assumption. Hence we can apply conclusion (6.8) to it, which yields

$$u(\xi)>v(\xi) \quad \text{for} \quad \xi\in(0,\xi_1). \tag{6.15}$$

It is not difficult to solve the comparison equation (6.13) on the interval $(0,\min(\xi_1,\xi_2))$ where $q(\xi)=3/4\xi^2$. We find that there are constants $\gamma_\pm$ such that

$$v(\xi)=\gamma_+\xi^{3/2}+\gamma_-\xi^{-1/2}. \tag{6.16}$$

The maximum principle [1a] together with the Cauchy data (6.12) and assumption (6.7) shows that at each point ξ_3 of the interval $(0,\min(\xi_1,\xi_2))$

$$v(\xi_3)>0 \quad \text{and} \quad v'(\xi_3)\leqslant 0. \tag{6.17}$$

Inserting relation (6.17) in formula (6.16) we find that γ_- is strictly positive. Hence v is not in $\mathfrak{L}_2(0,\min(\xi_1,\xi_2))$, and because of inequality (6.15) neither is u. This proves conclusion (6.9) and completes the proof of Lemma 6.1.

It is a consequence of Lemma 6.1 that under an additional assumption, the operator $L(p)$ is invertible. Since we shall make essential use of this fact in the proof of Theorem 4.2, we formulate it as a corollary.

Corollary 6.1: *Suppose that the potential p satisfies the assumptions of* Lemma 6.1 *and in addition*

$$\inf_{\xi\in\mathscr{R}_+} p(\xi)\neq 0. \tag{6.18}$$

Then the closure of the operator of $L(p)$ *is invertible in* $\mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_+))$ *so that*

$$0\in\rho(L(p)). \tag{6.19}$$

To prove this corollary first we claim that this operator is one to one. To see this note that according to assumption (6.18) the equation,

$$\gamma=\inf_{\xi\in\mathscr{R}_+} p(\xi), \tag{6.20}$$

defines a strictly positive constant. Inserting this fact in definition (6.2) shows that the quadratic form of this operator is bounded below by γ. In other words,

$$L(p)\geqslant\gamma I>0 \quad \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+). \tag{6.21}$$

It is an elementary consequence of relation (6.21) that the closure of this operator is one to one, proving our first claim.

Secondly we claim that this operator is onto,

$$\mathscr{R}(L(p))=\mathfrak{L}_2(\mathscr{R}_+). \tag{6.22}$$

It follows from relation (6.21) that this range is closed. Hence it equals the orthocomplement of the null space of its adjoint. [5b] Next we show that this null space is trivial. In other words,

$$f^*\in\mathfrak{D}(L(p))^* \quad \text{and} \quad (L(p))^*f^*=0 \tag{6.23}$$

implies

$$f^* = 0. \tag{6.24}$$

We prove relation (6.24) indirectly. Accordingly we assume that $f^* \neq 0$, and show that this contradicts Lemma 6.1. Since the potential satisfies the basic assumption (6.1) we can apply the Weyl lemma [5d,11] to the equation of assumption (6.23). This shows that the function f^* is a pointwise solution of the equation

$$-(f^*)''(\xi) + p(\xi) f^*(\xi) = 0, \quad \xi \in \mathscr{R}_+. \tag{6.25}$$

Since the coefficients of this equation are real, the real and imaginary parts of f^* each satisfy Eq. (6.25). Hence there is no loss of generality to assume that f^* is real. Because $f^* \in \mathfrak{L}_2(\mathscr{R}_+)$, and $f^* \not\equiv 0$ the derivative of $(f^*)^2$ cannot be nonnegative everywhere. Therefore there is a point ξ_1, where $f^*(\xi_1)$ and $(f^*)'(\xi_1)$ have opposite signs. Hence either f^* or $-f^*$ satisfies assumption (6.7) of Lemma 6.1. So we can conclude from Lemma 6.1 that f^* is not in $\mathfrak{L}_2(\mathscr{R}_+)$. This contradicts the first half of assumption (6.23). This contradiction, in turn, completes the proof of relation (6.24) and hence of relation (6.22).

Thus the operator $L(p)$ is one to one and onto; hence it has an algebraic inverse. Clearly the algebraic inverse of a closed operator is closed. This fact allows us to apply the closed graph theorem[10] to this algebraic inverse and conclude that it is bounded. This completes the proof of conclusion (6.20) and of Corollary 6.1.

As is well known, Lemma 6.1 implies that this inverse is an integral operator. [10e] In the following lemma we estimate its kernel.

Lemma 6.2: *Suppose that the potential p satisfies the basic assumption* (6.1) *and the positivity assumption* (6.3). *Suppose further that the closure of the corresponding operator of definition* (6.2) *is such that*

$$0 \in \rho(L(p)). \tag{6.26}$$

Then the kernel of this inverse is positive, that is to say

$$L(p)^{-1}(\xi,\eta) \geqslant 0. \tag{6.27}$$

If in addition q is another potential which satisfies assumptions (6.1), (6.3) *and* (6.26) *and*

$$q \leqslant p, \tag{6.28}$$

then

$$L(p)^{-1}(\xi,\eta) \leqslant L(q)^{-1}(\xi,\eta). \tag{6.29}$$

We prove conclusion (6.27) indirectly. Accordingly we assume that there is a point (ξ_0, η_0) such that

$$L(p)^{-1}(\xi_0,\eta_0) < 0, \tag{6.30}$$

and show that this contradicts the minimum principle. [1b]

The Weyl representation theorem for the resolvent kernel [5e] shows that this Green's function is continuous. Hence the point η_0 has a neighborhood $\mathscr{N}$ such that

$$\eta \in \mathscr{N} \quad \text{implies} \quad L(p)^{-1}(\xi_0,\eta) < 0. \tag{6.31}$$

Let f be a function such that

$$f \in \mathfrak{C}_0(\mathscr{R}_+) \quad f \geqslant 0 \quad \operatorname{supp} f \subset \mathscr{N}, \tag{6.32}$$

and set

$$g = L(p)^{-1} f. \tag{6.33}$$

Then clearly

$$g(\xi_0) < 0. \tag{6.34}$$

We show that there is another point ξ_1 in the open axis $\mathscr{R}_+$, such that

$$\min_{\xi \in \mathscr{R}} g(\xi) = g(\xi_1) < 0. \tag{6.35}$$

To see this we note that definition (6.33) and assumption (6.26) togethe yield,

$$g \in \mathfrak{D}(L(p)) \subset \mathfrak{L}_2(\mathscr{R}_+). \tag{6.36}$$

This, in turn, yields

$$\liminf_{\xi \to \infty} |g(\xi)| = 0. \tag{6.37^{r}}$$

We claim that

$$g \in \mathfrak{C}(\mathscr{R}_+), \tag{6.38}$$

so that

$$\lim_{\xi \to 0} g(\xi) = 0. \tag{6.37^{l}}$$

To see this recall that we defined $\mathfrak{D}(L(p))$ by closure. Hence assumption (6.36) implies that there is a sequence $\{g_n\}$ such that

$$g_n \in \mathfrak{C}_0^\infty(\mathscr{R}_+) \quad \text{and} \quad g = \lim_{n \to \infty} g_n, \tag{6.39}$$

and $\{L(p) g_n\}$ is a Cauchy sequence. In view of the Schwarz inequality these two relations yield $\sup_n (L(p)g_n, g_n) < \infty$. Inserting this relation and the positivity assumption (6.3) in definition (6.2) we obtain

$$\sup_n \|Dg_n\| < \infty. \tag{6.40}$$

Another application of the Schwarz inequality yields,

$$|g_n(\xi_2) - g_n(\xi_1)| \leqslant |\xi_2 - \xi_1|^{1/2} \|Dg_n\|.$$

Thus this sequence of functions in equicontinuous. At the same time it follows that this sequence is uniformly bounded on the interval [0,1]. Hence, according to the Arzela–Ascoli compactness criterion [5a] there is a subsequence which converges uniformly over this closed interval. This proves relation (6.38). At the same time remembering relation (6.39) it follows that

$$g(0) = \lim_{n \to \infty} g_n(0) = 0.$$

This proves relation (6.37)^{l}. Then relations (6.37)^{l,r}, (6.38), and (6.34) together prove relation (6.35). Next we show that the minimum principle holds for g. To see this note that definitions (6.32) and (6.33) together yield

$$L(p) g = f.$$

Inserting definition (6.32) and relation (6.38) in this equation we find that

$$g \in \mathfrak{C}^2(\mathscr{R}_+). \tag{6.41}$$

At the same time we find that

$$L(p) g(\xi) \geqslant 0 \quad \text{for} \quad \xi \in \mathscr{R}_+. \tag{6.42}$$

The positivity assumption (6.3) and relation (6.41) together, allow us to apply the minimum principle [1b] to the solution of the differential inequality (6.42). We conclude from this principle and from relation (6.35) that g is a constant. Because of relation $(6.37)^1$ this constant has to equal zero. This contradicts the indirect assumption (6.30). This contradiction, in turn, proves conclusion (6.27).

We complete the proof of Lemma 6.2 by showing that conclusion (6.27) implies conclusion (6.29). For this purpose we need the following version of the second resolvent equation,[12]

$$L(q)^{-1} - L(p)^{-1} = L(q)^{-1} M(p-q) L(p)^{-1}, \quad \text{on } L(p)\mathfrak{C}_0^\infty(\mathscr{R}_+). \tag{6.43}$$

First we note that according to conclusion (6.27) the first and third factors on the right of relation (6.43) map positive functions into positive functions. Second, we note that according to assumption (6.28) this also holds for the second factor. Therefore, the left member of relation (6.43) maps each positive function in $L(p)\mathfrak{C}_0^\infty(\mathscr{R}_+)$ into another positive function. Since assumption (6.26) implies that this set is dense,the validity of conclusion (6.29) follows. This completes the proof of Lemma 6.2.

7. PROOF OF THEOREM 4.2

In this section we prove Theorem 4.2. We start this proof by formulating another theorem. For this purpose we define two potentials for each integer κ by setting

$$p_\pm(\kappa)(\xi) = \frac{\kappa(\kappa \pm 1)}{\xi^2} + 1. \tag{7.1}$$

Clarly these potentials satisfy the basic assumption (6.1). Hence the corresponding operator of definition (6.2) admits a closure which we denote by $L(p_\pm(\kappa))$ again. It is also clear that for $|\kappa| \geqslant 2$, these potentials satisfy assumption (6.4) and (6.18). Hence we can conclude from Corollary 6.1 that

$$0 \in \rho[L(p_\pm(\kappa))] \quad \text{and so} \quad -1 \in \rho[L(p_\pm(\kappa))]. \tag{7.2}$$

The following theorem estimates the norm of the adjusted resolvent of the operator $L(P(0,\kappa))$ in terms of the norms of the adjusted resolvents of the operators $L(p_\pm(\kappa))$.

Theorem 7.1: *Let $L(P(0,\kappa))$ and $L(p_\pm(\kappa))$ be the closures of the operators of definition* (4.4), (4.7) *and* (6.2), (7.1), *respectively. Then for each integer κ, with $|\kappa| \geqslant 2$*

$$\|M^{-1} R(0, L(P(0,\kappa)))\|^2 \leqslant \max_{+,-} \{ M^{-1} R(-1, L(p_\pm(\kappa))) M^{-1} \| \}. \tag{7.3}$$

As a first step in the proof of Theorem 7.1 we formulate a lemma. To do this recall definition (4.6) of part I. It shows that

$$\sigma(C_\infty) = \{+1, -1\}. \tag{7.4}$$

Next let $\mathscr{C}_\pm$ denote the corresponding eigenspaces. In other words,

$$x_\pm \in \mathscr{C}_\pm \quad \text{implies} \quad C_\infty x_\pm = \pm x_\pm. \tag{7.5}$$

Lemma 7.1: *The operators of* Theorem 7.1 *are such that for each integer κ,*

$$L(P(0,\kappa))^2 = \begin{cases} L(p_+(\kappa)) & \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+, \mathscr{C}_+), \\ L(p_-(\kappa)) & \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+, \mathscr{C}_-). \end{cases} \tag{7.6}$$

To prove Lemma 7.1 recall definition (4.7) of part I. It shows that

$$\begin{aligned} L(P(0,\kappa))^2 &= JDJD + M(P(0,\kappa))JD \\ &\quad + JDM(P(0,\kappa)) + M(P^2(0,\kappa)) \\ &\qquad \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+, \mathscr{C}_2). \end{aligned} \tag{7.7}$$

The product rule of differentiation yields,

$$DM(P(0,\kappa)) = M(P(0,\kappa))D + M(P'(0,\kappa)) \quad \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+, \mathscr{C}_2). \tag{7.8}$$

Inserting definition (4.3) and relation (7.8) into relation (7.7) we obtain

$$\begin{aligned} L(P(0,\kappa))^2 &= -D^2 + [M(P(0,\kappa))J + JM(P(0,\kappa))]D \\ &\quad + M(JP'(0,\kappa)) + M(P^2(0,\kappa)) \\ &\qquad \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+, \mathscr{C}_2). \end{aligned} \tag{7.9}$$

Next we claim that

$$P(0,\kappa)J + JP(0,\kappa) = 0. \tag{7.10}$$

To see this first we note that setting $e = 0$ in formula (5.11) and remembering definitions (4.2), (4.5), and (4.6) one finds

$$JC_0(0,\kappa) = -\kappa C_\infty. \tag{7.11}$$

Since according to these definitions $C_0(0,\kappa)$ and C_∞ are symmetric and J is skew symmetric, this yields

$$C_0(0,\kappa)J + JC_0(0,\kappa) = 0. \tag*{$(7.12)_{0,J}$}$$

Second we note that formula (5.12) and definitions (4.2), (4.6), and (4.5) together show that

$$JC_\infty = C_0(0,1).$$

Hence another application of these definitions yields

$$C_\infty J + JC_\infty = 0. \tag*{$(7.12)_{\infty,J}$}$$

Relations $(7.11)_{0,J}$, $(7.11)_{\infty,J}$ and definition (4.7) together prove relation (7.10). Inserting relation (7.10) in relation (7.9) we find

$$L(P(0,\kappa))^2 = -D^2 + M(JP'(0,\kappa) + P^2(0,\kappa)). \tag{7.13}$$

Definitions (4.5), (4.6), and (4.2) together show that

$$C_0(0,\kappa)C_\infty = \kappa J.$$

Hence another application of these definitions yields

$$C_0(0,\kappa)C_\infty + C_\infty C_0(0,\kappa) = 0. \tag*{$(7.12)_{0,\infty}$}$$

Combining relations (7.11), $(7.12)_{0,\infty}$ and definition (4.7) with

$$C_0^2(0,\kappa) = \kappa^2 I \quad \text{and} \quad C_\infty^2 = I,$$

we find

$$JP'(0,\kappa)(\xi) + P^2(0,\kappa)(\xi) = (1 + \kappa^2/\xi^2)I + (\kappa/\xi^2)C_\infty. \tag{7.14}$$

Combining relation (7.14) in turn, with definitions (7.1) and (7.5) we find

$$JP'(0,\kappa) + P^2(0,\kappa) = \begin{cases} p_+(\kappa) & \text{on} \quad \mathscr{C}_+, \\ p_-(\kappa) & \text{on} \quad \mathscr{C}_-. \end{cases} \tag{7.15}$$

Inserting relation (7.15) in relation (7.13) and using definition (6.2) proves conclusion (7.6). This, in turn, completes the proof of Lemma 7.1.

Since the operator $L(P(0,\kappa))$ is symmetric we see from definitions (4.4) and (4.7) of part I that

$$[iI-L(P(0,\kappa))]^*[iI-L(P(0,\kappa))] = I + L(P(0,\kappa))^2 \quad \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2).$$

On the other hand, we see from Lemma 7.1 that

$$I + L(P(0,\kappa))^2 = [I + (p_+(\kappa))] \oplus [I + L(p_-(\kappa))]$$
$$\text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_+) = \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_+) \oplus \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_-).$$

Combining these two relations, we find

$$[iI-L(P(0,\kappa))]^*[iI-L(P(0,\kappa))] = [I+L(p_+(\kappa))] \oplus [I+L(p_-(\kappa))]. \qquad (7.16)$$

As is well known,[13] if each term of an orthogonal sum of two operators is invertible, then so is the sum, and its inverse equals the orthogonal sum of the inverses. Clearly the closure of an orthogonal sum equals the orthogonal sum of the closures. Since we denote an operator and its closure by the same symbol, these facts and relation (7.2) allow us to conclude from relation (7.16) that

$$-\{[iI-L(P(0,\kappa))]^*[iI-L(P(0,\kappa))]\}^{-1} = R(-1,L(p_+(\kappa))) \oplus R(i,L(p_-(\kappa)))^*. \qquad (7.17)$$

As a second step in the proof of Theorem 7.1 we show that

$$-\{[iI-L(P(0,\kappa))]^*[iI-L(P(0,\kappa))]\}^{-1} = R(i,L(P(0,\kappa)))R(i,L(P(0,\kappa)))^*. \qquad (7.18)$$

To see this, we formulate an abstract lemma. In this particular lemma we denote the closure of a given operator T by $\overline{T}$.

Lemma 7.2: *Let the operator T on $\mathfrak{D}(T)$ in $\mathfrak{H}$ be closable. Suppose that*

$$0 \in \rho(\overline{T}), \qquad (7.19)$$

and that

$$0 \in \rho(\overline{T^*T}). \qquad (7.20)$$

Then

$$(\overline{T^*T})^{-1} = (\overline{T})^{-1}(T^*)^{-1}. \qquad (7.21)$$

To prove this lemma note that as is well known,[5j] assumption (7.19) implies that

$$0 \in \rho(T^*). \qquad (7.22)$$

Hence each of the three operators in conclusion (7.21) is in $\mathfrak{B}(\mathfrak{H})$. Thus it suffices to show that the right member extends the left member. In other words, denoting the graph of a given operator B by $\Gamma(B)$, suffices to show that

$$\Gamma((\overline{T^*T})^{-1}) \subset \Gamma((\overline{T})^{-1}(T^*)^{-1}). \qquad (7.23)$$

To prove this inclusion, assume that

$$(f,g) \subset \Gamma((\overline{T^*T})^{-1}). \qquad (7.24)$$

Then we see that $\overline{T^*T}\,g = f$. By definition this implies that there is a sequence $\{g_n\}$ such that

$$g_n \in \mathfrak{D}(T^*T), \quad n = 1,2,\cdots, \qquad (7.25)$$

and

$$\lim_{n\to\infty} g_n = g \quad \text{and} \quad \lim_{n\to\infty} T^*Tg_n = f. \qquad (7.26)$$

Because of relation (7.22), relation (7.25) yields

$$(T^*)^{-1}T^*Tg_n = Tg_n.$$

Because of assumption (7.19) this, in turn, yields

$$(\overline{T})^{-1}(T^*)^{-1}T^*Tg_n = g_n.$$

Inserting relation (7.26) in this relation we obtain

$$(\overline{T})^{-1}(T^*)^{-1}f = g.$$

Hence assumption (7.24) implies that

$$(f,g) \subset \Gamma((\overline{T})^{-1}(T^*)^{-1}).$$

This, in turn, implies inclusion (7.23) and completes and proof of Lemma 7.2. It is an interesting fact, observed by Klaus, that assumption (7.19) implies assumption (7.20). Since for our operator both of these assumptions will hold we shall not use this abstract fact.

The Rellich–Weidmann relation (4.2) of part I shows that the operator

$$T = iI - L(P(0,\kappa)) \quad \text{on} \quad \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2), \qquad (7.27)$$

satisfies assumption (7.19). Relation (7.17) shows that this operator also satisfies assumption (7.20). Therefore applying the abstract Lemma 7.2 to this operator proves relation (7.18). Multiplying both sides of relation (7.18) by M^{-1} yields

$$-M^{-1}R(i,L(P(0,\kappa)))R(i,L(P(0,\kappa))^*M^{-1} = M^{-1}R(-1,L(p_+(\kappa)))M^{-1} \oplus M^{-1} \times R(-1,L(p_-(\kappa)))M^{-1}. \qquad (7.28)$$

It is well known that the norm of the product of the adjoint of an operator and the operator equals the square of the norm of the operator.[5g] Since the norm of an operator equals the norm of its adjoint,[5j] for each operator B in $\mathfrak{B}(\mathfrak{H})$,

$$\|BB^*\| = \|B\|^2.$$

Applying this abstract relation to the operator

$$B = M^{-1}R(i,L(P(0,\kappa))),$$

we find

$$\|M^{-1}R(i,L(P(0,\kappa)))R(i,L(P(0,\kappa)))^*M^{-1}\| = \|M^{-1}R(i,L(P(0,\kappa)))\|^2. \qquad (7.29)$$

It is also well known that the norm of an orthogonal sum of two operators equals the maximum of the norms.[14] Hence

$$\|M^{-1}R(-1,L(p_-(\kappa)))M^{-1} \oplus M^{-1}R(-1,L(p_+(\kappa)))M^{-1}\| = \max_{\pm}\|M^{-1}R(-1,L(p_\pm(\kappa)))M^{-1}\|. \qquad (7.30)$$

Inserting relations (7.29) and (7.30) in relation (7.28) we arrive at the validity of conclusion (7.3). This completes the proof of Theorem 7.1.

In the following theorem we show that the right member of conclusion (7.3) tends to zero as $|\kappa|$ tends to infinity.

Theorem 7.2: *The operators $L(p_\pm(\kappa))$ of* Theorem 7.1 *are such that*

$$\lim_{|\kappa|\to\infty} \|M^{-1}R(-1,L(p_\pm(\kappa)))M^{-1}\| = 0. \qquad (7.31)$$

One might be tempted to prove this theorem by introducing the comparison potentials

$$q_{\pm}(\kappa)(\xi) = \kappa(\kappa \pm 1)/\xi^2, \tag{7.32}$$

and apply Lemma 6.2. However, these potentials do not satisfy assumption (6.26) since 0 is in the continuous spectrum of the operator $L(q_{\pm}(\kappa))$. Nevertheless we can use the usual construction for the Green's function,[5e,10b,15] to define a comparison kernel. More specifically we define two kernels $G_{\pm}(\kappa)(\xi,\eta)$ for each integer κ by the requirements that

$$\left(-\frac{d^2}{d\eta^2} + q_{\pm}(\kappa)(\eta)\right)G_{\pm}(\kappa)(\xi,\eta) = -\delta(\xi-\eta), \tag{7.33}$$

$$\lim_{\eta\to 0} G_{\pm}(\kappa)(\xi,\eta) = 0, \tag{7.34}$$

and

$$\lim_{\eta\to\infty} G_{\pm}(\kappa)(\xi,\eta) = 0. \tag{7.34$^{\rm r}$}$$

Note that the existence of such a kernel is not evident. Later we shall give it explicitly. In the following lemma we show that such a kernel majorizes the negative of the resolvent kernel of the operator $L(p_{\pm}(\kappa))$.

Lemma 7.3: *For each integer κ let the potentials $p_{\pm}(\kappa)$ be defined by relation* (7.1) *and let the kernels $G_{\pm}(\kappa)(\xi,\eta)$ be defined by relations* (7.33) *and* (7.34)$^{\rm l,r}$. *Then*

$$0 \leqslant -R(-1,L(p_{\pm}(\kappa)))(\xi,\eta) \leqslant G_{\pm}(\kappa)(\xi,\eta). \tag{7.35}$$

To prove the first inequality of conclusion (7.35) we note that relation (7.2) and definition (7.1) allow us to apply Lemma 6.2 to the operator $L(p_{\pm}(\kappa))$. Then conclusion (6.26) yields this inequality.

To prove the second inequality of conclusion (7.35) first we note that this resolvent kernel satisfies the differential equation

$$\left(-\frac{d^2}{d\eta^2} + p_{\pm}(\kappa)(\eta) + 1\right)R(-1,L(p_{\pm}(\kappa)))(\xi,\eta) = \delta(\xi-\eta), \tag{7.36}$$

for each ξ in in $\mathscr{R}_+$.[5c,10b] Second, we note that as is well known[5k]

$$R(-1,L(p_{\pm}(\kappa)))(\xi;\cdot)\in\mathfrak{L}_2(\mathscr{R}_+). \tag{7.37}$$

Next let $k_{\pm}^{\rm l,r}(\kappa)$ be two solutions of the homogeneous differential equation

$$\left(-\frac{d^2}{d\eta^2} + p_{\pm}(\kappa)(\eta) + 1\right)k_{\pm}^{\rm l,r}(\kappa)(\eta) = 0,$$

such that

$$k_{\pm}^{\rm l}(\kappa)\in\mathfrak{L}_2(0,1) \quad \text{and} \quad k_{\pm}^{\rm r}(\kappa)\in\mathfrak{L}_2(1,\infty).$$

Elementary algebra shows that then

$$\lim_{\eta\to 0} k_{\pm}^{\rm l}(\kappa)(\eta) = \lim_{\eta\to\infty} k_{\pm}^{\rm r}(\kappa)(\eta) = 0.$$

Inserting these facts in relations (7.36) and (7.37) yields for each ξ in $\mathscr{R}_+$

$$\lim_{\eta\to 0} R(-1,L(p_{\pm}(\kappa)))(\xi,\eta) = 0 \tag{7.38$^{\rm l}$}$$

and

$$\lim_{\eta\to\infty} R(-1,L(p_{\pm}(\kappa)))(\xi,\eta) = 0. \tag{7.38$^{\rm r}$}$$

To complete the proof of conclusion (7.35) we introduce a notation for this difference. Since this difference involves the negative of this resolvent kernel, this yields,

$$g_{\xi}(\eta) = G_{+}(\kappa)(\xi;\eta) + R(-1,L(p_{+}(\kappa)))(\xi;\eta). \tag{7.39}$$

In terms of this notation the second inequality of conclusion (7.35) reads,

$$g_{\xi}(\eta)\geqslant 0, \quad \eta\in\mathscr{R}_+. \tag{7.40}$$

We prove relation (7.40) indirectly. Accordingly we assume that there is a point η_0 such that

$$g_{\xi}(\eta_0) < 0, \quad \eta_0\in\mathscr{R}_+, \tag{7.41}$$

and show that this contradicts the minimum principle.[1b] To see this first we show that there is another point η_1 in the open axis $\mathscr{R}_+$ such that

$$\lim_{\eta\in\mathscr{R}_+} g_{\xi}(\eta) = g_{\xi}(\eta_1) < 0. \tag{7.42}$$

To see this, in turn, subtract relations (7.38)$^{\rm l,r}$ from definitions (7.34)$^{\rm l,r}$. This yields that the function of definition (7.39) also satisfies these boundary conditions:

$$\lim_{\eta\to 0} g_{\xi}(\eta) = 0, \tag{7.43$^{\rm l}$}$$

and

$$\lim_{\eta\to\infty} g_{\xi}(\eta) = 0. \tag{7.43$^{\rm r}$}$$

The Weyl formula for the resolvent kernel[5c,15] shows that the second term of definition (7.39) is continuous. Since the one-dimensional δ function is the second derivative of a continuous function[10a] we see from definition (7.33) that the first term in definition (7.39) is continuous. Therefore,

$$g_{\xi}\in\mathfrak{C}(\mathscr{R}_+). \tag{7.44}$$

Relations (7.44), (7.43)$^{\rm l,r}$, and the indirect assumption (7.41), together prove relation (7.42). Second, we show that for this function the minimum principle holds. To see this insert definition (7.32) in definition (7.1). This yields

$$p_{\pm}(\kappa) = q_{\pm}(\kappa) + 1.$$

Inserting this relation in the differential equation (7.36), adding the result to definition (7.33), and using the definition (7.39), we find

$$\left(-\frac{d^2}{d\eta^2} + q_{\pm}(\kappa)(\eta)\right)g_{\xi}(\eta) = -2R(-1,L(p_{\pm}(\kappa)))(\xi,\eta). \tag{7.45}$$

We have already seen that the right member of this equation is continuous. Combining this fact with relation (7.44) and definition (7.32) yields

$$g_{\xi}\in\mathfrak{C}^2(\mathscr{R}_+). \tag{7.46}$$

According to the already established first inequality of conclusion (7.35) this right member is also positive. Hence

$$\left(-\frac{d^2}{d\eta^2} + q_{\pm}(\kappa)(\eta)\right)g_{\xi}(\eta)\geqslant 0. \tag{7.47}$$

According to definition (7.32) the potential $q_{\pm}(\kappa)$ is positive and bounded on every compact subinterval of $\mathscr{R}_{+}$. This fact and relation (7.46) allows us to apply the minimum principle to the differential inequality (7.47). From this principle and from relation (7.42) we conclude that g_{ξ} is a constant. Because of relation (7.43)[1] this constant has to equal zero. This contradicts the indirect assumption (7.41). This contradiction, in turn, proves relation (7.40) and completes the proof of Lemma 7.3.

We return to the proof of Theorem 7.2. Using the comparison kernel of Lemma 7.3 for each integer κ we define a pair of kernels by

$$K_{\pm}(\kappa)(\xi,\eta) = \xi^{-1}G_{\pm}(\kappa)(\xi,\eta)\eta^{-1}. \qquad (7.48)_{\pm}$$

Then we see from Lemma 7.3 that

$$0 \leqslant -\xi^{-1}R(-1,L(p_{\pm}(\kappa)))(\xi;\eta)\eta^{-1} \leqslant K_{\pm}(\kappa)(\xi;\eta).$$

In other words the kernel of the operator of Theorem 7.2 is negative and its absolute value is majorized by the kernel of definition $(7.48)_{\pm}$.

Hence Theorem 7.2 is implied by the following lemma.

Lemma 7.4: For each integer κ let $K_{\pm}(\kappa)$ be the operator corresponding to the kernel of definition $(7.48)_{\pm}$. Then for $|\kappa| \geqslant 2$

$$K_{\pm}(\kappa) \in \mathfrak{B}(\mathfrak{L}_2(\mathscr{R}_{+})), \qquad (7.49)_{\pm}$$

and

$$\lim_{|\kappa| \to \infty} \|K_{\pm}(\kappa)\| = 0. \qquad (7.50)$$

For brevity we prove this lemma for positive integers and for the operator $K_{-}(\kappa)$ only. Then the usual construction for the Green's function[10b] yields,

$$G_{-}(\kappa)(\xi,\eta) = \frac{1}{-2\kappa+1}\begin{cases}\xi^{-\kappa+1}\eta^{\kappa}, & \eta < \xi,\\ \xi^{\kappa}\eta^{-\kappa+1}, & \eta > \xi.\end{cases}$$

Inserting this formula in definition $(7.48)_{-}$ we find that

$$\int_{\mathscr{R}_{+}} |K_{-}(\kappa)(\xi,\eta)|\, d\eta = \frac{1}{|2\kappa-1|}\left(\frac{1}{\kappa} + \frac{1}{\kappa-1}\right).$$

According to a result of Schur–Holmgren[16] this relation and the symmetry of this kernel implies conclusion $(7.49)_{-}$. At the same time it also implies that

$$\|K_{-}(\kappa)\| \leqslant \frac{1}{|2\kappa-1|}\left(\frac{1}{\kappa} + \frac{1}{\kappa-1}\right).$$

This estimate proves conclusion $(7.50)_{-}$ and completes the proof of Lemma 7.4. Lemma 7.4, in turn, completes the proof of Theorem 7.2.

Inserting Theorem 7.2 in Theorem 7.1 proves Theorem 4.2 of part I. Therefore the proof of the main Theorem 2.1 of part I is complete.

ACKNOWLEDGMENT

It is a pleasure to thank Professor Faris and Professor Spruck for valuable conversations. Special thanks are due to Professor Weinberger for his numerous suggestions concerning the maximum principle and to Professor Klaus for his elegant Appendix, which proves Theorem 7.2 via a Hardy type inequality.

APPENDIX: PROOF OF THEOREM 7.2 VIA A HARDY TYPE INEQUALITY, BY MARTIN KLAUS

We start this proof by defining a positive integer l by

$$l(l+1) = \kappa(\kappa+1) \quad \text{for } \kappa > 0, \qquad (\mathrm{A1})_{+}$$

and

$$l(l+1) = \kappa(\kappa-1) \quad \text{for } \kappa < 0. \qquad (\mathrm{A1})_{-}$$

Then the Hardy type inequality used by Schmincke[17-19] and definitions (6.2) and (7.1) together imply

$$L(p_{\pm}(\kappa)) + I \geqslant \tfrac{1}{4} + l(l+1)M^{-2} \quad \text{on } \mathfrak{C}_0^{\infty}(\mathscr{R}_{+}),$$
$$l = 0,1,\cdots. \qquad (\mathrm{A2})$$

Since

$$\tfrac{1}{4} + l(l+1) = (2l+1)^2/4$$

and the operator M^{-1} is symmetric, relation (A2), in turn, implies

$$([L(p_{\pm}(\kappa)) + 1]f,f) \geqslant \frac{(2l+1)^2}{4}(M^{-1}f, M^{-1}f). \qquad (\mathrm{A3})$$

Next we show that

$$0 \in \rho(L(p_{\pm}(\kappa) + I)) \text{ and hence } (L(p_{\pm}(\kappa)) + I)^{-1} \geqslant 0. \qquad (\mathrm{A4})$$

To see this relation we note the well-known fact[7b] that the restriction of the operator $-\Delta + I$ to the subspace of momentum l is unitarily equivalent to $(2l+1)$ copies of the operator $L(p_{\pm}(\kappa))$. More specifically this holds for the restriction of the operator $-\Delta + I$ to $\mathfrak{C}_0^{\infty}(\mathscr{R}_3\{0\})$. We know that this restriction is essentially self-adjoint.[9b]

We also know that for the closure of this operator,

$$0 \in \rho(-\Delta + I). \qquad (\mathrm{A5})$$

These facts together prove relation (A4). At the same time they show that

$$[L(p_{\pm}(\kappa)) + I]\mathfrak{C}_0^{\infty}(\mathscr{R}_{+}) \text{ is dense in } \mathfrak{L}_2(\mathscr{R}_{+}). \qquad (\mathrm{A6})$$

As a first consequence of relation (A4) we see that each of the two operators $L(p_{\pm}(\kappa)) + I$ admits a symmetric square root. Hence

$$(L(p_{\pm}(\kappa) + I)f,f)$$
$$= ([L(p_{\pm}(\kappa)) + I]^{1/2}f, [L(p_{\pm}(\kappa)) + I]^{1/2}f). \qquad (\mathrm{A7})$$

As a second consequence of relation (A4) we see that

$$0 \in \rho([L(p_{\pm}(\kappa)) + I]^{1/2}). \qquad (\mathrm{A8})$$

Hence defining

$$g = [L(p_{\pm}(\kappa)) + I]^{1/2}f,$$

relations (A3), (A7), and (A8) yield

$$\frac{4}{(2l+1)^2}(g,g) \geqslant (M^{-1}[L(p_{\pm}(\kappa)) + I]^{-1/2}g,$$
$$M^{-1}[L(p_{\pm}(\kappa)) + I]^{-1/2}g),$$
$$\text{for } g \in [L(p_{\pm}(\kappa)) + I]^{1/2}\mathfrak{C}_0^{\infty}(\mathscr{R}_{+}), \qquad (\mathrm{A9})$$

We know[9b] that a bounded operator which has a dense range maps a dense set onto another dense set. Therefore, relations

(A6) and (A8) imply that

$$[L(p_{\pm}(\kappa))+I]^{1/2}\mathfrak{C}_0^\infty(\mathscr{R}_+) \text{ is dense in } \mathfrak{L}_2(\mathscr{R}_+). \quad \text{(A10)}$$

Similarly to the concluding steps of Lemma 7.2 we see from relations (A9) and (A10) that

$$\|M^{-1}[L(p_{\pm}(\kappa)+I)^{-1}]M^{-1}\|\leqslant \frac{4}{(2l+1)^2}. \quad \text{(A11)}$$

Relation (A11) completes the proof of Theorem 7.2. Incidentally note that this estimate for the norm is sharp.

Notes added in proof:

Remark 1: According to a verbal communication of Behncke and Klaus *Theorem* 2.1 *is sharp* inasmuch as replacing the open interval by the closed interval in assumption (2.6) conclusion (2.9) need not hold.

Remark 2: *A weaker version of Lemma* 7.1 *suffices to prove Theorem* 4.2. To see this, insert relation (7.14) into relation (7.13). This yields

$$L(P(0,\kappa))^2$$
$$= -D^2+(\kappa^2I+\kappa C_\infty)M^{-2}+I, \quad \text{on } \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2).$$

As in Sec. 7, this, in turn, yields

$$[iI-L(P(0,\kappa))]^*[iI-L(P(0,\kappa))]$$
$$= -D^2+(\kappa^2I+\kappa C_\infty)M^{-2}+2I.$$

Since C_∞ is Hermitian, we see from relation (7.4) that

$$\kappa C_\infty \geqslant -|\kappa|I, \quad \text{on } \mathscr{C}_2.$$

Inserting this inequality into the previous one, we find

$$[iI-L(P(0,\kappa))]^*[iI-L(P(0,\kappa))]$$
$$\geqslant -D^2+[(|\kappa|-\tfrac{1}{2})^2-\tfrac{1}{4}]M^{-2}+2I.$$

Now we proceed as in the Klaus Appendix by inserting the Hardy inequality into this one. Then we find the key inequality

$$[iI-L(P(0,\kappa))]^*[iI-L(P(0,\kappa))]$$
$$\geqslant(|\kappa|-\tfrac{1}{2})^2M^{-2}, \quad \text{on } \mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2).$$

This inequility and relations (7.19) and (7.27) together imply that

$$\|g\|^2\geqslant(|\kappa|-\tfrac{1}{2})^2\|M^{-1}R(i,L(P(0,\kappa))g\|^2,$$
$$\text{for } g\in[iI-L(P(0,\kappa))]\mathfrak{C}_0^\infty(\mathscr{R}_+,\mathscr{C}_2),$$

and that this set is dense. Hence Theorem 4.2 follows by closure.

Remark 3: Note that the proof of Theorem 4.2 is based on the fact that *the larger the reducing subspace parameter* $|\kappa|$, *the larger the bound on e.* This remarkable fact was observed by Brownell in his Appendix to P.A. Rejto, "Some essentially self-adjoint one electron Dirac operators," Israel J. Math **9**, 144–171 (1971).

[1]M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations* (Prentice-Hall, New Jersey, 1967); a. Theorem 1.3; b. Apply Theorem 1.3 to $-u$.

[2]K.O. Friedrichs, *Criteria for the Discrete Character of the Spectra of Ordinary Differential Operators*, Courant Anniversary Volume (Interscience, New York, 1948), pp. 145–60.

[3]K.O. Friedrichs, "Criteria for discrete spectra," Commun. Pure Appl. Math. 439 (1950).

[4]D.B. Sears, "On the solutions of linear second order differential equation which are of integrable square," J. London Math. Soc. **24**, 201 (1949).

[5]N. Dunford and J.T. Schwartz, *Linear Operators Parts I and II*, 2nd ed. (Wiley, New York, 1963), a. Theorem IV.6.II; b. Lemma XII:1.6.(d); c. Corollary XII.4.13 and Theorem XII.4.18; d. Lemma XIII.2.9; e. Theorem XIII.3.16; f. Theorem XIII.6.23; g. Lemma IX.3.2; h. Theorem XII.4.19; i. Lemma VII.3.2; j. Lemma VI.2.10; k. Lemma XIII.3.4.

[6]W.G. Faris, "Self-Adjoint Operators," Springer Lect. Notes Math. **433**, (1975). See the reference to the maximum principle in the Notes to Sec. 9.

[7]A. Messiah, *Quantum Mechanics, Vols. I and II* (North-Holland, Amsterman, 1962), Chap. 1–21, a. Section XX.6; b. Sections IX.3 and IX.4.

[8]P.A.M. Dirac, *Principles of Quantum Mechanics* (Clarendon, Oxford, 1936).

[9]T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966), a. Relation (5.26); b. Subsection V.5.2; c. Problem III.2.9.

[10]B. Friedman, *Principles and Techniques of Applied Mathematics* (Wiley, New York, 1956), a. Equation (3.11); b. Equation (3.46).

[11]J. Hellwig, *Differential Operators of Mathematical Physics* (Addison-Wesley, Reading, Massachusetts, 1967), Sec. II.5.

[12]E. Hille and R.S. Phillips, *Functional Analysis and Semigroups*, Am. Math. Soc. Coll. XXXI (1957). See Subsection V.5.10.

[13]A.I. Plesner, *Spectral Theory of Linear Operators* (Ungar, New York, 1969), Vol. II, Eng. Transl. see Theorem 6.7.7.

[14]M.A. Naimark, *Normed Rings* (Noordhoff, Groningen, 1972). See Subsection I.5.15.

[15]H. Weyl, "Ober gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen," Math. Ann. **68**, 220 (1909).

[16]K.O. Friedrich, "Spectral Theory of Operators in Hilbert Space," Springer Lect. Notes Appl. Math. Sci. **9**, (1973). See Sec. 20.

[17]U.W. Schmincke, "Distinguished self-adjoint extensions of Dirac Operators," Math. Z. **129**, 335 (1972). See Lemma 6.

[18]H. Kalf, U.W. Schmincke, J. Walter, and R. Wüst, "On the Spectral Theory of Schrödinger and Dirac Operators with Strongly Singular Potentials," Springer Lect. Notes Math. **448**, (1975). See Lemma 1.

[19]A. Zygmund, *Trigonometric Series* (Cambridge U.P., Cambridge, 1959), Vol. 1. See Theorem 9.16.

Zero-mass behavior of Feynman amplitudes. I

Edward B. Manoukian
Department of National Defence, Royal Military College of Canada, Kingston, Ontario K7L 2W3

(Received 21 February 1979; accepted for publication 24 May 1979)

The zero-mass behavior of Feynman amplitudes with or without subtractions is investigated. Rules are given to determine the behavior of any Feynman amplitude, in the most general cases, when some or all of the masses in the underlying theory become small, in general, at different rates in Euclidean space. That we have to consider the approach of such masses to zero at different rates as well is clearly essential on physical grounds. Examples are worked out as illustrations for the rules. A general rule is also given which gives a sufficiency condition to guarantee the existence of the zero-mass behavior of *any* renormalized Feynman amplitude under the above same conditions which reduces to an elementary and direct inspection of the corresponding Feynman diagram.

I. INTRODUCTION

The purpose of this work is to investigate, in Euclidean space, the explicit zero-mass behavior of any Feynman amplitude with or without subtractions, where in the original Feynman amplitude all the masses in the underlying theory are nonzero and the subtractions are all performed at the origin directly on the Feynman integrand. The study is general enough to deal with the most general cases when some (not necessarily all) of the masses involved in the theory become small, in general, at different rates. That we have to consider the most general cases when (i) some of the masses as well and not necessarily all the masses become small and (ii) consider the approach of such masses to zero at different rates as well is clearly a physical requirement. To avoid problems with theories for which some of the masses are *a priori* zero (at least on experimental grounds), we initially write for the denominators of the propagators, corresponding to a pole term Q^2, $(Q^2+\mu^2)$ in the Feynman rules. The limit of such masses $\mu\to 0$ and other masses not *a priori* zero is rigorously studied with the approach to zero of these masses, necessarily at different rates.

In quantum electrodynamics, for example, one would be interested in determining the limit of a renormalized Feynman amplitude for $\mu\to 0$, $m\to 0$, *and* $(\mu/m)\to 0$, which is a nontrivial limit, where μ is a photon mass as defined above and m is the mass of the electron. To our knowledge all previous studies of the *behavior* of Feynman amplitudes at the zero-mass limit in Euclidean space either have not taken, explicitly, into account the complex and nontrivial role of subtractions or have considered the case only when all the masses in the theory approach zero and all at the same rate and were carried out by different methods.[1,2] Finally, we also give a general rule as a sufficiency condition for the existence of the zero-mass limit of any renormalized Feynman amplitude which reduces to an elementary and direct inspection of the corresponding Feynman diagram. The analysis is carried out with the external Euclidean momenta of a proper and connected Feynman graph under investigation being nonexceptional, i.e., no partial sums of the external momenta vanish, and with all the subtractions carried out at the origin directly on the Feynman integrand. The degree of divergence of a subdiagram is chosen to coincide with its dimensionality.

II. ZERO-MASS BEHAVIOR

The structure of a Feynman amplitude A, with or without subtractions, is of the form

$$A(p_1,\ldots,p_{4m};\mu_1,\mu_2,\ldots,\mu_\rho)$$
$$=\int dk_1\cdots dk_{4n}\,R(p_1,\ldots,p_{4m};k_1,\ldots,k_{4n};\mu_1,\mu_2,\ldots,\mu_\rho),\quad(1)$$

where

$$R(p_1,p_2,\ldots,p_{4m};k_1,k_2,\ldots,k_{4n};\mu_1,\mu_2,\ldots,\mu_\rho)$$
$$=\sum\frac{A^{i}_{m_{ij},\ldots,n_{it},\ldots,\tau_{in}}}{\Pi_l(Q_l^2+\mu_l^2)^{\sigma_l}}\prod_j(p_j)^{m_{ij}}\prod_t(k_t)^{n_{it}}\prod_n(\mu_n)^{\tau_{in}},\quad(2)$$

$\sigma_l>0$ and the sum is over all nonnegative integers m_{ij}, n_{it}, τ_{in}, and i; and

$$Q_l=\sum_j a_j^l p_j+\sum_j b_j^l k_j,\quad(3)$$

with $\{p_1,p_2,\ldots,p_{4m}\}$ and $\{k_1,k_2,\ldots,k_{4n}\}$ representing the sets of the external and internal independent momentum components, respectively, of a proper and connected graph G with which A is associated. The $A^{i}_{m_{ij},\ldots,n_{it},\ldots,\tau_{in}}$'s are suitable coefficients. $\{\mu_1,\mu_2,\ldots,\mu_\rho\}$ denotes the set of the masses in the theory appearing in the graph G ($\mu_n\neq 0$, $n=1,2,\ldots,\rho$). All subtractions in (1) are supposed to have been carried out at the origin as defined[3] directly on the unrenormalized Feynman integrand.

We scale the masses in an arbitrary subset $\{\mu_1,\mu_2,\ldots,\mu_s\}\subseteq\{\mu_1,\mu_2,\ldots,\mu_\rho\}$ of the masses as follows:

$$\begin{aligned}&\mu_1\to\lambda_1\mu_1,\\&\mu_2\to\lambda_1\lambda_2\mu_2,\\&\vdots\\&\mu_s\to\lambda_1\lambda_2\cdots\lambda_s\mu_s.\end{aligned}\quad(4)$$

The masses in the subset $\{\mu_1,\mu_2,\ldots,\mu_s\}$ have been arbitrarily labeled from 1 to s for convenience. Without loss of generality we may assume that any set of masses which approach zero at the same rate have been identified with either μ_1, or

 0022-2488/80/051218-06$01.00

$\mu_2, \ldots,$ or μ_s, depending on the rate at which we wish them to vanish in comparison to the other vanishing ones. Under the scaling (4), each term in the sum of the numerator of R in (2) is transformed as follows:

$$\prod_j (p_j)^{m_{ij}} \prod_t (k_t)^{n_{it}} (\mu_1)^{\tau_{i1}} (\mu_2)^{\tau_{i2}} \cdots (\mu_s)^{\tau_{is}} \cdots (\mu_\rho)^{\tau_{i\rho}}$$
$$\rightarrow (\lambda_1 \lambda_2 \cdots \lambda_s)^{d(N)} \prod_j (p'_j)^{m_{ij}} \prod_t (k'_t)^{n_{it}} \prod_n (\mu'_n)^{\tau_{in}},$$

where

$$\begin{aligned} p'_j &= p_j/\lambda_1 \lambda_2 \cdots \lambda_s, \quad k'_t = k_t/\lambda_1 \lambda_2 \cdots \lambda_s, \\ \mu'_1 &= \mu_1/\lambda_2 \lambda_3 \cdots \lambda_s, \\ \mu'_2 &= \mu_2/\lambda_3 \lambda_4 \cdots \lambda_s, \\ &\vdots \\ \mu'_{s-1} &= \mu_{s-1}/\lambda_s, \\ \mu'_s &= \mu_s, \\ \mu'_i &= \mu_i/\lambda_1 \lambda_2 \cdots \lambda_s, \quad \text{for } s < i \leqslant \rho, \end{aligned} \tag{5}$$

where $d(N)$ is the dimensionality of the numerator defined by

$$d(N) = (m_{i1} + m_{i2} + \cdots) + (n_{i1} + n_{i2} + \cdots) + (\tau_{i1} + \tau_{i2} + \cdots + \tau_{i\rho}), \tag{6}$$

and is a fixed number for all i for which $A^i_{m_{i1}, \ldots, n_{i1}, \ldots, \tau_{i1}} \cdots \neq 0$. Similarly, the denominator in Eq. (2) is transformed as follows:

$$\prod_l (Q_l^2 + \mu_l^2)^{\sigma_l} \rightarrow (\lambda_1 \lambda_2 \cdots \lambda_s)^{d(D)} \prod_l (Q_l'^2 + \mu_l'^2)^{\sigma_l}, \tag{7}$$

where

$$Q'_l = Q_l/\lambda_1 \lambda_2 \cdots \lambda_s,$$

and the μ'_l's are defined in Eq. (5). $d(D)$ is the dimensionality of the denominator, i.e., $d(D) = 2\Sigma_l \sigma_l$. Accordingly, the (renormalized) integrand R is transformed to

$$(\lambda_1 \lambda_2 \cdots \lambda_s)^{d(R)} R(p'_1, \ldots, p'_{4m}, k'_1, \ldots, k'_{4n}, \mu'_1, \mu'_2, \ldots, \mu'_s, \mu'_{s+1}, \ldots, \mu'_\rho), \tag{8}$$

where $d(R) = D(N) - d(D)$. Hence, finally the amplitude A is transformed to

$$\begin{aligned} &A(p_1, p_2, \ldots, p_{4m}; \lambda_1 \mu_1, \lambda_1 \lambda_2 \mu_2, \ldots, \lambda_1 \lambda_2 \cdots \lambda_s \mu_s, \mu_{s+1}, \ldots, \mu_\rho) \\ &= (\lambda_1 \lambda_2 \cdots \lambda_s)^{d(G)} A\left(\frac{p_1}{\lambda_1 \cdots \lambda_s}, \ldots, \frac{p_{4m}}{\lambda_1 \cdots \lambda_s}; \frac{\mu_1}{\lambda_2 \cdots \lambda_s}, \frac{\mu_2}{\lambda_3 \cdots \lambda_s}, \ldots, \frac{\mu_{s-1}}{\lambda_s}, \mu_s, \frac{\mu_{s+1}}{\lambda_1 \cdots \lambda_s}, \ldots, \frac{\mu_\rho}{\lambda_1 \cdots \lambda_s} \right) \\ &\equiv (\lambda_1 \lambda_2 \cdots \lambda_s)^{d(G)} A', \end{aligned} \tag{9}$$

where $d(G) = (d(N) - 2\Sigma_l \sigma_l + 4n)$, which is the dimensionality of the graph G, by definition.

Let the $4n$ integration variables, the $4m$ external momentum components[4] and, for convenience, the masses $\mu_1, \mu_2, \ldots, \mu_\rho$ [1] be considered as the components of a vector $\mathbf{P}$ in an $(4(n+m)+\rho)$ Euclidean space $R^{(4(n+m)+\rho)}$. For each line l in G with momentum–mass $(Q_l^1, \ldots, Q_l^4, \mu_l) \equiv \overline{Q}_l$, we introduce a vector $\mathbf{V}_l$ such that $\mathbf{V}_l \cdot \mathbf{P} = \overline{Q}_l$ for each of the five components of $\overline{Q}_l$. Let I be an arbitrary subspace[4] of $R^{(4(n+m)+\rho)}$ spanned by arbitrarily chosen $4n$ vectors associated with the $4n$ integration variables. We introduce E as the complement of I in $R^{(4(n+m)+\rho)\,4}$ as $R^{(4(n+m)+\rho)} = I + E$, with $\Lambda(I)$ as the operator of projection onto E along I. (See Ref. 4 for additional definitions.) In particular, we note that $\Sigma_{\mu=1}^5 (\mathbf{V}_l^\mu \cdot \mathbf{P})(\mathbf{V}_l^\mu \cdot \mathbf{P}) = Q_l^2 + \mu_l^2$. Let $S_i \subseteq E$ be a subspace spanned by the independent vectors $\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_i$, where $\mathbf{L}_1$ is with nonvanishing components $p_1, p_2, \ldots, p_{4m}, \mu_{s+1}, \ldots, \mu_\rho$; $\mathbf{L}_2$ with nonvanishing component μ_1; $\mathbf{L}_3$ with nonvanishing component $\mu_2, \ldots,$ and $\mathbf{L}_i$ with nonvanishing component μ_{i-1}.

To find the behavior of A' with respect to λ_i when $1/\lambda_i \rightarrow \infty$ we consider the following:

Let $T = \{G'_i, G''_i, \cdots\}$ be the totality of all subdiagrams $\subseteq G$ (with may include G itself if applicable) *such that*:

(i) All the masses in the lines of G/G'_i, $G/G''_i, \cdots$ (for those not empty) are in the set $\{\mu_i, \mu_{i+1}, \ldots, \mu_s\}$. G/G'_i is by definition the graph G with G'_i shrunk to a point, i.e., $I_G = I_{G/G'_i} I_{G'_i}$, where I_G denotes the unrenormalized Feynman integrand of G. Also, by definition, we have $I_{G/G} = 1$.

(ii) $G'_i, G''_i, \cdots$ contain all the vertices to which the external momenta (the lines) to the graph G join, i.e., they contain all the *external vertices* (though not necessarily all the lines) of G at which the external momenta of G flow. If there is an internal vertex in G to which is attached some of the lines in G'_1, not carrying any external momenta and not forming closed loops, then, necessarily, the masses carried by these lines must be from the set $\{\mu_1, \ldots, \mu_{i-1}, \mu_{s+1}, \ldots, \mu_\rho\}$.

(iii) $d(G'_i) = d(G''_i) = \cdots$.

(iv) All the lines in G'_i, $G''_i, \cdots$ have their $\mathbf{V}$'s not orthogonal to S'_1, $S''_1, \cdots$, respectively, with $\Lambda(I) S'_1 = S_i$, $\Lambda(I) S''_1 = S_i, \cdots$, where S'_1, $S''_1, \cdots$ are subspaces of $R^{[4(n+m)+\rho]}$ with which the subdiagrams $G'_i, G''_i, \cdots$, respectively, are associated[4] in a convenient notation.

(v) Any subdiagram G_0 which respects the conditions (i)–(iv) above is such that $d(G_0) \leqslant d(G'_i)$. If $d(G_0) = d(G'_i)$, *then* $G_0 \in T$ by definition of such a subdiagram G_0.

Let $\tau_0 = \{S'_1, S''_1, \cdots\}$ denote the totality of the subspaces with which the subdiagrams $G'_i, G''_i, \cdots$ are associated as defined through (i)–(v) above.

Consider the subdiagram G'_i. Let $g'_1, \ldots, g'_j, \cdots$ be the set of *all possible* proper but not necessarily connected subdiagrams of G'_i (including G'_i itself, if applicable) *such that*

(vi) Each proper and connected part of each of the subdiagrams $g'_1, \ldots, g'_j, \cdots$ is *divergent*. (By definition, a proper but disconnected subdiagram is meant that the number of its connected parts does not increase by cutting any one of its lines.)

(vii) All the masses in the lines in each of the subdiagrams $g'_1, \ldots, g'_j, \cdots$ are necessarily in the set $\{\mu_i, \mu_{i+1}, \ldots, \mu_s\}$.

(viii) Let $S'_2, \ldots, S'_{j+1}, \cdots$ be the subspaces associated with $(G'_i|g'_1), \cdots, (G'_i|g'_j), \cdots$, respectively, where $(G'_i|g'_1)$ denotes the subgraph of G'_i with $I_{g'_1}$ replaced by

$$V_{g'_1} = \prod_i V_{g'_{1i}} = \prod_i \left[(-T_{g'_{1i}}) \sum_{Z_i} \prod_{g \in Z_i} (-T_g) I_{g'_{1i}} \right], \tag{10}$$

where[3] the sum is over all sets Z_i of divergent proper and connected subdiagrams in g'_{1i} and where $g'_{11}, g'_{12},\cdots$ denote the (divergent) proper and connected parts of g'_1. $T_{g'_{11}}$, for example, denotes the Taylor operation in the external momenta of the subdiagram g'_{11} up to $d(g'_{11})$. [Note that $d(g'_1)\equiv\Sigma_i\, d(g'_{1i})$, by definition.] Hence, $I_{(G'_i|g'_1)}$ denotes $I_{(G'_i/g'_1)}$ multiplied by the generalized vertex as defined in Eq. (10) *as a function of the external momenta of* g'_1 *only*, i.e, the expression (10) is a polynomial of degree $d(g'_1)$ in the external variables of g'_1.

Effectively, the subdiagram g'_1 in G'_i has been replaced by a vertex parts $V_{g'_1}$ [Eq. (10)] with dimensionalities $d(g'_{11}), d(g'_{12}),\cdots$ and total dimensionality $d(g'_1)$. By momentum conservation, the external variables of g'_1, i.e., of $g'_{11}, g'_{12},\cdots$, will then be written as linear combinations of the external variables of G'_i and the internal variables of G'_i/g'_1 as usual. We remark that *if* the subdiagrams $g'_1,\ldots,g'_j,\cdots$ were defined by removing the restriction on their masses in their lines and the latter were also from the set $\{\mu_1,\ldots,\mu_{i-1}, \mu_{s+1},\ldots,\mu_\rho\}$, which are scaled by the parameter $1/\lambda_i$ according to the definition of the amplitude A', then the dimensionality of $V_{g'_1}$ with respect to $1/\lambda_i$ would have been

$$\dim_{1/\lambda_i} V_{g'_1}\leqslant d(g'_1) - \text{number of internal lines in } g'_1 \text{ with masses in the set } \{\mu_1,\ldots,\mu_{i-1},\mu_{s+1},\cdots\mu_\rho\}, \tag{11}$$

if at least some of the external momenta of g'_1 are scaled by $1/\lambda_i$ as well instead of the higher dimensionality of $d(g'_1)$. This shows that the dimensionality of $V_{g'_1}$, with respect to $1/\lambda_i$, would have been reduced from $d(g'_1)$ if any of its lines had masses from the set $\{\mu_1,\ldots,\mu_{i-1},\mu_{s+1},\ldots,\mu_\rho\}$. This is the reason for defining the subgraphs $g'_1,\ldots,g'_j,\cdots$ as in (vii) above with *all* their masses in the set $\{\mu_i,\ldots,\mu_{i+1},\ldots,\mu_s\}$.

Repeat the above procedure to construct the subdiagrams $g''_1,\ldots,g''_j,\ldots;\ldots$ and the subspaces $S''_2,\ldots,S''_{j+1},\ldots;\ldots$ by considering all the remaining subdiagrams $G''_i;\ldots$ in T. We denote the totality of the subspaces $S'_1, S'_2,\ldots,S'_{j+1},\ldots; S''_1, S''_2,\ldots;\ldots$ by τ. By definition, $\tau_0\subseteq\tau$. To recollect, we note that $S'_1, S'_2,\ldots; S''_1, S''_2,\ldots$ are the subspaces with which are associated the subdiagrams G'_i, $(G'_i|g'_1),\ldots; G''_i$, $(G''_i|g''_1)$, $\ldots;\ldots$, respectively, as defined through (i)–(viii) above with

$$\begin{aligned} &A(I)S'_1=S_i, && A(I)S'_2=S_i,\ldots,\\ &A(I)S''_1=S_i, && A(I)S''_2=S_i,\ldots,\\ &\cdots && \cdots \end{aligned}$$

Accordingly, we can state[5,4,6] the rules for determining the behavior of

$$A(p_1,\ldots,p_{4m};\lambda_1\mu_1,\lambda_1\lambda_2\mu_2,\ldots,\lambda_1\lambda_2\ldots\lambda_s\mu_s,\mu_{s+1},\ldots,\mu_\rho)$$

when $1/\lambda\, i\to\infty$ as follows in an elementary fashion:

Rules for determining the behavior of A for $\lambda_i\to 0$: Let $T=\{G'_i, G''_i,\ldots\}$ and $g'_1,\ldots,g'_j,\ldots;g''_1,\ldots;\ldots$ be as defined through (i)–(viii) above, and $\tau_0=\{S'_1,S''_1,\ldots\}$ and $\tau=\{S'_1,S'_2,\ldots;S''_1,S''_2,\ldots\}$ as defined below (v) and (viii) above, respectively. The asymptotic coefficients[5,4,6] of A' with respect to the parameters $1/\lambda_i$ are given through

$$\alpha_I(S_i)=d(G'_i) \tag{12}$$

and $\beta_I(S_i)$ is given through the following: Write I as a decomposition to, arbitrarily chosen and arbitrarily labeled, one-dimensional $4n$ disjoint subspaces $I=I_1+I_2+\cdots+I_{4n}$, associated with $4n$ integration variables. Choose arbitrarily a subspace S from the set $\tau_0=\{S'_1,S''_1,\cdots\}$. Let $S^1,S^2,\cdots$ be all those subspaces in τ *such that* for a fixed $j\in[1,\ldots,4n]$

$$\begin{aligned} A(I_j+\cdots+I_{4n})S^1 &= A(I_j+\cdots+I_{4n})S^2\\ &=\cdots=A(I_j+\cdots+I_{4n})S; \end{aligned} \tag{13}$$

then,

$$\beta_I(S_i)=\sum_{j=1}^{4n}p_j, \tag{14}$$

with $p_j=0$, if all the elements in

$$\begin{aligned}\{&\dim A(I_{j+1}+\cdots+I_{4n})S^1-\dim S_i, \dim A(I_{j+1}+\cdots\\ &+I_{4n})S^2-\dim S_i,\ldots,\dim A(I_{j+1}+\cdots+I_{4n})S-\dim S_i\}\end{aligned} \tag{15}$$

are equal, and $p_j=1$ otherwise for each $j\in[1,2,\ldots,4n]$. Here $\dim S$ denotes the dimension of the subspace S. Hence,[5,4,6]

$$\begin{aligned} &A(p_1,\ldots p_{4m};\lambda_1\mu_1,\ldots\lambda_1\ldots\lambda_s\mu_s,\mu_{s+1},\ldots,\mu_\rho)\\ &\quad = O\Big((\lambda_1)^{d(G/G'_1)}\ldots(\lambda_i)^{d(G/G'_i)}\ldots(\lambda_s)^{d(G/G'_s)}\\ &\quad\times\sum_{\gamma_1,\ldots,\gamma_s}\Big(\ln\frac{1}{\lambda_{\pi_1}}\Big)^{\gamma_1}\Big(\ln\frac{1}{\lambda_{\pi_2}}\Big)^{\gamma_2}\cdots\Big(\ln\frac{1}{\lambda_{\pi_s}}\Big)^{\gamma_s}\Big), \end{aligned} \tag{16}$$

where the sums range over all nonnegative integers $\gamma_1,\gamma_2,\ldots,\gamma_s$ satisfying

$$\begin{aligned} &\gamma_1\leqslant\beta(\mathbf{L}_1,\ldots,\mathbf{L}_{\pi_1}),\\ &\gamma_1+\gamma_2\leqslant\beta(\mathbf{L}_1,\ldots,\mathbf{L}_{\pi_2}),\\ &\vdots\\ &\gamma_1+\ldots+\gamma_s\leqslant\beta(\mathbf{L}_1,\ldots,\mathbf{L}_{\pi_s}), \end{aligned} \tag{17}$$

where the asymptotic coefficients are arranged in increasing order such that

$$\beta(\mathbf{L}_1,\ldots,\mathbf{L}_{\pi_1})\leqslant\beta(\mathbf{L}_1,\ldots,\mathbf{L}_{\pi_2})\leqslant\ldots\leqslant\beta(\mathbf{L}_1,\ldots,\mathbf{L}_{\pi_s}), \tag{18}$$

and $\pi_1,\ldots,\pi_2$ is a permutation of the integers $1,\ldots,s$.

In the next section examples will be worked out in detail which demonstrates the full applications of these rules. Finally, we wish to remark directly from Eq. (16) the following:

(1) If the set $\{g'_1,\ldots,;g''_1,\ldots\}$ is empty, then no logarithmic divergences $[(\ln 1/\lambda_i)^{\delta_i}, \delta_i>0]$ occurs in reference to the parameter λ_i when $\lambda_i\to 0$.

(2) If the set $\{g'_1,\ldots;g''_1,\ldots\}$ is empty and $d(G'_i)\leqslant d(G)$, then the limit of A exists (as opposed to the divergence of A) when $\lambda_i\to 0$ and with $\lambda_1,\ldots\lambda_{i-1},\lambda_{i+1},\ldots,\lambda_s$ held fixed and nonzero.

By repeating the statement (2) for each of the parameters λ_i, $i=1,2,\ldots,s$, we may readily infer the existence of the limit of A when $\lambda_1\to 0,\lambda_2\to 0,\ldots,\lambda_s\to 0$, independently. In Sec. IV we summarize this result which gives a sufficiency condition for the existence of any renormalized Feynman amplitude in the zero mass limit and as we see the rule reduces to an elementary and direct inspection of the graph G.

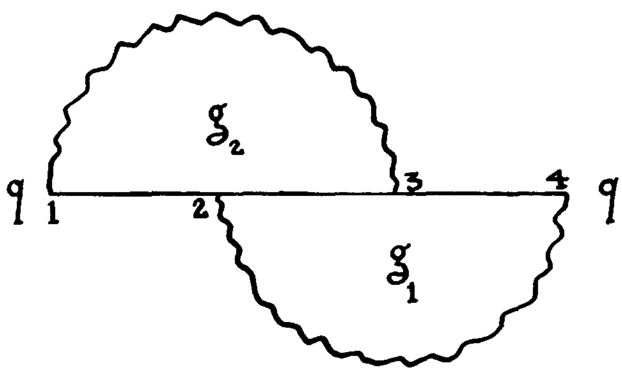

FIG. 1. A fourth order electron self-energy graph.

electrodynamics in Fig. 1, where the denominator Q^2 of the photon propagator has been written as $(Q^2+\mu^2)$, i.e.,

$$D^0_{\mu\nu}(Q)=\left(\delta_{\mu\nu}-G\frac{Q_\mu Q_\nu}{Q^2}\right)\frac{1}{(Q^2+\mu^2)},$$

in a covariant gauge with gauge constant G.

The momenta Q_{ij} flowing from the vertex i to j may be written as

$$\begin{aligned}
Q_{12} &= Ak_1+Ck_2+9/2,\\
Q_{13} &= -Ak_1-Ck_2+9/2,\\
Q_{32} &= -(A-B)k_1-(C-D)k_2,\\
Q_{34} &= -Bk_1-Dk_2+9/2,\\
Q_{24} &= Bk_1+Dk_2+9/2,
\end{aligned}\tag{19}$$

where Q_{13} and Q_{24} denote the photon momenta. In reference to the subgraphs g_1 and g_2, the internal (k^{g_1},k^{g_2}) and the external (q^{g_1},q^{g_2}) variables for g_1 and g_2, respectively, may be written

$$\begin{aligned}
k^{g_1}_{24} &= \left(B-\frac{A}{3}\right)k_1+\left(D-\frac{C}{3}\right)k_2,\\
k^{g_1}_{34} &= -\left(B-\frac{A}{3}\right)k_1-\left(D-\frac{C}{3}\right)k_2 \quad (=-k^{g_1}_{24}),\\
k^{g_1}_{32} &= \left(B-\frac{A}{3}\right)k_1+\left(D-\frac{C}{3}\right)k_2 \quad (=k^{g_1}_{24}),\\
k^{g_2}_{13} &= -\left(A-\frac{B}{3}\right)k_1-\left(C-\frac{D}{3}\right)k_2,\\
k^{g_2}_{12} &= \left(A-\frac{B}{3}\right)k_1+\left(C-\frac{D}{3}\right)k_2 \quad (=-k^{g_2}_{13}),\\
k^{g_2}_{32} &= -\left(A-\frac{B}{3}\right)k_1-\left(C-\frac{D}{3}\right)k_2 \quad (=k^{g_2}_{13}),
\end{aligned}\tag{20}$$

$$\begin{aligned}
q^{g_1}_{24} &= \frac{A}{3}k_1+\frac{C}{3}k_2+\frac{q}{2},\\
q^{g_1}_{34} &= -\frac{A}{3}k_1-\frac{C}{3}k_2+\frac{q}{2},\\
q^{g_1}_{32} &= -\frac{2A}{3}k_1-\frac{2C}{3}k_2,\\
q^{g_2}_{13} &= -\frac{B}{3}k_1-\frac{D}{3}k_2+\frac{q}{2},\\
q^{g_2}_{12} &= \frac{B}{3}k_1+\frac{D}{3}k_2+\frac{q}{2},\\
q^{g_2}_{32} &= \frac{2B}{3}k_1+\frac{2D}{3}k_2,
\end{aligned}\tag{21}$$

where a Taylor operation, for example T_{g_1}, is carried out with respect to the variables $q^{g_1}_{24},q^{g_1}_{34},q^{g_1}_{32}$ up to $d(g_1)$. Let $I_1+\ldots+I_4$ be arbitrarily associated with the variable $k^{g_1}_{24}$ and $I_5+\ldots+I_8$ with $k^{g_2}_{13}$, with nonvanishing Jacobian $J(k_1,k_2/k^{g_1}_{24},k^{g_2}_{13})$. The constants $A,B,C,$ and D in Eqs. (19) are arbitrary to the extent that $k^{g_i}_{rs}\neq 0, k^{G}_{rs}(=k_{rs})\neq 0$ for arbitrary k_1 and k_2, $i=1,2$ and $r,s=1,2,3,4$.

A. Example 1

Consider the behavior of the renormalized Feynman amplitude corresponding to the graph G in Fig. 1 for $\mu\to 0$ with $m\neq 0$ and $q\neq 0$ (in Euclidean space) and both fixed, where m is the mass of the electron. We scale the photon mass μ by λ. Accordingly, we have to consider first the behavior of the amplitude $A(q/\lambda,m/\lambda,\mu)$ in the limit $\lambda\to 0$. According to the rules (i)–(v) in Sec. II, we easily see that $T=\{G\}$, with $d(G)=1$, and G is associated with a subspace S' spanned by independent vectors $\mathbf{L}_1,\ldots,\mathbf{L}_8,\mathbf{L}_9$ and one representation of these vectors is $\mathbf{L}_1$ with nonvanishing component k^1_1, $\mathbf{L}_2$ with nonvanishing component $k^2_1,\ldots,$ and $\mathbf{L}_8$ with nonvanishing component k^4_2, and finally $\mathbf{L}_9$ with nonvanishing components $q^1,\ldots,q^4,m$. The subspace S', of course, does not depend on a particular representation. We note that $\Lambda(I)S'=\{\mathbf{L}_9\}\equiv S$.

G contains no divergent proper and connected subdiagram such that the mass in all its lines is μ [condition (vii) in the rules]. Hence, no logarithmic divergences occur in the parameter $1/\lambda$ and the limit $A(q,m,\lambda\mu)$ for $\lambda\to 0$ exists in Eq. (16) since $d(G/G)=0$, by definition. Technically, $\tau=\tau_0$, and Eqs. (13) and (15) shrink to $\Lambda(I_j+\ldots+I_8)S'$ and $\{\dim\Lambda(I_{j+1}+\ldots+I_8)S'-\dim S\}$, respectively, for all $j=1,2,\ldots,8$. Hence, $p_j=0$ in Eq. (14) for all $j=1,2,\ldots,8$ and $\beta_I(S)=0$. Accordingly

$$\begin{aligned}
A(q,m,\lambda\mu) &= (\lambda)^1 A(q/\lambda,m/\lambda,\mu)\\
&= O\left((\lambda)^1\left(\frac{1}{\lambda}\right)^1\left(\ln\frac{1}{\lambda}\right)^0\right)\\
&= O(1),
\end{aligned}\tag{22}$$

i.e., the limit of $A(q,m,\mu)$ exist for $\mu\to 0$ with $m\neq 0$ and $q\neq 0$ (in Euclidean space) and both fixed.

It is easy to see that the lowest order self-energy graph in Fig. 2 also exists in the limit $\mu\to 0$ under the same conditions as above. For the convenience of the reader we give the explicit expression for the renormalized amplitude corresponding to Fig. 2 with subtractions at the origin and $\mu=0$:

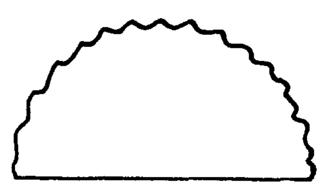

FIG. 2. Lowest order electron self-energy graph.

$$A(q,m,0) = -\frac{\alpha}{2\pi}\gamma_q \int_0^1 x\,dx \ln(1+\frac{q^2}{m^2}x) - \frac{m\alpha}{\pi}\left[\int_0^1 dx \ln(1+\frac{q^2}{m^2}x)\right], \qquad (23)$$

which is obviously well defined with $q^2>0, m\neq 0$, and both fixed. The general rule for the existence of the limit of *any* Feynman amplitude is given in a compact and elementary way in Sec. IV.

B. Example 2

Consider the behavior of the renormalized amplitude $A(q,m,\mu)$ corresponding to Fig. 1 for $m\to 0, \mu\to 0$, *and* $(\mu/m)\to 0$. In this case we have to consider first the behavior of $A(q/\lambda_1\lambda_2, m/\lambda_2, \mu)$ for $\lambda_1\to 0, \lambda_2\to 0$.

In reference to the parameter $1/\lambda_1$, we easily see from (i)-(v) in the rules that $T=\{G\}$, $d(G)=1$, and G is associated with the subspace $S_1' = \{\mathbf{L}_1, \mathbf{L}_2, ..., \mathbf{L}_8, \mathbf{L}_9\}$, where $\mathbf{L}_1, ..., \mathbf{L}_9$ are independent vectors, arbitrarily chosen, and $\mathbf{L}_1$ may be chosen with a nonvanishing component k_1^1, $\mathbf{L}_2$ with nonvanishing component k_1^2,..., and $\mathbf{L}_8$ with nonvanishing component k_2^4, and finally $\mathbf{L}_9$ with nonvanishing components $q^1,...,q^4$. Then $\Lambda(I)S_1' = S_1$, where $S_1 = \{\mathbf{L}_9\}$. As in example 1, we may associate $I_1 + ... + I_4$ with $k_{24}^{g_1}$ and $I_5 + ... + I_8$ with $k_{13}^{g_2}$, for convenience.

The graph G contains the following divergent subdiagrams which satisfy conditions (vi)–(viii), in reference to $1/\lambda_1: g_1, g_2$, and G itself, i.e., $\tau = \{S_1', S_2', S_3', S_4'\}$ with which are associated the diagrams in $\{G, (G|g_1), (G|g_2), (G|G)\}$, respectively. The subdiagram $(G|G)$ is a generalized vertex defined by[7]

$$(-T_G)[1 - T_{g_1} - T_{g_2}]I_G, \qquad (24)$$

in the external momentum q/λ_1, corresponding to A', and hence is just a polynomial of degree $d(G)$ in $(1/\lambda_1)$. The subspaces S_2', S_3', and S_4' may be easily defined as follows:

$$S_2' = \{\alpha_2\mathbf{L}_1 + \beta_2\mathbf{L}_5, \alpha_2\mathbf{L}_2 + \beta_2\mathbf{L}_6, \alpha_2\mathbf{L}_3 + \beta_2\mathbf{L}_7, \alpha_2\mathbf{L}_4 + \beta_2\mathbf{L}_8, \mathbf{L}_9\},$$

where α_2 and β_2 are any consistent and nontrivial solutions of

$$\alpha_2(B - A/3) + \beta_2(D - C/3) = 0, \qquad (25)$$

where $[B - A/3]$ and $[D - C/3]$ are the coefficients of k_1 and k_2 in $k_{24}^{g_1}$, respectively. We note in particular that $\Lambda(I_i + ... + I_4)S_2' = S_2'$ for $i = 1,2,3,4$. Similarly, S_3' is defined with α_2 replaced by α_3 and β_2 replaced by β_3 as nontrivial and consistent solutions of

$$\alpha_3(B/3 - A) + \beta_3(D/3 - C) = 0. \qquad (26)$$

Finally, S_4' is nothing but the subspace $S_1 = \{\mathbf{L}_9\}$. It is important to note that these subspaces do *not* depend on a particular representation of the $\mathbf{L}$ vectors, by definition. We have only chosen the representations given above only for clarity. For example, we could have chosen a representation of $\mathbf{L}_1, ..., \mathbf{L}_8$ in S_1' with nonvanishing components $k_{24}^{g_1 1}, ..., k_{13}^{g_2 4}$, respectively. We now readily see that

$$\{\dim\Lambda(I_2 + ... + I_8)S_1' - \dim S_1, ..., \dim\Lambda(I_2 + ... + I_8)S_4' - \dim S_1\} = \{1,0,1,0\}, \qquad (27)$$

and hence $p_1 = 1$,

$$\{\dim\Lambda(I_{i+1} + ... + I_8)S_1' - \dim S_1\} = \{0\}, \qquad (28)$$

for $i = 2,3,4,6,7,8$, and $p_i = 0$, and finally, for $i = 5$,

$$\{\dim\Lambda(I_6 + ... + I_8)S_1' - \dim S_1, \dim\Lambda(I_6 + \cdots + I_8)S_3' - \dim S_1\} = \{5,4\} \qquad (29)$$

and hence $p_5 = 1$, i.e., $\beta_I(S_1) = \sum_{i=1}^{8} p_i = 2$, from Eq. (14). Hence

$$A(q, \lambda_1 m, \lambda_1\lambda_2\mu) = O\left((\lambda_1\lambda_2)^1\left(\frac{1}{\lambda_1}\right)^1\left(\frac{1}{\lambda_2}\right)^1 \sum_{n=0}^{2} a_n\left(\ln\frac{1}{\lambda_1}\right)^n\right) = O\left(\sum_{n=0}^{2} a_n\left(\ln\frac{1}{\lambda_1}\right)^n\right), \qquad (30)$$

where we have used the result in example 1 for the behavior corresponding to $\lambda_2 = \lambda\to 0$, and the fact that $d(G) = 1$. This example is rich enough to illustrate the basic rules of Sec. II. Needless to say, the above examples have been chosen from quantum electrodynamics, but the rules may be applied to other theories as well.

IV. CONCLUSION

We have given the rules to determine, in Euclidean space, the behavior of any Feynman amplitude with or without subtractions when the masses of an arbitrary subset of the masses in the underlying theory become small and, in general, at different rates. The external momenta of the corresponding graph are nonexceptional and all subtractions are performed at the origin. These rules are given in Sec. II and are straightforward to apply. We summarize these rules as follows.

A. Rules for determining the behavior of A for $\lambda_1\to 0, ..., \lambda_s\to 0$

In reference to the parameter $\lambda_i (i = 1,...,s)$, let G_i' be any subdiagram of G (which may coincide with G itself if applicable) which contains all the external vertices of G (though not necessarily all the lines) and is such that (1) all the masses in G/G_i' (if not empty) are in the set $\{\mu_i, \mu_{i+1}, ...\mu_s\}$; (2) if there is an internal vertex in G to which is attached some of the lines in G_i' not carrying any external momenta and not forming closed loops, then, necessarily, the masses carried by these lines must be from the set $\{\mu_1, ..., \mu_{i-1}, \mu_{s+1}, ..., \mu_\rho\}$; (3) let T denote the totality of all subdiagrams $\{G_i', G_i'', ...\}$ as defined in (1) and (2) and $d(G_i') = d(G_i'') = ...$ such that if $\mathbf{G}$ is any subdiagram respecting (1) and (2), then $d(\mathbf{G}) \leqslant d(G_i')$. Then, only if $d(\mathbf{G}) = d(G_i')$, then $\mathbf{G}\in T$. Let $g_1', g_2', ...$ be the set of all possible proper but not necessarily connected subdiagrams of G_i' (which may include G_i' if applicable) such that (4) each proper and connected part of each of $g_1', g_2', ...$ is divergent; and (5) all the masses in $g_1' g_2', ...$ are in the set $\{\mu_i, \mu_{i+1}, ..., \mu_s\}$. Reconsider the steps (4) and (5) for the remaining subdiagrams $G_i', ...$ in T. Repeat the above steps for each of the parameters $\lambda_i, i = 1,2,...,s$. The behavior of $A(p_1, ..., p_{4m}, \lambda_1\mu_1, ...\lambda_1 ... \lambda_s\mu_s, \mu_{s+1}, ..., \mu_\rho)$ is then given by Eq. (16).

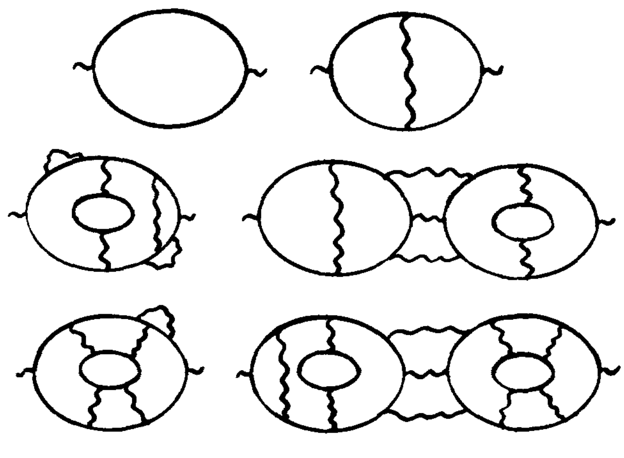

FIG. 3. Some low and high order photon self-energy graphs.

A proper but disconnected subdiagram is meant that the number if its connected part does not increase upon cutting any one of its lines. For additional details of the above rules refer to Sec. II. For applications of these rules refer to the examples in Sec. III.

From remark (2) at the end of Sec. II, we may also summarize the rule which gives a sufficiency condition for the existence of the limit of $A(p_1,...,p_{4m},\lambda_1\mu_1,...,\lambda_1 ...\lambda_s\mu_s,\mu_{s+1},...,\mu_\rho)$ for $\lambda_1\to 0,...,\lambda_s\to 0$, independently.

B. Rule (sufficiency condition) for the existence of lim$A$

If the following two conditions are true, in reference to *each* of the parameters λ_i $(i=1,...,s)$, then the lim$A$ exists:

(1) *Any* subdiagram G_0 of the whole graph G, which contains all the external vertices of G for which (a) all the lines in G/G_0 depend on the masses only from the set $\{\mu_i,\mu_{i+1},...,\mu_s\}$ and (b) if there is an internal vertex in G to which is attached some of the lines in G_0 not carrying any external momenta and not forming closed loops, then, necessarily, the mass carried by this particular line must be from the set $\{\mu_1,...,\mu_{i-1},\mu_{s+1},...,\mu_\rho\}$, then G_0 is *such that* $d(G_0)\leqslant d(G)$. This inequality must be true for all such G_0's as just defined.

(2) Only for those G_0's as defined in (1), for which $d(G_0)=d(G)$, then *no* divergent proper and connected subdiagram $g\subseteq G_0$ results in G with *all* the masses in g from the set $\{\mu_i,\mu_{i+1},...,\mu_s\}$.

The above rule is elementary and reduces only to a direct inspection of the graph G and needs no reference to any other details. As an example, we may apply it again to quantum electrodynamics, and we note that photon self-energy parts, e.g., a few of which are shown in Fig. 3, exist in the photon zero-mass limit since both (1) and (2) in the above rule are trivially true (see also example 1 in Sec. III). Other examples and theories may be also readily inspected.

[1]D.A. Slavnov, Teor. Mat. Fiz. **19**, 3 (1974); **17**, 169 (1973) [English Translation J. Theor. Math. Phys. **19**, 307 (1974): **17**, 1067 (1973)].
[2]M.C. Bergère and Y.-M.P. Lam, Commun. Math. Phys. **48**, 267 (1976).
[3]W. Zimmermann, Commun. Math. Phys. **15**, 208 (1969).
[4]S. Weinberg, Phys. Rev. **118**, 838 (1960).
[5]E.B. Manoukian, J. Math. Phys. **19**, 917 (1978).
[6]J.P. Fink, J. Math. Phys. **9**, 1389 (1968).
[7]E.B. Manoukian, Phys. Rev. D **14**, 966, 2202 E (1976).

Lorentz deformation in the O(4) and light-cone coordinate systems

Y. S. Kim
Center for Theoretical Physics, Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742

Marilyn E. Noz
Department of Radiology, New York University, New York, New York 10016

S. H. Oh
Laboratory for Nuclear Science and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 14 August 1979; accepted for publication 18 January 1980)

A formalism is developed for describing Lorentz deformation properties of extended hadrons in terms of solutions of the harmonic oscillator equation in the O(4) and light-cone coordinate systems. The physical hadronic wave function discussed in previous papers is written as a linear expansion of orthonormal functions in those coordinates which form representations of compact groups. The separability of the oscillator equation is shown to play the essential role in developing the proposed mathematics.

I. INTRODUCTION

In constructing models of relativistic quantum mechanics, we face the following two mathematical problems. First, the relativistic dynamics should be constructed in a four-dimensional Minkowskian space-time. Second, we have to impose a covariant subsidiary condition in order to obtain a three-dimensional Euclidean space in which Schrödinger quantum mechanics is reproduced.[1] Because of its mathematical simplicity, the harmonic oscillator model has been very effective in studying these mathematical problems.[2,3]

One of the special features of the covariant oscillator formalism is that the starting differential equation is separable in many different coordinate systems.[4] As is amply demonstrated in the work of Kalnins and Miller,[5] the technique of separating partial differential equations is one of the most useful techniques in mathematical physics.[6] While the authors of Ref. 5 are primarily concerned with the Klein–Gordon equation for which the separation problem is that of the D'Alembertian, the relativistic oscillator equation contains, in addition, the Lorentz-invariant "potential." Since this quadratic term is separable in the same coordinate system as the Klein–Gordon equation, the techniques developed in Ref. 5 can be applied to studying the relativistic oscillator equation.

While this remains as an interesting future problem, we should also note the following difference. Unlike those of the Klein–Gordon equation which run over the entire space–time, solutions of the oscillator equation are localized within a specified space–time region due to their Gaussian factors. This localization region, in some cases, undergoes Lorentz deformation. The oscillator wavefunctions that were discussed in our previous publications[3,7] are localized within a Lorentz-deformable space–time region, as is demonstrated in Fig. 1 of Ref. 7, and this deformation property has been shown to be consistent with what we observe in the real world.[3]

As was pointed out in Ref. 7, the mathematical formalism of the covariant oscillator model is based on the kinematics of a moving Lorentz frame in which the hadron is at rest. The purpose of the present paper is to discuss mathematical tools with which we can understand and interpret the Lorentz deformation property in terms of solutions of the wave equation in more traditional coordinate systems. Among several coordinate systems proposed for studying relativistic dynamics, the O(4) and light-cone coordinates are frequently discussed in the literature.

The convenience of the O(4) coordinate system was noted originally by Wick in connection with the study of the Bethe–Salpeter equation.[8,9] Because it includes the time variable, O(4) coordinates are used often for interpreting high-energy data or for studying possible new dynamical symmetries.[10] There has also been an attempt to employ these coordinates to construct a relativistic oscillator model.[11] In this paper, we study first the Lorentz transformation property of extended hadrons using solutions of the harmonic oscillator equation in the O(4) coordinate system.

Because its Lorentz transformation takes a convenient form, the light-cone coordinate system is regarded as one of the most promising coordinate systems for constructing relativistic quantum mechanics.[1,12] It is therefore of interest to see what form the above-mentioned oscillator wave function takes in light-cone coordinates.

In this paper, we show that the harmonic oscillator equation is again separable in both the O(4) and light-cone coordinate variables, and that the original wave function can be represented as a linear combination of the solutions in these coordinate systems. The mathematics of the harmonic oscillator or equivalently of the Hermite polynomials and of the Gaussian function is well known, as is the kinematics of Lorentz transformation. However, the mathematics of oscillator wavefunctions combined with Lorentz kinematics is not yet well known. We shall develop this relatively new mathematics using the technique of variable separation.

In Sec. II, we write down the harmonic oscillator differential equations in Lorentz-invariant form, in O(4) coordi-

 0022-2488/80/051224-05$01.00

nates, and in light-cone coordinates. The problem of this paper is then to write the physical wave function representing the Poincaré group in terms of the O(4) and light-cone solutions. Section III contains a detailed analysis of the O(4) coordinates. The "moving O(4)" coordinates are briefly discussed in Sec. IV. Section V deals with solutions of the oscillator equation in the light-cone coordinate system. In Sec. VI, we discuss mathematical and physical implications of the results obtained in this paper.

II. FORMULATION OF THE PROBLEM

In our previous paper,[7] we considered a free hadron consisting of two quarks bound together by a harmonic oscillator force of unit strength. We started with the equation

$$\{2[\Box_1 + \Box_2] - (x_1 - x_2)^2/16 + m_0^2\}\ \phi(x_1,x_2) = 0, \tag{1}$$

where x_1 and x_2 are the space–time coordinates for the two spinless quarks inside the hadron. We then used the two usual independent variables:

$$X = (x_1 + x_2)/2,$$
$$x = (x_1 - x_2)/2\sqrt{2}. \tag{2}$$

The solution $\phi(x_1,x_2)$ took the form

$$\phi(x_1,x_2) = \phi(x,X) = \psi(x,P)\exp(\pm iP\cdot X), \tag{3}$$

where P is the four momentum of the hadron and $\psi(x,P)$ is the internal wavefunction describing the motion of the quarks inside the hadron. The internal wavefunction satisfies the differential equation

$$(\tfrac{1}{2})[\Box - x_\mu^2]\psi(x,P) = (\lambda + 1)\psi(x,P), \tag{4}$$

and

$$P^\mu a_\mu^\dagger\ \psi(x,P) = 0, \tag{5}$$

where

$$a_\mu^\dagger = x_\mu + \frac{\partial}{\partial x^\mu}.$$

In Ref. 7, we showed that normalizable solutions of the partial differential equation of Eq. (4) satisfying the subsidiary condition of Eq. (5) form a representation of the Poincaré group. It was observed there that the solution $\psi(x,P)$ takes a very simple form if we use moving Lorentz coordinates in which the hadron is at rest:

$$x' = x, \quad y' = y,$$
$$z' = (z - \beta t)/(1 - \beta^2)^{1/2},$$
$$t' = (t - \beta z)/(1 - \beta^2)^{1/2}, \tag{6}$$

where the hadron is assumed to be moving along the z direction with velocity parameter β. We showed further that the wavefunctions $\psi(x,P)$ are diagonal in the Casimir operators of the Poincaré group.

Since we are not yet familiar with the moving coordinate variables, we still have to explain the result of our previous paper[7] in terms of solutions of the oscillator equation in more traditional coordinate systems. Because the O(4) coordinates are frequently discussed in the literature,[8-11] we are interested in describing the Lorentz deformation property of $\psi(x,P)$ in terms of solutions of the O(4) oscillator equation:

$$(\tfrac{1}{2})\Big[-(\nabla^2 + \frac{\partial^2}{\partial t^2} + (\vec{x}^2 + t^2)\Big]u(x) = (\sigma + 2)u(x). \tag{7}$$

Solutions of the above differential equation form a complete set in the four-dimensional Euclidean space of x and t. The Lorentz deformed wavefunction $\psi(x,P)$ is well localized in this space. Therefore, $\psi(x,P)$ can be written as a linear expansion of the orthonormal functions satisfying Eq. (7).

The Lorentz transformation takes a very simple form in the light-cone coordinate system. The transformation of Eq. (6) can be written as

$$x' = x, \quad y' = y,$$
$$(z' + t') = [(1 - \beta)/(1 + \beta)]^{1/2}(z + t),$$
$$(z' - t') = [(1 + \beta)/(1 - \beta)]^{1/2}(z - t). \tag{8}$$

Because this light-cone coordinate system is commonly used in the literature, we are interested in translating the Lorentz deformation of $\psi(x,P)$ into the language of solutions of the oscillator equation in the light-cone coordinate system:

$$\frac{1}{2}\Big[-\Big(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}\Big) + (x^2 + y^2 + \xi^2 + \eta^2)\Big]g(x) = (\epsilon + 2)g(x),$$

where

$$\xi = (z' + t')/\sqrt{2}, \tag{9}$$

$$\eta = (z' - t')/\sqrt{2}. \tag{10}$$

Here again, solutions of this differential equation form a complete set, and the Lorentz deformed $\psi(x,P)$ can be expressed in terms of a orthonormal set of wavefunctions satisfying Eq. (9).

Compared with Eq. (4), which serves as the starting point for the oscillator formalism, the O(4) and light-cone differential equations given in Eqs. (7) and (9) do not appear to carry any direct physical interpretation. However, the basic mathematical advantage of using these equations is that their solutions form representations of compact groups.

III. OSCILLATORS IN THE O(4) COORDINATE SYSTEM

The usual approach to the solution of O(4)-invariant equations is to use polar variables:

$$t = \rho\cos\alpha,$$
$$x = \rho\sin\alpha\sin\theta\cos\phi,$$
$$y = \rho\sin\alpha\sin\theta\sin\phi,$$
$$z = \rho\sin\alpha\cos\theta. \tag{11}$$

We can then write the solution of Eq. (7) as

$$u_{\mu b}^{lm}(x) = S_{\mu b}(\rho)\,Z_{b+1}^{lm}(\alpha,\theta,\phi), \tag{12}$$

where

$$Z_{b+1}^{lm}(\alpha,\theta,\phi) = P_{b+3/2}^{-b-1/2}(\cos\alpha)(\sin\alpha)^{-1/2}\,Y_l^m(\theta,\phi),$$

and $S_{\mu b}(\rho)$ satisfies the "radial" differential equation

$$\Big(\frac{d^2}{d\rho^2} + \frac{3}{\rho}\frac{d}{d\rho} - \frac{(b+1)^2 - 1}{\rho^2} - \rho^2 + 2(\sigma + 2)\Big)$$

$$\times S_{\mu b}(\rho) = 0, \tag{13}$$

where $\sigma = 2\mu + b$. The solution of this equation takes the form

$$S_{\mu b}(\rho) = N\rho^b \exp(-\rho^2/2)\, L_\mu^{b+1}(\rho^2), \tag{14}$$

where L_μ^{b+1} is the associated Laguerre polynomial.[13]

With the above form, it is easy to perform rotations in the O(4) space, and this rotation mixes the spatial coordinates with the time variable. It is, however, well known that this O(4) rotation is quite different from the Lorentz transformation. The purpose of this section is to see precisely how these two transformations are different, using the oscillator formalism as an illustrative example.

The unique feature of the harmonic oscillator "potential" is that the differential equation is separable also in the Cartesian coordinate system, where solutions of Eq. (7) can take the form

$$u(x) = f_s(x) f_w(y) f_n(z) f_k(t), \tag{15}$$

$$\text{with } \sigma = s + w + n + k.$$

The $f_s, f_w, \cdots$ are normalized one-dimensional oscillator wavefunctions, and can be written as

$$f_n(z) = (\sqrt{\pi}\, 2^n n!)^{-1/2} H_n(z) \exp(-z^2/2), \quad \text{etc.} \tag{16}$$

The above solution is not really different from the wave function given in Eq. (12). As in the well-known case of the three-dimensional oscillator, the spherical form is a linear combination of the Cartesian wavefunctions. For $\beta = 0$, $s = w = n = k = 0$, and the wavefunction becomes

$$\mu_0(x) = (1/\pi) \exp(-\rho^2/2)$$

$$= (1/\pi) \exp[-(x^2 + y^2 + z^2 + t^2)/2]. \tag{17}$$

Appendix A contains a detailed discussion of the degeneracies in the Cartesian and spherical coordinates for $\sigma = 1$ and 2. When we discuss Lorentz transformations, it is more convenient to use the Cartesian solutions.

In Ref. 7, the physical wavefunction took the form

$$\psi(x) = R_{vl}(r') Y_l^m(\theta', \phi'), \tag{18}$$

where r', θ', ϕ' are the spherical variables in the x', y', z' coordinate. R_{vl} is the normalized radial wavefunction for the three-dimensional isotropic oscillator, and its form is well known. This expression can also be written in terms of the moving Cartesian coordinate variables:

$$\psi(x) = f_s(x') f_w(y') f_n(z') f_0(t'), \tag{19}$$

with

$$v = s + w + n.$$

$f_s, \cdots$ are defined in Eq. (16). The t' excitation is suppressed by the subsidiary condition of Eq. (5).

Our original problem was to express the physical wavefunction of Eq. (18) as a linear combination of the O(4) solutions $u_{\mu b}^{lm}(x)$ of Eq. (12). However, thanks to the separability of the oscillator potential, the problem can be reduced to that of relating their respective Cartesian forms given in Eqs. (15) and (19). We note first that the transverse components can be dropped, because they are not affected by Lorentz transformations, and write $\psi(x)$ and $u(x)$ as

$$\psi_\beta^n(z,t) = f_n(z') f_0(t'), \quad \text{with } \lambda = n,$$

$$u_k^a(z,t) = f_a(z) f_k(t), \quad \text{with } \sigma = a + k. \tag{20}$$

The problem then becomes that of writing

$$\psi_\beta^n(z,t) = \sum_{a,k} B_a^{n,k}(\beta) u_k^a(z,t), \tag{21}$$

and then of calculating the coefficient $B^{n,k}(\beta)$:

$$B_a^{n,k}(\beta) = \int dt\, dz\, u_k^a(z,t) \psi_\beta^n(z,t). \tag{22}$$

The integrand in the above expression contains three Hermite polynomials and four Gaussian factors that are quite familiar to us. What is new here is that the arguments of these functions are the coordinate variables for two different Lorentz frames. The detailed calculation based on the generating function of the Hermite polynomial is given in Appendix B. The coefficient $B_a^{n,k}$ takes the form

$$B_a^{n,k}(\beta) = A_k^n(\beta) \delta_{a,n+k}, \tag{23}$$

where

$$A_k^n(\beta) = \beta^k (1 - \beta^2)^{(n+1)/2} [(n+k)!/n!k!]^{1/2}. \tag{24}$$

Equation (21) can now be written as a summation over a single index:

$$\psi_\beta^n(z,t) = \sum_{k=0}^{\infty} A_k^n(\beta)\, u_k^{n+k}(z,t), \tag{25}$$

$$\text{with} \quad \sum_k |A_k^n(\beta)|^2 = 1. \tag{26}$$

If we write the above sum explicitly,

$$\psi_\beta^n(z,t)$$

$$= (\pi 2^n n!)^{-1/2} (1 - \beta^2)^{(n+1)/2} \exp[-(z^2 + t^2)/2]$$

$$\times \sum_{k=0}^{\infty} \frac{\beta^k}{2^k k!} H_{n+k}(z) H_k(t). \tag{27}$$

The purpose of this section was to construct a representation of the Poincaré group describing the hadronic Lorentz deformation in terms of the harmonic oscillator solutions in the O(4) coordinate system. We noted first that the oscillator equation is separable in the Cartesian variables, and then showed that the representation can be constructed explicitly.

It is well known that the O(4) rotation is quite different from the Lorentz transformation. The explicit form given in Eq. (27) consisting of an infinite sum of the O(4) solutions indicates precisely how these two transformations are different.

IV. MOVING O(4) COORDINATE

If the hadron is at rest with $\beta = 0$, only the $k = 0$ term contributes in Eq. (27). This term takes a functional form identical to that of the physical wave function. This leads us to consider the moving O(4) coordinate system consisting of the coordinate variables x', y', z', t', and solutions of the oscillator equation

$$\frac{1}{2}\left[-\left(\nabla'^2 + \frac{\partial^2}{\partial t'^2}\right) + (\vec{x}'^2 + t'^2)\right] g(x) = (\epsilon + 2) g(x). \tag{28}$$

If we solve this equation with the restriction

$$\left(\frac{\partial}{\partial t'} + t'\right)g(x) = 0, \tag{29}$$

then $g(x)$ becomes identical to $\psi(x)$ of Eq. (18). The physical wavefunction therefore represents a solution of two entirely different partial differential equations. This is because the t' variable can be separated both in the original equation of Eq. (4) and in the moving O(4) equation given in Eq. (27).

V. LIGHT-CONE COORDINATE

We noted in Sec. IV that Eq. (28) together with the subsidiary condition of Eq. (29) can generate the physical wavefunction. If we rewrite Eq. (27) in terms of the light-cone variables of Eq. (10), then the resulting differential equation is the light-cone oscillator equation given in Eq. (9), with solutions of the form

$$g(x) = f_s(x)f_w(y)f_i(\xi)f_j(\eta). \tag{30}$$

Here again we can ignore the x and y components. The problem now becomes that of writing

$$\psi_\beta^n(z,t) = \sum_{i,j}^{\infty} C_{i,j} f_i(\xi)f_j(\eta), \tag{31}$$

and then of calculating the coefficient $C_{i,j}$.

As far as the Gaussian factors are concerned, we note that

$$z'^2 + t'^2 = \xi^2 + \eta^2. \tag{32}$$

As for the Hermite polynomial $H_n(z')$, we use the addition formula[13]

$$H_n(z') = H_n((\xi + \eta)/\sqrt{2}) = \left(\frac{1}{2}\right)^{n/2} \sum_{m=0}^{\infty} \binom{n}{m} H_{n-m}(\xi)H_m(\eta). \tag{33}$$

Thus the explicit form for the physical wavefunction becomes

$$\psi_\beta^n(z,t) = \left(\frac{1}{2}\right)^n (1/\pi n!)^{1/2} \exp[-(\xi^2 + \eta^2)/2] \times \left[\sum_{m=0}^{n} \binom{n}{m} H_{n-m}(\xi)H_m(\eta)\right]. \tag{34}$$

Unlike the case of the O(4) expansion, the summation in the above formula is finite, and is restricted to the total quantum number

$$\lambda = \epsilon = n, \tag{35}$$

where λ and ϵ are defined in Eqs. (4) and (9), respectively. The expansion coefficient $\binom{n}{m}$ is independent of β. The dependence on this velocity parameter is in the definition of the ξ and η variables given in Eq. (10). We realize that a more traditional definition of the light-cone variables is to use z and t, instead of z' and t', in Eq. (10).[1] However, this difference lies only in the elongation/contraction along the ξ/η axis, as is specified in Eq. (8).

The purpose of this section was to construct a representation of the Poincaré group describing the hadronic Lorentz deformation in terms of the harmonic oscillator solutions in the light-cone coordinates. This light-cone coordinate system is natural for constructing Dirac's "front form" quantum mechanics.[1] For this reason, some authors have started their effort from this coordinate system, with hopefully a simpler expression for the hadronic wave function.[12] However, the moving Lorentz frame in which the physical wave function $\psi(x)$ takes its simplest form is essentially the coordinate system for Dirac's "instant form" quantum mechanics.[14] The present paper is based on the physical prejudice gained from our previous papers[3-4] that the physics starts in the "instant" form. The mathematical formalism given in this section can also be used to translate the physics from the front form to the instant form, if and when there is enough experimental evidence to indicate that the origin of physics is in the light-cone coordinate. In either case, the physics remains covariant, and the formalism presented here remains useful.

VI. CONCLUDING REMARKS

Because the Minkowskian world with its hyperbolic topology is not convenient for visualization of a distribution localized in a specified space–time region,[15] other coordinate systems such as the O(4) and light-cone coordinates are commonly discussed in the literature. The O(4) system is convenient because it is Euclidean. The light-cone coordinates are very attractive because the Lorentz transformation simply elongates one light-cone axis and contracts the other.

In our previous papers,[3,4,15] we studied the harmonic oscillator wave function, which is compatible with the established laws of quantum mechanics and special relativity, and which can explain the basic hadronic phenomena observed in high-energy experiments. This wave function was simplest in terms of coordinate variables expressed in the Lorentz frame moving with the hadronic velocity. This oscillator wave function is localized in a space-time region and undergoes a Lorentz deformation. In this paper, we translated this deformation property into a language based on solutions of the oscillator equation in O(4) coordinates, and into that in the light-cone coordinate system.

Separation of variables is one of the oldest mathematical arts known to physicists. It is also well known that the harmonic oscillator potential is separable in many different coordinate systems, including the Cartesian system. In this paper, we combined these two well-known features to study a mathematical formalism with which we can visualize the Lorentz deformation property of relativistic extended hadrons.

APPENDIX A

As in the well-known case of the three-dimensional isotropic harmonic oscillator, the O(4) oscillator equation is separable in both the Cartesian and spherical coordinate systems. We note that the radial variable ρ is related to the Cartesian variables by

$$\rho = (x^2 + y^2 + z^2 + t^2)^{1/2} \tag{A1}$$

and that the ground-state solution is given in Eq. (17).

The first excited states are

$$\psi_1(x) = (\sqrt{2}/\pi)x_i \exp(-\rho^2/2), \tag{A2}$$

where $x_i = x,y,z$, or t. In the spherical coordinate system,

$\sigma = b = 1$, and $\mu = 0$. Thus the wavefunction of Eq. (12) takes the form

$$u_{01}^{lm}(x) = \rho \exp(-\rho^2/2) Z_2^{lm}(\alpha,\theta,\phi), \tag{A3}$$

where

$$Z_2^{00}(\alpha,\theta,\phi) = (2/\pi^2)^{1/2} \cos\alpha,$$

$$Z_2^{lm}(\alpha,\theta,\phi) = (8/3\pi)^{1/2} \sin\alpha \, Y_1^m(\theta,\phi), \text{ with } m = 1,0,-1.$$

Here again, the wavefunction is fourfold degenerate. These spherical wavefunctions can be written as linear combinations of the Cartesian forms given in Eq. (A2).

For $\sigma = 2$, there are four Cartesian wavefunctions of the form

$$\psi_i(x) = (1/\pi\sqrt{2})(2x_i^2 - 1) \exp(-\rho^2/2), \tag{A4}$$

and six wavefunctions of the form

$$\psi_{ij}(x) = (2/\pi) x_i x_j \exp(-\rho^2/2), \quad \text{with } i \neq j. \tag{A5}$$

There are therefore ten degenerate Cartesian wavefunctions.

Let us now look at the $\sigma = 2$ states in the O(4)-symmetric coordinate system. We have to consider here two different values of the radial quantum number μ:

$$\text{(a) } \mu = 1, \quad \text{with } b = 0;$$

$$\text{(b) } \mu = 0, \quad \text{with } b = 2. \tag{A6}$$

If $\mu = 1$ and $b = 0$, the spherical wavefunction of Eq. (12) becomes

$$u(x) = (1/\pi\sqrt{2})(2 - \rho^2) \exp(-\rho^2/2). \tag{A7}$$

If $\mu = 0$ and $b = 2$, the wavefunction is

$$u(x) = (\rho^2/\sqrt{3}) \exp(-\rho^2/2) Z_3^{lm}(\alpha,\theta,\phi). \tag{A8}$$

For $l = 0$, there is one wavefunction with

$$Z_3^{00} = (1/2\pi^2)^{1/2}(4\cos^2\alpha - 1). \tag{A9}$$

For $l = 1$, there are three wavefunctions with

$$Z_3^{lm} = (16/\pi)^{1/2} \cos\alpha \sin\alpha \, Y_l^m(\theta,\phi), \quad m = 1,0,-1. \tag{A10}$$

For $l = 2$, there are five wavefunctions with

$$Z_3^{2m} = (16/5\pi)^{1/2} \sin^2\alpha \, Y_2^m(\theta,\phi), \quad m = 2,1,0,-1,-2. \tag{A11}$$

There are therefore ten degenerate O(4) symmetric wavefunctions.

We can write the above wavefunctions in terms of the Cartesian variables of Eq. (11), and as linear combinations of the Cartesian wavefunctions given in Eqs. (A4) and (A5).

APPENDIX B

In order to evaluate the integral of Eq. (22), we use the generating function for Hermite polynomials[16]:

$$G(r,z) = \exp(-r^2 + 2rz) = \sum_{m=0}^{\infty} \frac{r_m}{m!} H_m(z). \tag{B1}$$

The calculation is then reduced to the evaluation of the integral

$$I = \int dt\, dz\, G(r,z) G(s,t) G(r',z') \times \exp[-(z^2 + t^2 + z'^2 + t'^2)/2]. \tag{B2}$$

The integrand of the above expression is an exponential function whose argument is quadratic in z and t.

We can diagonalize this quadratic form using the light-cone variables. The integral then becomes

$$I = \pi(1 - \beta^2)^{1/2} \exp(2\beta rs) \exp[2rr'(1 - \beta^2)^{1/2}]. \tag{B3}$$

By expanding the above exponential factors, we arrive at the result given in Eqs. (23) and (24).

Let us next check whether the expansion coefficient $A_k^n(\beta)$ in Eq. (25) satisfies the unitarity condition. We start with

$$\sum_{k=0}^{\infty} |A_k^n(\beta)|^2 = (1 - \beta^2)^{n+1} \sum_{k=0}^{\infty} \beta^{2k} \frac{(n+k)!}{n!k!}. \tag{B4}$$

However, from the binomial expansion,

$$\left(\frac{1}{1-\beta^2}\right)^{(n+1)} = \sum_{k=0}^{\infty} \frac{\beta^{2k}(n+k)!}{n!k!}. \tag{B5}$$

Thus

$$\sum_{k=0}^{\infty} |A_k^n(\beta)|^2 = 1, \tag{B6}$$

and Eq. (27) is indeed an expansion is a complete orthonormal set.

[1] P.A.M. Dirac, Rev. Mod. Phys. **21**, 392 (1949).
[2] H. Yukawa, Phys. Rev. **19**, 416 (1953); M. Markov, Suppl. Nuovo Cimento **3**, 760 (1956); R.P. Feynman, M. Kislinger, and F. Ravndal, Phys. Rev. D **3**, 2706 (1971).
[3] For a review paper, see Y.S. Kim, and M.E. Noz, Am. J. Phys. **46**, 484 (1978).
[4] Y.S. Kim and M.E. Noz, Phys. Rev. D **8**, 3521 (1973).
[5] E.G. Kalnins and W. Miller, J. Math. Phys. **18**, 1 (1977), and references therein.
[6] W. Miller, *Symmetry and Separation of Variables* (Addison and Wesley, Reading, MA, 1977).
[7] Y.S. Kim, M.E. Noz, and S.H. Oh, J. Math. Phys. **20**, 1341 (1979).
[8] G.C. Wick, Phys. Rev. **96**, 1124 (1954). For a review paper, see N. Nakanishi, Suppl. Prog. Theor. Phys. **43**, 1 (1969).
[9] G. Domokos and P. Suranyi, Nucl. Phys. **54**, 529 (1964).
[10] O. Nachtman, Nucl. Phys. **B63**, 237 (1973); M. Ninomiya and K. Watanabe, Nuovo Cimento **37A**, 343 (1977); V.B. Serebrennikov, A.E. Shabad, and Y.S. Smorodinsky, Nuovo Cimento A **47**, 333 (1978). The references listed here are not meant to be complete. They represent some recent publications on this subject.
[11] N.P. Chang and C.A. Nelson, Phys. Rev. D **13**, 1345 (1976).
[12] G. 't Hooft, Nucl. Phys. B **75**, 461 (1974); M. Ida, Prog. Theor. Phys. **54**, 1519 (1975); H. Yabuki, Prog. Theor. Phys. **58**, 1935 (1977); H. Leutwyler and J. Stern, Ann. Phys. **112**, 94 (1978), and references therein.
[13] W. Magnus and F. Oberhettinger, *Formulas and Theorems for the Functions of Mathematical Physics* (Chelsea, New York, 1949).
[14] Y.S. Kim and M.E. Noz, Found. Phys. **9**, 375 (1979).
[15] Y.S. Kim, M.E. Noz, and S.H. Oh, Am. J. Phys. **47**, 892 (1979).
[16] M.J. Ruiz, Phys. Rev. D **10**, 4306 (1974).

The Baker–Campbell–Haussdorff formula for the SU(2) supergroups

C. Aragone, J. Chela-Flores, and A. Restuccia
Depto. de Física, Universidad Simón Bolívar, Apartado 80659, Caracas 108, Venezuela

(Received 27 February 1979; accepted for publication 13 November 1979)

In the theory of supersymmetric SU(2) Yang–Mills fields described on the 8th dimensional superspace, the local gauge transformations constitute a group whose Lie algebra has its coefficients in the Weyl-spinorial-Grassmann algebra. The coefficients can be taken to be chiral elements, real or complex. This last case is the most general one and contains all others of interest for developing physical models. We present here a Baker–Campbell–Haussdorff formula for the complex SU(2) supergroup. The formula gives the finite form for each element of the supergroup in terms of the local fields entering into the infinitesimal (complex) superscalar generator. Thereafter we exhibit, as a particular case, a BCH-expression for the real SU(2) supergroup which completes previous results already reported for the chiral SU(2) supergroups. Moreover, it is shown that, despite the appearance of apparently nonperiodic polynomial terms, the global variety of the real SU(2) supergroup is contained in a variety having the same compact bosonic width as the usual SU(2) pure bosonic variety.

1. INTRODUCTION

Any supersymmetric formulation of Yang–Mills theories[1,2] has to deal with supergroups, that is groups generated through the exponentiation of a Lie algebra with Grassmann valued coefficients.

The same initial Lie algebra offers different possibilities, according to whether the coefficients are chosen to be chiral, real or complex Grassmann elements.[3]

Moreover, it is well known that, besides the local structure of the supergroup, its global structure is also needed[4,5] in order to understand some of the physical implications of the different supersymmetric models being developed.

In this article we obtain a closed finite expression for any element of the complex SU(2) supergroup ($SSU(2)_C$). This expression shows the very different role played by the components parallel to the bosonic part of the complex- Lie valued- scalar generator compared with their orthogonal components.

As an immediate consequence of the complex case we shall write down a BCH formula for the real SU(2) supergroup ($SSU(2)_R$) and we shall exhibit a parametrization of this supergroup which allows us to show why, in spite of the nonperiodic (rational) functions in the BCH-formula, the supervariety $SSU(2)_R$ is contained in the Cartesian product of a three-dimensional compact subset K times R^{45}. This is due to a strictly supersymmetric effect responsible for the absorption of the nonperiodic behavior in the purely bosonic coordinates by some of the remaining supersymmetric coordinates, as it will be shown in Sec. III.

The whole set of results for $SSU(2)_C$ constitute the natural generalization of the simpler results already communicated for the case of the chiral supergroups.[6]

Section II is devoted to establishing the notation and to obtaining BCH formulas for the complex $SSU(2)_C$ supergroup and, in Sec. III, we carefully discuss the $SSU(2)_R$ case and some aspects concerning its global structure, as mentioned above.

Finally, in the last section we shall summarize and discuss the results already obtained.

II. THE BCH-EXPRESSION FOR $SSU(2)_C$

Our aim is to give a closed expression for[7] e^{Ξ},

$$e^{\Xi}, \quad \Xi \equiv \Xi^a X_a, \quad X_a \equiv 2^{-1} i\sigma_a, \tag{1a}$$

$$\begin{aligned}
\Xi^a X_a \equiv {} & \zeta^a X_a + \zeta^{a\alpha}\theta_\alpha X_a + \eta^a_{\dot\beta}\bar\theta^{\dot\beta} X_a \\
& + \zeta^a_+ X_a \theta_+ + \eta^a_- X_a\theta_- + \zeta^{a\mu}\theta_\mu X_{a+} \\
& + \zeta^{a\alpha}_- X_a \theta_{\alpha -} + \eta^a_{\dot\beta +} X_a \bar\theta^{\dot\beta}_+ \\
& + \zeta^a_{+-} X_a \theta_{+-} \\
\equiv {} & \zeta\{ n_a X_a + n^{a\alpha}\theta_\alpha X_a \\
& + m^a_{\dot\beta}\bar\theta^{\dot\beta} X_a + n^a_+ \theta_+ X_a \\
& + m^a_- \theta_- X_{a+} \\
& + n^{a\mu}\theta_\mu X_a + n^{a\alpha}_- X_a \theta_{\alpha -} \\
& + m^a_{\dot\beta +} X_a \bar\theta^{\dot\beta}_+ + n^a_{+-} X_a \theta_{+-} \},
\end{aligned} \tag{1b}$$

$$n^a n^b \delta_{ab} = 1.$$

We shall achieve this goal in two steps: first we will look for Lie valued elements Z and ξ' such that

$$S \equiv e^Z \xi' e^{-Z} = \Xi, \tag{2}$$

and thereafter, having Z and ξ' that verify Eq. (2) one can straightforwardly exponentiate this equation obtaining that e^{Ξ} has the value

$$e^S = e^Z e^{\xi'} e^{-Z} = e^{\Xi}, \tag{3}$$

which is easily determined if one is previously able to calculate e^Z (and e^{-Z}) and $e^{\xi'}$ separately and then making their product as indicated in Eq. (3).

In order to find out Z and ξ' which verify Eq. (2), we make the ansatz[8]:

$$\begin{aligned}
Z \equiv {} & \xi^\alpha \theta_\alpha + \chi_{\dot\beta}\bar\theta^{\dot\beta} + \xi_+ \theta_+ \\
& + \chi_- \theta_- + \xi^\mu \theta_\mu + \xi^\alpha_- \theta_{\alpha -} \\
& + \chi_{\dot\beta +}\bar\theta^{\dot\beta}_+ + \xi_{+-}\theta_{+-} \equiv Z^a X_a,
\end{aligned} \tag{4a}$$

$$\xi' \equiv (1 + C)\xi, \quad \xi \equiv \xi^a X_a, \tag{4b}$$

where we are omitting the internal indices (i.e., $\xi^\alpha\theta_\alpha = \xi^{a\alpha}\theta_\alpha X_a, \ldots, \xi = \xi^a X_a$) and C begins with θ^1:

$$C \equiv C^\alpha \theta_\alpha + d_{\dot\beta}\dot\theta^\beta + C_+\theta_+ + d_-\theta_- + C^\mu\theta_\mu$$

 0022-2488/80/051229-05$01.00

$$+ C^{\alpha}_{-}\theta_{\alpha -} + d_{\beta +}\bar{\theta}^{\beta}_{+} + C_{+-}\theta_{+-}, \quad (5)$$

and the symbol ${}^{p}[Z,A] \equiv [Z,[Z,[Z...[Z,A]]...]$ is used to denote p commutations with the same operator Z acting on the left.

After the definition (2) of S, by virtue of the properties of the Grassman coefficients of the Lie algebra which make ${}^{p}[Z,\xi'] = 0$ for any $p \geqslant 5$, it turns out that

$$S = \xi' + [Z,\xi'] + (2!)^{-1}\,{}^{2}[Z,\xi'] + (3!)^{-1}\,{}^{3}[Z,\xi'] + (4!)^{-1}\,{}^{4}[Z,\xi']. \quad (6)$$

Moreover, in the particular case of SU(2), as[9] $[Z,\xi'] = -(Z\wedge\xi')^{c}X_{c}$, Eq. (6) can be written in the form:

$$S = \xi' + \sum_{p=1}^{p=4} (-1)^{p}(p!)^{-1}({}^{p}(Z\wedge)\xi')^{c}X_{c}. \quad (7)$$

In order to ease the calculations we restrict the search for (Z,ξ,C) of Z's orthogonal to ξ'[10]:

$$Z\cdot\xi = 0 = Z\cdot\xi'. \quad (8)$$

Then S becomes

$$S = \xi'\{1 - (2!)^{-1}Z^{2} + (4!)^{-1}Z^{4}\} + \{-1 + (3!)^{-1}Z^{2}\}Z\wedge\xi' \quad (9)$$

and, as we stated before, we want to solve Eq. (2) finding out suitable Z,ξ,C such that $S(Z,\xi,C) = \Xi$, a fixed element of $SSU(2)_{C}$.

Taking terms proportional to θ^{0} on both sides of this equation we have:

$$\xi^{a} = \zeta n^{a}. \quad (10)$$

Introduction of this result into Eq. (2) gives us a more convenient representation of Eq. (2):

$$\begin{aligned}(1+C)\zeta\hat{n}\{1 - 2^{-1}Z^{2} + 24^{-1}Z^{4}\} &+ \{-1 + 6^{-1}Z^{2}\}Z\wedge(1+C)\zeta\hat{n} \\ &= \zeta\{\hat{n} + n^{\alpha}\theta_{\alpha} + m_{\beta}\bar{\theta}^{\beta} + n_{+}\theta_{+} + m_{-}\theta_{-} \\ &\quad + n^{\mu}\theta_{\mu} + n^{\alpha}_{-}\theta_{\alpha -} + m_{\beta +}\bar{\theta}^{\beta}_{+} \\ &\quad + n_{+-}\theta_{+-}\} \equiv \zeta\{1 + N_{\parallel})\hat{n} + N_{\perp}\},\end{aligned} \quad (11a)$$

$$N_{\perp}\cdot n = 0, \quad (11b)$$

(internal indices have been omitted).

Projecting Eq. (11) along $\hat{n}$ and along the subspace orthogonal to $\hat{n}$ two equations are obtained:

$$(1+C)(1 - 2^{-1}Z^{2} + 24^{-1}Z^{4}) = 1 + N_{\parallel}, \quad (12a)$$

$$(-1 + 6^{-1}Z^{2})(1+C)Z = n\wedge N_{\perp}. \quad (12b)$$

The first one is very simple to interpret: Since Z starts with terms which are linear in θ, it is straightforward to calculate the inverse of $1 - 2^{-1}Z^{2} + 24^{-1}Z^{4}$

$$(1 - 2^{-1}Z^{2} + 24^{-1}Z^{4})^{-1} = 1 + 2^{-1}Z^{2} + 5\times 24^{-1}Z^{4}. \quad (13)$$

Then, Eq. (12a) establishes a bijective correspondence between $N_{\parallel}$ and C which can be taken in both directions, either when having C to obtain $N_{\parallel}$, the component of Ξ parallel to n, or when starting from the value of $N_{\parallel}$ to get C:

$$C = (1 + N_{\parallel})(1 + 2^{-1}Z^{2} + 5\times 24^{-1}Z^{4}) - 1. \quad (14)$$

In the following we are going to think in terms of Ξ being represented by ζ; n; C and $N_{\perp}$ as independent variables.

Then one has to exploit Eq. (12b) which can be written

$$(1 - 6^{-1}Z^{2})Z = -n\wedge N_{\perp}(1+C)^{-1} \equiv A, \quad (15)$$

where $(1+C)^{-1}$ is given in Appendix A and the supervector A has the form:

$$\begin{aligned}A \equiv a^{\alpha}\theta_{\alpha} + b_{\beta}\bar{\theta}^{\beta} + a_{+}\theta_{+} + b_{-}\theta_{-} \\ + a^{\mu}\theta_{\mu} + a^{\alpha}_{-}\theta_{\alpha -} + b_{\beta +}\bar{\theta}^{\beta}_{+} + a_{+-}\theta_{+-},\end{aligned} \quad (16a)$$

with its coefficients $a^{\alpha}, b_{\beta} \sim c^{0}, a_{+}, b_{-}$ and $a_{\mu} \sim c^{0} + ac^{1}$; a^{α}_{-} and $b_{\beta +} \sim c^{0} + ac^{1} + bc^{2}$ and $a_{+-} \sim c^{0} + ac^{1} + bc^{2} + dc^{3}$, where c^{p} represents any homogeneous monomial of degree p in the coefficients of C.

Equation (15) can be solved giving a unique value for Z:

$$\begin{aligned}Z = a^{\alpha}\theta_{\alpha} + b_{\beta}\bar{\theta}^{\beta} + a_{+}\theta_{+} + b_{-}\theta_{-} \\ + a^{\mu}\theta_{\mu} + \{a^{\alpha}_{-} + 6^{-1}b_{\beta}(a^{\alpha}\cdot b^{\beta}) \\ - 6^{-1}2^{-1}(b_{\beta}b^{\beta})a^{\alpha}\}\,\theta_{\alpha -} + \{b_{\beta +} - 6^{-1}a^{\alpha}(a_{\alpha}\cdot b_{\beta}) \\ - 6^{-1}2^{-1}(a^{\alpha}\cdot a_{\alpha})b_{\beta}\}\,\bar{\theta}^{\beta}_{+} + \xi_{+-}\theta_{+-},\end{aligned} \quad (16b)$$

(here we did not calculate ξ_{+-} because it is not needed in the rest of this article).

If one be interested in obtaining both Z and C in terms of the initially given generator Ξ, one has to substitute the value of C given by Eq. (14) into Eq. (12b) and therefore solve an equation similar to (15):

$$(1 + 3^{-1}Z^{2})Z = -n\wedge N_{\perp}(1 + N_{\parallel})^{-1}. \quad (17)$$

Once this equation is solved for Z,[11] going back to Eq. (14) and introducing this value of Z, C is found as an explicit function of n, $N_{\parallel}$ and N.

We can now proceed to compute e^{Ξ}, according to Eq. (3). First e^{Z} and e^{-Z} are evaluated. Due to the fact that $Z^{5} = 0$, these two exponentials take respectively the values

$$\begin{aligned}e^{Z} = (1 - (2!)^{-1}2^{-2}Z^{2} + (4!)^{-1}2^{-4}Z^{4})\mathbb{1} \\ + i2^{-1}(1 - (3!)^{-1}2^{-2}Z^{2})Z^{a}\sigma_{a} \equiv \alpha\mathbb{1} + i\beta Z^{a}\sigma_{a},\end{aligned} \quad (18a)$$

$$e^{-Z} = \alpha\mathbb{1} - i\beta Z^{a}\sigma_{a} \equiv \alpha\mathbb{1} - i\beta \not{Z}, \quad V^{a}\sigma_{a} \equiv \not{V} \quad (18b)$$

Thereafter one can compute $e^{\xi'} = e^{(1+C)\zeta n^{a}X_{a}}$. Since $[n^{a}X_{a}, Cn^{b}X_{b}] = 0$

$$\begin{aligned}e^{\xi'} = e^{\zeta n^{a}X_{a}}e^{\zeta Cn^{a}X_{a}} = (\cos(\zeta/2) + i\sin(\zeta/2)\cdot\not{n}) \\ \times\{1 - (2!)^{-1}2^{-2}C^{2}\zeta^{2} + (4!)^{-1}2^{-4}C^{4}\zeta^{4} \\ + i2^{-1}(C\zeta - (3!)^{-1}2^{-2}C^{3}\zeta^{3})\not{n}\} \equiv \gamma\mathbb{1} + i\delta\not{n},\end{aligned} \quad (19a)$$

where

$$\begin{aligned}\gamma &\equiv \cos(\zeta/2)\{1 - (2!)^{-1}2^{-2}C^{2}\zeta^{2} + (4!)^{-1}2^{-4}C^{4}\zeta^{4}\} \\ &\quad - \sin(\zeta/2)\{2^{-1}C\zeta - (3!)^{-1}2^{-3}C^{3}\zeta^{3}\}, \\ \delta &\equiv \sin(\zeta/2)\{1 - (2!)^{-1}2^{-2}C^{2}\zeta^{2} + (4!)^{-1}2^{-4}C^{4}\zeta^{4}\} \\ &\quad + \cos(\zeta/2)\{2^{-1}C\zeta - (3!)^{-1}2^{-3}C^{3}\zeta^{3}\}.\end{aligned} \quad (19b)$$

Then, substitution of the values (18) and (19) of the three exponentials factorizing e^{Ξ} gives BCH-expressions

$$\begin{aligned}e^{\Xi} &= \gamma\mathbb{1} + i\delta(1 - 2^{-1}Z^{2} + 3^{-1}2^{-3}Z^{4})\not{n} + i\delta(1 - 6^{-1}Z^{2})\not{n}\wedge Z \\ &= \gamma\mathbb{1} + i\delta(1 - 2^{-1}Z^{2} + 3^{-1}2^{-3}Z^{4})\not{n} + i\delta(1+C)^{-1}\not{N}_{\perp} \\ &= \gamma\mathbb{1} + i\delta(1+C)^{-1}\{(1 + N_{\parallel})\not{n} + \not{N}_{\perp}\}.\end{aligned} \quad (20)$$

The last expression came out from Eq. (12). If C is thought of as being solved using Eqs. (17) and (14), this last expression is the BCH-formula in term of the original data of Ξ. If one concentrates on ζ, n, C, and $N_{\perp}$ as a good parametrization of the supergroup, then the second expression is more appropriate, where one has to think of Z as the solution of Eq. (15) given in terms of n, $N_\perp$ and C, as in Eq. (16b).

Introduction of the value (16b) of Z into Eq. (20) yields

$$
\begin{aligned}
e^{\Xi} = \gamma \mathbb{1} + i\delta \{ 1 &+ 2^{-2} a^\alpha a_\alpha \theta_+ + 2^{-2} b_{\dot\beta} b^{\dot\beta} \theta_- \\
&+ 2^{-1} a_\alpha b_{\dot\beta} \sigma_\mu{}^{\alpha\dot\beta} \theta^\mu \\
&- a^\alpha b_- \theta_{\alpha -} - 2^{-1} \sigma^{\mu\alpha\dot\beta} b_{\dot\beta} a_\mu \theta_{\alpha -} \\
&- b_{\dot\beta} a_+ \bar\theta^{\dot\beta}_+ + 2^{-1} \sigma^\mu_{\alpha\dot\beta} a^\alpha a_\mu \bar\theta^{\dot\beta}_+ \\
&+ (2^{-1} a^\alpha_- a_\alpha + 2^{-1} b_{\dot\beta +} b^{\dot\beta} + 2^{-2} a^\mu a_\mu \\
&- a_+ b_- - 2^{-4} a^\alpha a_\alpha b_{\dot\beta} b^{\dot\beta} \\
&- 2^{-3} (b_{\dot\beta} a^\alpha)(a_\alpha b^{\dot\beta}) \theta_{+-} \} \not{n} + i\delta (1 + C)^{-1} \not{N}_\perp,
\end{aligned} \tag{21}
$$

where $a^\alpha, b_{\dot\beta}, a_+, a^\mu, b_-, a^\alpha_-, b_{\dot\beta +}$ are given in the Appendix in terms of C and $N_\perp$, and γ and δ are determined by Eq. (19b).

This last expression is one of the BCH-formulas for the SU(2) complex supergroup which shall be exhibited in this article. From this formula similar expressions for $\mathrm{SSU}(2)_R$ and chiral $\mathrm{SSU}(2)_L$ (or $\mathrm{SSU}(2)_R$) can easily be written, just by making the appropriate specialization.

III. THE UNITARY SUPERGROUP $\mathrm{SSU}(2)_R$

If one demands that e^{Ξ} be unitary

$$(e^{\Xi})^+ = e^{-\Xi} = (e^{\Xi})^{-1}, \tag{22}$$

Ξ cannot have the general form (1b) with all the coefficients being arbitrary complex spinors. In order to satisfy Eq. (22) Ξ must have the structure

$$
\begin{aligned}
\Xi = \zeta^a X_a &+ \zeta^{a\alpha} \theta_\alpha X_a + \bar\zeta^a{}_{\dot\beta} \bar\theta^{\dot\beta} X_a \\
&+ \zeta^a_+ \theta_+ X_a + \bar\zeta^a_+ \theta_- X_a \\
&+ \zeta^{a\mu} \theta_\mu X_a + \zeta^{a\alpha}_- X_a \theta_{\alpha -} \\
&+ \bar\zeta^a_{\dot\beta -} X_a \theta^{\dot\beta +} + \zeta^a_{+-} \theta_{+-} X_a,
\end{aligned} \tag{23a}
$$

$$\Xi = \Xi^a X_a : \Xi^{a+} = \Xi^a \text{ with } X^+_a = -X_a, \tag{23b}$$

$$\zeta^a, \zeta^{a\mu}, \zeta^a_{+-} \quad \text{reals}, \tag{23c}$$

which means that the three Grassmann scalars Ξ^a are real elements of the algebra.

Reality of Ξ^a imposes reality of Z, ξ', and C. They respectively become

$$
\begin{aligned}
Z = \xi^a \theta_\alpha &+ \bar\xi_{\dot\beta} \bar\theta^{\dot\beta} + \xi_+ \theta_+ + \bar\xi_+ \theta_- + \xi^\mu \theta_\mu \\
&+ \xi^\alpha_- \theta_{\alpha -} + \bar\xi_{\dot\beta -} \bar\theta^{\dot\beta}_+ + \xi_{+-} \theta_{+-},
\end{aligned} \tag{24a}
$$

$$\xi' = (1 + C)\zeta n^a, \quad \zeta \equiv (\zeta^a \zeta^b \delta_{ab})^{1/2}, \quad \zeta \text{ real},$$

$$n^a n^b \delta_{ab} = +1, \quad n^a \text{ real}, \tag{24b}$$

$$
\begin{aligned}
C \equiv C^\alpha \theta_\alpha &+ \bar C_{\dot\beta} \bar\theta^{\dot\beta} + C_+ \theta_- + C^\mu \theta_\mu \\
&+ C^\alpha_- \theta_{\alpha -} + \bar C_{\dot\beta -} \bar\theta^{\dot\beta}_+ + C_{+-} \theta_{+-},
\end{aligned} \tag{24c}
$$

with

$$\xi^\mu, \quad \xi_{+-}, \quad C^\mu, \quad C_{+-} \quad \text{real}.$$

Moreover, $N_\parallel$ and $N_\perp$ turn out to be real elements too and the BCH-formula (21) still holds, with $A = -n \wedge N_\perp \times (1 + C)^{-1}$ being now a real Lie-valued element.

In this case, a natural question can be asked: *what is the global structure of the* $\mathrm{SSU}(2)_R$ *-variety*? Since $\mathrm{SSU}(2)_R$ descends from a compact group, whose compactness is strongly related to the quantization of the charges of the different matter gauge fields,[4] and since this is the unique unitary superextension of the bosonic SU(2) case, we believe it is relevant to try answering this question.

From Eq. (19) we notice that even if γ and δ are periodic functions of ζ, for a given $E \equiv C\zeta$, the coefficient of $\not{n}$ and $\not{N}$ shall contain a rational dependence upon ζ which does not allow to say whether the global structure of $\mathrm{SSU}(2)_R$ is of the type $K_{\text{bosonic}} \times R^{45}$ or, for instance $R^3 \times R^{45}$.

However suppose we take as independent variables $\zeta, n, E \equiv C\zeta$ and $M_\perp \equiv (1 + C)^{-1} N_\perp$. Then, according to Eq. (19):

$$
\begin{aligned}
\gamma &= \cos\zeta/2 \{ 1 - (2!)^{-1} 2^{-2} E^2 + (4!)^{-1} 2^{-4} E^4 \} \\
&\quad - \sin\zeta/2 \{ 2^{-1} E - (3!)^{-1} 2^{-3} E^3 \}, \\
\delta &= \sin\zeta/2 \{ 1 - (2!)^{-1} 2^{-2} E^2 + (4!)^{-1} 2^{-4} E^4 \} \\
&\quad - \sin\zeta/2 \{ 2^{-1} E - (3!)^{-1} 2^{-3} E^3 \}.
\end{aligned} \tag{25a}
$$

Since the vector A has the value

$$
\begin{aligned}
A = -n \wedge M_\perp \equiv -n \wedge \{ & m^\alpha_\perp \theta_\alpha + \bar m_{\dot\beta\perp} \theta^{\dot\beta} + m_{+\perp} \theta_+ \\
&+ \bar m_{+\perp} \theta_- + m^\mu_\perp \theta_\mu + m^\alpha_{\perp -} \theta_{\alpha -} \\
&+ \bar m_{\perp - \dot\beta} \bar\theta^{\dot\beta}_+ + m_{\perp +-} \theta_{+-} \},
\end{aligned} \tag{25b}
$$

any of the inner products appearing in Eq. (21) can be replaced by the corresponding inner products in the $M_\perp$ components. Therefore e^{Ξ}, Ξ real, using the ζ, n, E, $M_\perp$ parametrization can be cast in the form:

$$\Xi \text{ real}, \ e^{\Xi} = \gamma(\zeta, E) \mathbb{1} + i\delta(\zeta, E) \{ 1 + f(M_\perp) \not{n} + \not{M}_\perp \}, \tag{25c}$$

explicitly showing, because the periodic dependence of both $\gamma(\zeta, E)$ and $\delta(\zeta, E)$ in ζ, that for a given set $(n, E, M_\perp)$ gives all the different elements that (25c) can provide if ζ runs over a circumference S_1 of length 4π.

It is interesting to point out that $M_\perp$ fixed, E fixed can be obtained if both $N_\perp$ and ζ vary such that

$$M_\perp = N_{\perp 2} (1 + \zeta_2^{-1} E)^{-1} = N_{\perp 1} (1 + \zeta_1^{-1} E)^{-1}. \tag{26}$$

Equivalently, if $\zeta_2 = \zeta_1 + 4\pi$ the corresponding increment of $N_\perp$ keeping $M_\perp$ constant is

$$\Delta N_\perp \equiv N_{\perp 2} - N_{\perp 1} = -E M_\perp 4\pi \zeta_1^{-1} (\zeta_1 + 4\pi)^{-1}. \tag{27}$$

As both E and $M_\perp$ start with θ^1 terms, $\Delta M_\perp$ starts with θ^2 terms. In terms of the last set of parameters we have introduced, Ξ assumes the form:

$$\Xi = (\zeta + E)(n + M_\perp) = \zeta \{ (1 + C)n + N_\perp \}, \tag{28a}$$

where

$$\zeta + E = \Xi \cdot n, \quad M_\perp = \Xi (\Xi \cdot n)^{-1} - n. \tag{28b}$$

It is worth remarking that $E = 0 = \Xi \cdot n$ gives the supersymmetric generalization of the SU(2) rotation with axis n: $E = 0$ makes the elements e^{Ξ} to have the values

$$e^{\Xi}(\Xi \cdot n = 0) = \cos\zeta/2 \mathbb{1} + i \sin\{ 1 + f(M_\perp) \not{n} + \not{M}_\mu \}, \tag{29}$$

which clearly constitute an additive Abelian subgroup, homeomorphic to $S_1(0,4,\pi)$. They represent the superrotations of axis $n + M_\perp$.

In general one could say that E is the more interesting element in the analysis of the supergroups, which could be the source of nontrivial intrinsically supersymmetric properties which deserve further studies, specially in the case of Yang–Mills models.

It is also worthwhile to point out that there are similar in spirit but slightly different alternative parametrizations leading to BCH expressions too which are given in Appendix B for the sake of completeness.

IV. DISCUSSIONS AND COMMENTS

We have been able to obtain BCH-formulas for the more general supersymmetric extension of the local SU(2) gauge group, which is $SSU(2)_C$. Therefore and, as particular cases, BCH-formulas for the real (unitary) and chiral supergroups arising from SU(2) can be straightforwardly written down.

One of the methods exhibited in this article gave as a byproduct a canonical factorization for every element e^{Ξ} of $SSU(2)_C$ which, as Eq. (3) shows up can always be written as a "transverse" element times a "longitudinal" bosonic element times the inverse of the same "tranverse" element.

Then we studied the supersymmetric unitary supergroup $SSU(2)_R$, specially looking at the connections its global structure may have with the original bosonic SU(2) local gauge group, where we know the physical importance of its compactness.

That point of view led us to look at other parametrizations, like (28a), which splits up an element of the generating Lie algebra into a superscalar times a superaxis $n + M_\perp$, and where the bosonic elements of each factor belong to the parent SU(2) variety (a compact one) and the strictly supersymmetric components E and $M_\perp$ range unrestricted over R^{45}. These results confirm and extend the observations already made[6] for the chiral supergroups.

Finally, let us express that we conjecture that the BCH-formulas presented here shall be useful when dealing with supersymmetric Yang–Mills models, especially regarding their quantization, and choosing the appropriate gauges in the search of physically relevant symmetric topological configurations.

ACKNOWLEDGMENT

This work received partial support from the CONICIT (Consejo Nacional de Investigaciones Científicas y Tecnológicas, Venezuela), through Grant S1-972.

APPENDIX A

Given $C = c^\alpha\theta_\alpha + d_{\dot\beta}\bar\theta^{\dot\beta} + c_+\theta_+ + d_-\theta_- + c^\mu\theta_\mu + c^\alpha_-\theta_{\alpha-} + d_{\dot\beta+}\bar\theta^{\dot\beta}_+ + c_{+-}\theta_{+-}$,

it is straightforward to obtain $(1+c)^{-1}$

It turns out to be:

$$\begin{aligned}(1+C)^{-1} = {}& 1 - c^\alpha\theta_\alpha - d_{\dot\beta}\bar\theta^{\dot\beta} - (c_+ + 2^{-1}c^\alpha c_\alpha)\theta_+ - (d_- + 2^{-1}d_{\dot\beta}d^{\dot\beta})\theta_- - (c^\mu + c\sigma^\mu d)\theta_\mu \\ & + (2c^\alpha d_- + 2^{-1}c^\alpha dd + c^\mu\sigma_\mu{}^{\alpha\dot\beta} d_{\dot\beta} + 2^{-1}c\sigma_\mu dd_{\dot\beta}\sigma^{\mu\alpha\dot\beta} - c^\alpha_-)\theta_{\alpha-} + (2c_+ d_{\dot\beta} + 2^{-1}cc\, d_{\dot\beta} \\ & - c^\mu c^\alpha\sigma_{\mu\alpha\dot\beta} - 2^{-1}c\sigma^\mu dc^\alpha\sigma_{\mu\alpha\dot\beta} - d_{\dot\beta+})\bar\theta^{\dot\beta}_+ + (2^{-1}(cc)(dd) - 2^{-1}(c\sigma_\mu d)(c\sigma^\mu d) - 3\times2^{-1}c^\mu(c\sigma_\mu d) - 2^{-1}c^\mu c_\mu \\ & + 3\times2^{-1}c_+ dd + 3\times2^{-1}d_- cc + c_+ d_- + d_- c_+ - c^\gamma_- c_\gamma - d_{\dot\beta+}d^{\dot\beta} - c_{+-})\theta_{+-}.\end{aligned} \tag{A1}$$

In a shorter form:

$$(1+C)^{-1} = 1 - c^\alpha\theta_\alpha - d_{\dot\beta}\bar\theta^{\dot\beta} - h_+\theta_+ - k_-\theta_- - h^\mu\theta_\mu + h^\alpha_-\theta_{\alpha-} + k_{\dot\beta+}\bar\theta^{\dot\beta}_+ + h_{+-}\theta_{+-}, \tag{A2}$$

where $h_+; h_-, h_\mu \sim c^1 + ac^2; h^\alpha_-$ and $k_{\dot\beta+} \sim c^1 + a'c^2 + b'c^3$ and $h_{+-} \sim c^1 + \cdots + d''c^4$, by $c^1, c^2, \cdots, c^4$ terms homogeneous in the first power of C are meant, in the second power up to the fourth power in the coefficients of C.

The BCH expression (21) has been given in terms of $A = -n\wedge(1+c)^{-1}N_\perp$.

Replacement in this expression of $(1+c)^{-1}$ as above and of $N_\perp$ as obtained from Eq. (11) gives A the value:

$$\begin{aligned}A = {}& -n\wedge\{n^\alpha_\perp\theta_\alpha + m_{\dot\beta\perp}\bar\theta^{\dot\beta} + (n_{+\perp} + 2^{-1}c^\alpha n_{\alpha\perp})\theta_+ + (m_{-\perp} + 2^{-1}d_{\dot\beta}m^{\dot\beta}_\perp)\theta_- + (n^\mu_\perp + 2^{-1}c\sigma^\mu m_\perp + 2^{-1}n_\perp\sigma^\mu d)\theta_\mu \\ & + [n^\alpha_{-\perp} - c^\alpha m_{-\perp} - n^\alpha_\perp(d_- + 2^{-1}d_{\dot\beta}d^{\dot\beta}) - 2^{-1}(c^\mu + c\sigma^\mu d)\sigma_\mu{}^{\alpha\dot\beta}m_{\dot\beta\perp} - 2^{-1}n^\mu_\perp\sigma_\mu{}^{\alpha\dot\beta}d_{\dot\beta}]\theta_{\alpha-} \\ & + [m_{\dot\beta+\perp} - n_{+\perp}d_{\dot\beta} - m_{\dot\beta\perp}(c_+ + 2^{-1}c^\alpha c_\alpha) + 2^{-1}n^\alpha_\perp(c^\mu + c\sigma^\mu d)\sigma_{\mu\alpha\dot\beta} \\ & + 2^{-1}c^\alpha n^\mu_\perp\sigma_{\mu\alpha\dot\beta}]\bar\theta^{\dot\beta}_+ + ((1+c)^{-1}N_\perp)_{+-}\theta_{+-}\},\end{aligned} \tag{A3}$$

where, for instance, $n^\alpha_\perp$ is the projection of n^α in the subspace orthogonal to n and $((1+c)^{-1}N_\perp)_{+-}$ does not matter in the remaining results.

APPENDIX B

We shall sketch here alternative BCH-formulas for $SSU(2)_R$. We start from the representation[11] of Ξ:

$$\Xi = \zeta\{(1+N_\parallel)n + N_\perp\} = \Xi^a X_a \equiv S s^a X_a, \tag{B1}$$

where S is the modulus (supersymmetric) of Ξ^a and s^a the unitary vector defined by Ξ^a, that is,

$$S^2 \equiv \Xi^a\Xi^b\delta_{ab}, \quad s^a \equiv S^{-1}\Xi^a, \quad s^a s^b\delta_{ab} = +1. \tag{B2}$$

It is completely straightforward to show that:

$$e^{\Xi} = \cos(2^{-1}S) + i\sin(2^{-1}S)s^a\sigma_a. \tag{B3}$$

From Eq. (B1)

$$S^2 = \zeta^2(1+N_\parallel)^2\{1 + (1+N_\parallel)^{-2}N^2_\perp\}. \tag{B4}$$

Then, $S = (S^2)^{1/2}$ and $S^{-1} = (S^2)^{-1/2}$ can easily be calculated using the fact that $(1+N_\parallel)^{-2}N^2_\perp$ starts with θ^2:

$$S = \zeta + \zeta\{N_\parallel + 2^{-1}N^2_\perp(1+N_\parallel)^{-1} - 2^{-3}N^4_\perp\}, \tag{B5}$$

$$S^{-1}=\zeta^{-1}(1+N_{\parallel})^{-1}\{1-2^{-1}N_{\perp}^{2}(1+N_{\parallel})^{-2}+3\times 2^{-3}N_{\perp}^{4}\},\tag{B6}$$

$$\hat{s}=S^{-1}\Xi=\{1-2^{-1}N_{\perp}^{2}(1+N_{\parallel})^{-2}+3\times 2^{-3}N_{\perp}^{4}\}\times(n+N_{\perp}(1+N_{\parallel})^{-1}).\tag{B7}$$

From Eq. (B5) we have that:

$$2^{-1}S=2^{-1}\zeta+2^{-1}\zeta\{N_{\parallel}+2^{-1}N_{\perp}^{2}(1+N_{\parallel})^{-1}-2^{-3}N_{\perp}^{4}\}\equiv 2^{-1}\zeta+2^{-1}D,\tag{B8}$$

where between $N_{\parallel}$, $N_{\perp}$ and D, $P_{\perp}$ there is a bijective correspondence:

$$N_{\parallel}=\zeta^{-1}D(1-2^{-1}P_{\perp}^{2})-2^{-1}P_{\perp}^{2}+3\times 2^{-3}P_{\perp}^{4},\qquad P_{\perp}\equiv N_{\perp}(1+N_{\parallel})^{-1}.\tag{B9}$$

The unit vector $\hat{s}$ (B7) can be cast in the shorter form:

$$\hat{s}=\{1-2^{-1}P_{\perp}^{2}+3\times 2^{-3}P_{\perp}^{4}\}(n+P_{\perp}).\tag{B10}$$

As it also straightforward to prove that $\cos(a+b)$ and $\sin(a+b)$ follow the same rule for a,b Grassmannian elements than for $a\cdot b$ real, one has that Eq. (B3) can be written down, in terms of $\zeta,n,D,P_{\perp}$:

$$e^{\Xi}=\cos(2^{-1}\zeta)\cos(2^{-1}D)-\sin(2^{-1}\zeta)\sin(2^{-1}D)+i\{\sin(2^{-1}\zeta)\cos(2^{-1}D)+\cos(2^{-1}\zeta)\sin(2^{-1}D)\}\times\{1-2^{-1}P_{\perp}^{2}+3\times 2^{-3}P_{\perp}^{4}\}(n+P_{\perp}),\tag{B11}$$

which has the advantage, with respect to (B3) that the compactness in the bosonic variables appears evident $\cos(2^{-1}D)$ and $\sin(2^{-1}D)$ are very simple expressions polynomials due to the fact that D starts with θ^{1} and therefore $D^{5}=0$.

[1]S. Ferrara and B. Zumino, Nucl. Phys. B **79**, 413 (1974).

[2]J. Wess, GIFT Seminar Lectures, Salamanca 1977, M. Sohnius, Nucl. Phys. B **136**, 461 (1978).

[3]F. Mansouri, J. Math. Phys. **18**, 2395 (1977).

[4]C.N. Yang, Phys. Rev. D **1**, 2360 (1970).

[5]T.T. Wu and C.N. Yang, Phys. Rev. D **12**, 3843 (1975).

[6]C. Aragone and A. Restuccia, Lett. Math. Phys. **3**, 29–34 (1979).

[7]Throughout this paper we shall use Weyl dotted and undotted two-dimensional spinors $\theta_{+}\equiv\theta^{\alpha}\theta_{\alpha}$, $\theta_{-}\equiv\bar{\theta}_{\dot\beta}\bar{\theta}^{\dot\beta}=\epsilon_{\dot\alpha\dot\beta}\bar{\theta}^{\dot\alpha}\bar{\theta}^{\dot\beta}=(\theta_{+})^{+}$, $\theta_{\alpha-}\equiv\theta_{\alpha}\theta_{-}$, $\bar{\theta}^{\dot\beta}_{+}\equiv\bar{\theta}^{\dot\beta}\theta_{+}$, $\theta_{+-}\equiv\theta_{+}\theta_{-}$. With respect to SU(2), X_a are its real generators $[X_a,X_b]=-\epsilon_{abc}X_c$.

[8]Notice that we are using a bilinear inner product for three-dimensional complex vector, i.e., $u\cdot v\equiv u^{a}v^{b}\delta_{ab}$ for $u,v\in C^{3}$.

[9]$(Z\wedge\xi')^{c}=\epsilon_{abc}Z^{a}\xi'^{b}$.

[10]As $Z\cdot\xi'=Z\cdot\xi(1+c)=0$, then $Z\cdot\xi=(1+c)^{-1}Z\cdot\xi'=0$, too.

[11]The solution is completely straightforward and its derivation does not require any deep consideration.

Models for SU(N)$\times$SU(N) symmetry breaking with extremum constraints

O. Gherman and L. Saliu
Department of Physics, University of Craiova, Craiova, Romania

(Received 15 May 1979; accepted 27 July 1979)

The extremum properties of some functions defined on the compact groups or their representations are investigated. The extremum constraints for the vacuum expectation value of U(N)$\times$U(N) and SU(N)$\times$SU(N) $\square$ [Z_2(P)$\times$$Z_2$(C)] symmetry breaking perturbation Hamiltonian are used to determine the models compatible with nonnegative particle mass spectrum in the first perturbation order. Some model-independent properties inferred by extremum constraints of chiral and CP-symmetry breaking are also examined.

1. INTRODUCTION

In this paper we investigate the extremum properties of some functions defined on a compact group (representation). The functions are built up as real parts of the average on a fixed vector of a tensorial operator, having both well defined transformation properties with respect to the considered compact group.

In a particular form such a problem has been investigated within the framework of a $(3,3^*)\oplus(3^*,3)$-model of the SU(3)$\times$SU(3) breaking.

Some authors succeeded in giving the most general expression of the symmetry breaking Hamiltonian density compatible with an extremum principle, showing that the breaking of a Nambu–Goldstone realization of SU(3)$\times$SU(3) symmetry may be accompanied by a breakdown of CP-invariance.[1,2] The connection of the mentioned results with some observed CP-violating effects have also been investigated.[3]

In order to get the best insight into the subject we start by studying the extremum properties of some adequate functions in the general case of a compact group (representation) with special regard to U(N)$\times$U(N) and SU(N)$\times$SU(N). We hope that our approach will allow nontrivial applications to some models of large dimensions and reveal some model-independent properties inferred by extremum constraints.

This paper is divided as follows:

In Sec. 2 we formulate the extremum problem and get some general properties of the extremum solutions.

In Sec. 3 we show the connection of this problem with the Hamiltonian theories with a spontaneous breaking of the symmetry and we define the unitary (chiral) $(D,D^*)\oplus(D^*,D)$-models for the group (SU(N)$\times$SU(N))$\square$($Z_2(P)\times Z_2(C)$). The extremum problem is formulated in a suitable for a global approach matricial form via a matricial realization of the irreducible representations (I.R.) of the direct product of unitary groups. Matricial equations satisfied by the extremum solutions are determined.

An explicit form of the extremum solutions and their properties for the $(N,N^*)\oplus(N^*,N)$ models in the U(N)$\times$U(N) and SU(N)$\times$SU(N) cases are obtained in Secs. 4 and 5, respectively.

In the U(N)$\times$U(N) case the CP-symmetry breaking term is eliminated by the extremum condition; the same condition might infer a CP-symmetry breaking term with vacuum symmetry in the $\epsilon=0$ limit for the SU(N)$\times$SU(N) case.

In Sec. 6 we enumerate the physical implication of the previous results, like the Okubo–Mathur's[4] and Dashen's[1] extremum domains, the connection between the theory parameters existing for the spontaneous CP-symmetry breaking and the implications of the vacuum symmetry for the extremum solutions.

2. THE EXTREMUM PROBLEM

The operators (particularly the Hamiltonian densities) with well defined properties with respect to a group G are introduced by the usual covariance scheme.[5–8]

Given the (continuous) linear representations T_I of the group G:

$$T_I: G\mapsto \mathscr{L}(V_I),\quad I=1,2, \tag{1}$$

by bounded operators in the vectorial spaces V_I, we are interested in a linear map

$$F: V_1\mapsto \mathscr{L}(V_2) \tag{2}$$

equivalent with respect to G, i.e., satisfying the conditions

$$T_2(g)F(x)T_2(g^{-1})=F(x'(g)),$$
$$g\in G,\quad x, x'(g)\in V_1, \tag{3}$$

where

$$x'(g)\equiv T_1(g)x. \tag{3'}$$

V_I are considered unitary spaces with the scalar products (SP) denoted by $\langle -,- \rangle_I$. Let now $w\in V_2$ and $v\in V_1$ be fixed vectors. We build up the function φ: $G\mapsto R$ defined by

$$\varphi(g)\equiv \operatorname{Re}\langle w,F(v'(g))w\rangle_2. \tag{4}$$

We consider compact the Lie groups such that $g=g(\gamma)$ where $\gamma\equiv\{\gamma_a\}$ $a=1{:}n$ are the (real) parameters of the and the T_I representations may be taken unitary with respect to SP $\langle -,- \rangle_I$ in V_I.[5–8]

Instead of Eq. (4) we write

$$\varphi(\gamma)\equiv \operatorname{Re}\langle w,F(v'(\gamma))w\rangle_2, \tag{5}$$

where

$$v'(\gamma)\equiv T_1(g(\gamma))v\equiv T_1(\gamma)v. \tag{5'}$$

 0022-2488/80/051234-09$01.00

Given group G, mappings T_I, F and vectors v, w, our problem is the determination of the element $v'(\gamma_0)$on the G-orbit of v in such a way that

$$\varphi(\gamma_0) = \text{extremum of } \varphi(\gamma), \quad g(\gamma)\in G. \tag{6}$$

As, by hypothesis the function $\varphi(\gamma)$ is defined on a compact, it attains its extremum.

It is obvious that

$$\mathscr{F}(x) \equiv \langle w, F(x)w\rangle_2, \quad x\in V_1, \ w\in V_2 \tag{7}$$

is a conjugate-linear bounded functional on V_1, hence there is a unique $\xi\in V_1$ such that

$$\mathscr{F}(x) = \langle \xi, x\rangle_1, \quad \varphi(\gamma) = \text{Re}\mathscr{F}(v'(\gamma)). \tag{8}$$

The vector $\xi\in V_1$ may be easily identified if we write Eq. (7) in an orthonormal basis (o.n.b.) $\{e_i\}$ of V_1 with respect to SP $\langle \cdot,\cdot\rangle_1$.

By using the linearity of the F mapping we obtain

$$\mathscr{F}(x) = x_i^* \xi_i = \langle \xi, x\rangle_1, \tag{9}$$

where

$$\xi_i \equiv \langle w, F(e_i)w\rangle_2 \tag{10}$$

are the coordinates of x and ξ respectively in o.n.b. $\{e_i\}$ and the star denotes the complex conjugation.

Consequently we obtain

$$\varphi(\gamma) = \text{Re}\langle \xi, v'(\gamma)\rangle_1. \tag{11}$$

If γ_0 is an extremum point of $\varphi(\gamma)$, then the function

$$\phi(\gamma) \equiv \text{Re}\langle \xi, T_1(\gamma)v_0\rangle_1 = \text{Re}\langle T_1^+(\gamma)\xi, v_0\rangle_1 \tag{12}$$

[where

$$v_0 \equiv v'(\gamma_0) = T_1(\gamma_0)v] \tag{13}$$

attains its extremum for $\gamma = 0$; hence we obtain the necessary extremum conditions

$$\left.\frac{\partial \phi(\gamma)}{\partial \gamma_a}\right|_{\gamma=0} = \text{Re}\langle \xi, \mathscr{T}_a v_0\rangle_1 = \text{Re}\langle \mathscr{T}_a^+ \xi, v_0\rangle_1, \tag{14}$$

where

$$\mathscr{T}_a \equiv \left.\frac{\partial T_1(\gamma)}{\partial \gamma_a}\right|_{\gamma=0}, \quad a = 1{:}n \tag{15}$$

are the generators of the T_1-representation of G. Corresponding to the kind of the extremum we shall require that the matrix of the second derivatives of ϕ

$$\left.\frac{\partial^2 \phi(\gamma)}{\partial \gamma_a \partial \gamma_b}\right|_{\gamma=0} = \text{Re}\langle \xi, \mathscr{T}_a \mathscr{T}_b v_0\rangle_1 \tag{16}$$

be positive or negative (semi) definite.

We can reformulate the same problem by using the real orthogonal representation associated with the (complex) unitary representation T_1. This second approach is advantageous when we want to deal with Hemitian operatorial images, for intance, in the Hamiltonian theories.

Let $T_1: G \mapsto \mathscr{L}(V_1)$ be a D-dimensional G-representation (unitary with respect to the SP $\langle \cdot,\cdot\rangle_1$) and let $\{e_i\}$ be an o.n.b. (with respect to the same SP) in V_1. Then

$$\{f_j : f_j = e_j \text{ if } j \leqslant D, \quad f_j = ie_{j-D} \equiv d'_{j-D} \quad \text{if } j > D\}_{j=1:2D} \tag{17}$$

is an o.n.b. with respect to SP

$$(x_R, y_R) \equiv \text{Re}\langle x, y\rangle_1 \tag{18}$$

in the real $2D$-dimensional space V_{1R} associated with V_1. Here we read $x,y\in V_1$ and $x_R, y_R \in V_{1R}$ with

$$x_R = (\text{Re}x_i)\, e_i + (\text{Im}x_i) e'_i = x_{R_j} f_j. \tag{19}$$

The linear real $2D$-dimensional representation $T_{1R}: G \mapsto \mathscr{L}(V_{1R})$ defined by

$$T_{1R}(\gamma)x_R = (T_1(\gamma)x)_R = x'_R(\gamma) \tag{20}$$

is orthogonal with respect to the SP(18). Let T_1 be the representation contragradient to T_1 with $T_1^*(g) = (T_1(g))^*$. It is simple to show that T_{1R} is complex equivalent to $T_1 \oplus T_1^*$, the direct sum of representations T_1 and T_1^*. Hence T_{1R} representation is irreducible as a real one. Furthermore, with $F: V_1 \mapsto \mathscr{L}(V_2)$ we should associated the linear map $F_R: V_{1R} \mapsto \mathscr{L}(V_2)$ with all $F_R(x_R)$ Hermitian defined by the law

$$F_R(x_R) \equiv \text{ the Hermitian part of } F(x) = \tfrac{1}{2}[F(x) + F(x)^+]. \tag{21}$$

The from Eqs. (5), (11), (17)–(21) we obviously obtain

$$\varphi(\gamma) \equiv \text{Re}\langle w, F(v'(\gamma))w\rangle_2 = \text{Re}\langle \xi, v'(\gamma)\rangle_1 = (\xi_R, v'_R(\gamma)) = (\xi_R, T_{1R}(\gamma)v_R) = \langle w, F_R(v'_R(\gamma))w\rangle_2 \tag{22}$$

and the necessary extremum conditions (14) for $\varphi(\gamma)$ in the point γ_0 become

$$(\xi_R, \mathscr{T}_a^{(R)} v_{0R}) = (\mathscr{T}_a^{(R)+} \xi_R, v_{0R}) = 0, \tag{23}$$

where

$$\mathscr{T}_a^{(R)} \equiv \left.\frac{\partial T_{1R}(\gamma)}{\partial \gamma_a}\right|_{\gamma=0} \tag{24}$$

are the generators of the T_{1R}-representation with the action on V_{1R} obvious from (20),

$$\mathscr{T}_a^{(R)} x_R = (\mathscr{T}_a x)_R. \tag{25}$$

By choosing the parametrization γ in such a way that the generators $\mathscr{T}_a, (\mathscr{T}_a^{(R)})$, are anti-Hermitian, (anti-symmetric), from (14) [(23)] we immediately read that the extremum solution v_{0R} belonging to the G-orbit of v_R is orthogonal [with respect to the (SP) (18)] to the (real) linear subspace $V_{\xi_R} \subset V_{1R}$ spanned by the vectors $\{\mathscr{T}_a^{(R)} \xi_R\}$.

This property may be obtained in a different way. Let $\mathscr{O}_{\xi_R}$ be the G-orbit of $\xi_R \in V_{1R}$:

$$\mathscr{O}_{\xi_R} \equiv \{T_{1R}(\gamma)\xi_R\} \subset S_{2D-1}^{\xi_R}, \tag{26}$$

where $S_{2D-1}^{\xi_R}$ is the sphere

$$S_{2D-1}^{\xi_R} \equiv \{y_R \in V_{1R} : \|y_R\| = \|\xi_R\|\}, \quad \|y_R\| \equiv (y_R, y_R)^{1/2}. \tag{27}$$

(Obviously we can take from the beginning $\|\xi_R\| = 1$ or $\|v_R\| = 1$.) In the previous conditions we showed that the extreumum solution $v_{0R} = T_{1R}(\gamma_0)v_R$, (ξ_R), is orthogonal to V_{ξ_R}, $(V_{v_{0R}})$, the tangent space to orbit $\mathscr{O}_{\xi_R}$, $(\mathscr{O}_{v_{0R}})$, in the "point" γ_0 i.e., in $\xi_R \in S_{2D-1}^{\xi_R}$, $(v_{0R} \in S_{2D-1}^{v_R})$. If $\mathscr{O}_{v_R}$ would cover the sphere $S_{2D-1}^{v_R}$ (we suppose $\|v_R\| = 1$) it is well known that

$$\max(\xi_R, T_{1R}(\gamma)v_R) = \|\xi_R\|, \tag{28}$$

the maximum (minimum) being reached for v_{0R} parallel (antiparallel) to ξ_R; this means that v_0, (v_{0R}) and implicity the operator $F(v_0)$, $(F_R(v_{0R}))$ have the same stabilizer G as ξ, (ξ_R). In this case the tangent space V_{ξ_R} in γ_0, orthogonal to ξ_R is $(2D-1)$-dimensional and as $v_{0R} \in V_{\xi_R}^{\perp}$ we get the same conclusion.

Such a situation would occur for instance for the identical N-dimensional representation T of the unitary group $G = \mathrm{U}(N)$ where[6]

$$T(U) = U \in \mathrm{U}(N). \tag{29}$$

Nevertheless, in the general case the G-orbits do not cover the sphere, $\dim V_{\xi_R} < 2D - 1$ and getting the extremum is much more involved.

As soon as v_0 is determined, we get the operator $F(v_0)$ with the desired properties:

$$\begin{aligned}\mathrm{Re}\langle w, F(v_0)w\rangle_2 &= \langle w, F_R(v_{0R})w\rangle_2, \\ &= \text{extremum}\langle w, F_R(T_{1R}(g)v_R)w\rangle_2 \\ &= \text{extremum } \mathrm{Re}\langle w, F(T_1(g)v)w\rangle_2, \quad g \in G.\end{aligned} \tag{30}$$

3. THE MODEL

A similar case to the above formulated extremum problem appears in the Hamiltonian theories with a degenerate fundamental (lowest energy) state.[1,9]

Let G be the fundamental symmetry group (of the dominant interaction) of the theory.

The Hamiltonian operators with well defined properties with respect to a group G is introduced by the usual covariance scheme. Now consider that the fundamental G-symmetry is broken by two mechanisms.

We define the intrinsic breaking of the G-symmetry by taking the Hamiltonian density of the form

$$\mathscr{H} = \mathscr{H}_0 + \epsilon\mathscr{H}_B,$$

where $\mathscr{H}_0$ (dominant interaction) is invariant with respect to G and $\mathscr{H}_B$, in the breaking G-invariance perturbation $\epsilon\mathscr{H}_B$, is the Hermitian operatorial image by a map F_R of some vector $v_R \in V_{1R}$

$$\mathscr{H}_B = F_R(v_R). \tag{31}$$

On the other hand, we define the spontaneous breaking of G supposing that in the $\epsilon = 0$ limit the fundamental state of $\mathscr{H}_0$, the vacuum $w \in V_2$ is G'-symmetric where G' is a proper subgroup of G, i.e.,

$$T_2(g')w = w \quad \text{for all} \quad g' \in G'. \tag{32}$$

As $\mathscr{H}_0$ is G-invariant, the fundamental state of $\mathscr{H}_0$ is degenerate, $T_2(g)w$ ($g \in G$), being also fundamental states.

Since $\mathscr{H}_B$ determines which of the vacua of $\mathscr{H}_0$ is the relevant one [namely $w \equiv \lim_{\epsilon \to 0} w(\epsilon)$ where $w(\epsilon)$ is the unique vacuum state of the operator $\mathscr{H}$] it follows that the G'-group is not independent of $\mathscr{H}_B$.[1]

Conversely, given the vacuum w, out of many

$$\begin{aligned}\mathscr{H}'_B(g) &\equiv F_R(v'_R(g)) = F_R(T_{1R}(g)v_R) \\ &= T_2(g)F_R(x_R)T_2(g^{-1}) = T_2(g)\mathscr{H}_B T_2(g^{-1})\end{aligned} \tag{33}$$

which can be obtained from each other by G-transformations we must select the one $\mathscr{H}'_{B0} \equiv \mathscr{H}'_B(g_0)$ which extremizes the energy,[1] namely,

$$\begin{aligned}\epsilon\varphi(g_0) &\equiv \langle w, \epsilon\mathscr{H}'_B(g_0)w\rangle_2 = \min\langle w, \epsilon\mathscr{H}'_B(g)w\rangle_2 \\ &= \min \epsilon\varphi(g), \quad g_0, g \in G;\end{aligned} \tag{34}$$

or with the parametrization $g(\gamma)$ [see Eq. (5)]

$$\epsilon\varphi(\gamma_0) = \min \epsilon\varphi(\gamma), \quad g(\gamma) \in G. \tag{34'}$$

with $\epsilon\mathscr{H}'_B(g_0)$ as a perturbation term, this condition assures nonnegative masses in the first order of perturbation theory.[1]

Now we analyze the extremum problem previously formulated when the fundamental symmetry group of the models is $G \equiv (\mathrm{SU}(N) \times \mathrm{SU}(N)) \square (Z_2(P) \times Z_2(C))$. The symbols $\times$ and $\square$ describe standard operations of direct and semidirect products respectively.[5–8]

The $Z_2(P) \times Z_2(C)$ group is constructed on the basis of the outer automorphisms C and P associated in the unitary models to the discrete transformations of charge conjugation and space reflexion respectively, each generating a cyclic group of second order, whilst $CP = PC$.[4,5]

The discrete transformations C and P induce on $\mathrm{SU}(N) \times \mathrm{SU}(N)$ the following outer automorphisms

$$(U,V) \overset{C}{\mapsto} (V^*, U^*), \quad (U,V) \overset{P}{\mapsto} (V, U), \quad (U,V) \underset{PC}{\overset{CP}{\mapsto}} (U^*, V^*) \tag{35}$$

where $U, V \in \mathrm{SU}(N)$ hence $(U,V) \in \mathrm{SU}(N) \times \mathrm{SU}(N)$. If T is a representation of $\mathrm{SU}(N) \times \mathrm{SU}(N)$ and X is an automorphism of $\mathrm{SU}(N) \times \mathrm{SU}(N)$, then $(U,V) \mapsto T(X(U,V)) \equiv T^x(U,V)$ is a representation of $\mathrm{SU}(N) \times \mathrm{SU}(N)$. In this way we define the action of discrete transformations on the $\mathrm{SU}(N) \times \mathrm{SU}(N)$ irreducible representations (IR) set.

By denoting an IR of $\mathrm{SU}(N) \times \mathrm{SU}(N)$ by the IR's dimensions of both $\mathrm{SU}(N)$ groups we have

$$\begin{aligned}&(D_1, D_2)^C = (D_2^*, D_1^*), \quad (D_1, D_2)^P = (D_2, D_1), \\ &(D_1, D_2)^{CP} = (D_1^*, D_2^*),\end{aligned} \tag{36}$$

where D^* is the representation contragradient to D. In a theory containing $\mathrm{SU}(N)^d$ invariants where $\mathrm{SU}(N)^d$ is the diagonal $\mathrm{SU}(N)$ subgroup of $\mathrm{SU}(N) \times \mathrm{SU}(N)$, we have to operate with representations of the type T_{1R} taking the identification

$$V_{1R} \equiv (D, D^*) \oplus (D^*, D) \tag{37}$$

where D is a unitary D-dimensional IR of $\mathrm{SU}(N)$. We saw that this is an IR as a real representation of $\mathrm{SU}(N) \times \mathrm{SU}(N)$ (or as a representation of G) associated to the unitary (complex) D^2-dimensional IR $T_1 : \mathrm{SU}(N) \times \mathrm{SU}(N) \mapsto \mathscr{L}(V_1 \equiv (D, D^*))$ of $\mathrm{SU}(N) \times \mathrm{SU}(N)$.

The breaking G-invariance $\mathscr{H}_B$ is the Hermitian operatorial image by a map F_R of a vector in $(D, D^*) \oplus (D^*, D)$, i.e.,

$$\mathscr{H}_B = F_R(v_R), \quad v_R \in (D, D^*) \oplus (D^*, D). \tag{38}$$

The (D, D^*) representation of $\mathrm{SU}(N) \times \mathrm{SU}(N)$ can be conveniently handled on the D^2-dimensional complex space $\mathscr{M}^{D^2}$ of $D \times D$ complex matrices. On a matrix $x \in \mathscr{M}^{D^2}$, a $g \in \mathrm{SU}(N) \times \mathrm{SU}(N)$ transformation, described by the pair $g \equiv (U, V)$ of unitary unimodular $N \times N$ matrices, acts as

$$x \mapsto x'(g) = x'(U,V) \equiv T_1(U,V)x = t(U)xt(V^+) = t(U)xt^+(V), \tag{39}$$

U and V being represented by $t(U)$ and $t(V)$ in the D-representation space of SU(N). The T_1 - representation of SU(N)$\times$SU(N) is a unitary one with respect to the following scalar product in $\mathscr{M}^{D^2}$:

$$\langle x_1, x_2\rangle \equiv \tfrac{1}{2}\,\mathrm{Tr}\,x_1 x_2^+.$$

In $\mathscr{M}^{D^2}$ we introduce an o.n.b. (with respect to this SP) consisting of D^2 Hermitian matrices

$$\{\Lambda_j^{(D)}\}_{j=1:D^2}. \tag{40}$$

Then $\{\Lambda_j^{(D)}, \Lambda_j^{(D)'} \equiv i\Lambda_j^{(D)}\}$ is an o.n.b. in the $2D^2$-dimensional real space $\mathscr{M}_R^{D^2}$ of $D\times D$ complex matrices with the SP (18),

$$(x_{1R}, x_{2R}) = \mathrm{Re}\langle x_1, x_2\rangle = \tfrac{1}{4}\,\mathrm{Tr}(x_1x_2^+ + x_2x_1^+). \tag{41}$$

In the above considered bases we have

$$v = (c_j + id_j)\Lambda_j^{(D)}, \tag{42a}$$

$$v_R = c_j\Lambda_j^{(D)} + d_j\Lambda_j^{(D)'} \tag{42b}$$

[Summation over repeated (dependent on D) j-indices is always implied.]

We observe that the (real) subspaces of the Hamiltonian and anti-Hermitian matrices, $\mathscr{E}$ and $\mathscr{A}$ respectively, from $\mathscr{M}^{D^2}$ are invariant with respect to the restriction of representation T_1, defined in Eq. (39) to SU(N)$^d \equiv \{(U,U): U \in$ SU(N)$\}$ and CP operation acts on $\mathscr{M}_R^{D^2}$ as a Hermitian conjugation, therefore the Hermitian and anti-Hermitian parts of $x \in \mathscr{M}_R^{D^2}$ are even and odd, respectively under CP.

We suppose that $(D,D^*) \oplus (D^*,D)$ is reduced up to IR of the SU(N)^d subgroup of SU(N)$\times$SU(N) and that in Eq. (42a) the vector v is decomposed after the even and odd IR's (with respect to CP) of SU(N)^d contained in the $(D,D^*) \oplus (D^*,D)$ representation. In this sense we observe that under SU(N)^d, Λ_j and $\Lambda_j' \equiv i\Lambda_j$ transform in the same way and by the previous assumption, according to IR's of SU(N)^d contained in $(D,D^*) \oplus (D^*,D)$. Then from Eqs. (38) and (42) emerges

$$\mathscr{H}_B = F_R(v_R) = c_ju_j + d_jv_j, \tag{43}$$

where u_j and v_j are the Hermitian operatorial images of Λ_j and Λ_j' respectively,

$$u_j \equiv F_R(\Lambda_j), \quad v_j \equiv F_R(\Lambda_j'). \tag{44}$$

Hence the CP and SU(N)$\times$SU(N)-properties of Λ_j and Λ_j' are transferred to u_j and v_j respectively:

$$u_j \overset{\mathrm{CP}}{\mapsto} u_j, \quad v_j \overset{\mathrm{CP}}{\mapsto} -v_j, \tag{45}$$

$$\mathscr{H}_B \overset{(U,V)}{\mapsto} \mathscr{H}_B'(U,V) = F_R(v_R'(U,V)) = c_j'u_j + d_j'u_j, \tag{46}$$

where c_j' and d_j' are defined by

$$\begin{aligned} v'(U,V) &\equiv T_1(U,V)\,v = (c_j' + id_j')\Lambda_j, \\ v_R'(U,R) &\equiv T_{1R}(U,V)v_R = c_j'\Lambda_j + d_j'\Lambda_j'. \end{aligned} \tag{47}$$

Therefore, the terms d_jv_j in $\mathscr{H}_B$ and $d_j'v_j$ in $\mathscr{H}_B'$ violate the CP-invariance.

Now we define the spontaneous breaking of SU(N)$\times$SU(N) supposing that in the $\epsilon = 0$ limit, the vacuum w is $G_\xi \equiv Q\,\square(Z_2(C)\times Z_2(P))$-symmetric, where $Q \subset$ SU(N)^d. Then the only vacuum matrix elements different from zero in this limit are

$$\langle w, u_{j_s} w\rangle_2 \equiv \xi_{j_s}, \tag{48}$$

where u_{j_s} are Q-singlets, hence the corresponding Λ_{j_s} and the associated vector ξ [see Eq. (10)] satisfy:

$$\begin{aligned} t(u)\Lambda_{j_s}^{(D)}t(u^+) &= \Lambda_{j_s}^{(D)}, \\ t(u)\xi t(u^+) &= \xi, \quad u \in Q \subset \mathrm{SU}(N)^d, \end{aligned} \tag{49}$$

in this case ξ being the Hermitian matrix

$$\xi \equiv \xi_{j_s}\Lambda_{j_s}. \tag{50}$$

The subgroup Q of SU(N)$\times$SU(N) in the stabilizer G_ξ of the vacuum in the $\epsilon = 0$ limit can be used to classify particle states. When ϵ is zero we have Q-multiplets of Goldstone bosons.[1,4] Let $\{\alpha_a\}_{a=1:N^2-1} \equiv \alpha$ be now a (real) parametrization of SU(N) and let

$$Q_a^{(N)'} \equiv \left.\frac{\partial U(\alpha)}{\partial \alpha_a}\right|_{\alpha=0} \tag{51}$$

be a basis of the Lie algebra su(N) of SU(N) with

$$[Q_a^{(N)'},\ Q_b^{(N)'}] = f_{abc}^{(N)}Q_c^{(N)'}. \tag{52}$$

We associate with α_a's the generators $Q_a^{(D)'}$'s of D-dimensional IR t of SU(N):

$$Q_a^{(D)'} \equiv \left.\frac{\partial t(\alpha)}{\partial \alpha_a}\right|_{\alpha=0}, \quad [Q_a^{(D)'},\ Q_b^{(D)'}] = f_{abc}^{(N)}Q_c^{(D)'}. \tag{53}$$

We can take such a parametrization that $\{Q_a^{(N)'}\}$, ($\{Q_a^{(N)} \equiv -iQ_a^{(N)'}\}$), be an o.n.b. of anti-Hermitian (Hermitian), traceless matrices in SU(N) and

$$(U(\alpha))^+ = U(-\alpha), \quad (t(\alpha))^+ = t(-\alpha) \tag{54}$$

so that $Q_a^{(D)'}$ ($Q_a^{(D)} \equiv -iQ_a^{(D)'}$), are anti-Hermitian (Hermitian), traceless $D\times D$ matrices. [In the U(N) case we should renounce the traceless condition for the matricial Lie algebra elements.] Then for $g \equiv (U,V) \in$ SU(N)$\times$SU(N) we obtain the parametrization

$$g(\gamma) = (U(\alpha), V(\beta)), \text{ i.e., } \gamma \equiv (\alpha,\beta). \tag{55}$$

Let us return not to the previously formulated extremum problem. We are interested in finding on the G-orbit of v_R, (v), an element

$$\begin{aligned} v_{0R} &\equiv T_{1R}(U_0,V_0)v_R = T_{1R}(\gamma_0)v_R, \\ v_0 &\equiv T_1(U_0,V_0)v = T_1(\gamma_0)v, \end{aligned} \tag{56}$$

such that

$$\begin{aligned} \varphi(U_0,V_0) &= \langle w, \mathscr{H}_B'(U_0,V_0)w\rangle = \langle w, F_R(v_{0R})w\rangle = \text{extremum}\ \varphi(U,V) \\ &= \text{extremum}\langle w, \mathscr{H}_B'(U,V)w\rangle \\ &= \text{extremum}\langle w, F_R(v_R'(U,V))w\rangle, \end{aligned} \tag{57}$$

$$\varphi(\gamma_0) = \text{extremum}\ \varphi(\gamma), \tag{57'}$$

where, for our model

$$\varphi(U,V) \equiv \mathrm{Tr}[t(U)vt(V^+)\xi + \xi t(V)v^+t(U^+)],$$

$U, V \in SU(N),$ (58)

$$\varphi(\gamma) \equiv \varphi(\alpha,\beta) = \mathrm{Tr}[t(\alpha)vt(-\beta)\xi + \xi t(\beta)v^{+}t(-\alpha)], \quad U(\alpha), \quad V(\beta) \in SU(N). \tag{58'}$$

According to Eqs. (46) and (47) the Hamiltonian extremizing the vacuum expectation value has the structure

$$\mathscr{H}'_{B0} \equiv \mathscr{H}'_B(U_0,V_0) = F_{1R}(v_{0R}) = c_{0j}u_j + d_{0j}v_j, \tag{59}$$

where c_{0j} and d_{0j} are defined by

$$v_{0R} = c_{0j}\Lambda_j^{(D)} + d_{0j}\Lambda_j^{(D)'},$$

$$(v_0 = (c_{0j} + id_{0j})\Lambda_j^{(D)}). \tag{60}$$

As the function defined in Eq. (12) is now

$$\phi(\alpha,\beta) \equiv \mathrm{Re\,Tr}\, t(\alpha)v_0 t(-\beta)\xi$$

and attains its extremum for $\alpha = \beta = 0$, Eqs. (14) and (23) take the form

$$\mathrm{Tr}Q_a^{(D)'}(v_0\xi - \xi v_0^{+}) = \mathrm{Tr}Q_a^{(D)'}(\xi v_0 - v_0^{+}\xi) = 0, \quad a = 1{:}N^2 - 1. \tag{61}$$

We emphasize that the matricial form of the extremum condition is suggestive and has all the advantages of a global approach. By splitting the matrix v_0 in its Hermitian (C) and anti-Hermitian (D') components

$$v_0 = C + D' = C + iD, \tag{62}$$

where D is Hermitian, From Eqs. (61) we get

$$\mathrm{Tr}(Q_a^{(D)'}[C,\xi]) = \mathrm{Tr}(Q_a^{(D)'}\{D',\xi\}) = 0, \quad a = 1{:}N^2 - 1. \tag{63}$$

Let us denote by $\mathscr{A}$ and $\mathscr{E}$ the real subspaces of all anti-Hermitian and Hermitian $D \times D$ matrices respectively. These subspaces are orthogonal with respect to the SP (41) and $\mathscr{M}_R^{D^2} = \mathscr{A} \oplus \mathscr{E}$. Let it be now $\mathscr{A}_1 \subset \mathscr{A}$ ($\mathscr{E}_1 \subset \mathscr{E}$), the subspace generated by the generators $\{Q_a^{(D)'}\}$, $(\{-iQ_a^{(D)'} \equiv Q_a^{(D)}\})$, and $\mathscr{A}_2 \subset \mathscr{A}$ ($\mathscr{E}_2 \subset \mathscr{E}$), its orthogonal complements in $\mathscr{A}$ ($\mathscr{E}$), with repect to the same SP. Then according to Eqs. (63)

$$[C,\xi] \in \mathscr{A}_2 \quad \{D',\xi\} \in \mathscr{A}_2 \tag{64}$$

$$(i[C,\xi] \in \mathscr{E}_2 \quad \{D,\xi\} \in \mathscr{E}_2). \tag{65}$$

The subspaces $\mathscr{A}$ and $\mathscr{E}$ are invariant under the restriction of the T_{1R} representation to $SU(N)^d$, i.e., when $U = V$ in Eq. (39). Similarly, from the commutation relations (53) for $Q_a^{(D)'}$ it comes out that $\mathscr{A}_1$ and $\mathscr{E}_1$ (therefore also the orthogonal complements $\mathscr{A}_2$ and $\mathscr{E}_2$) are $SU(N)^d$-invariant.

As ξ is Q-invariant and $Q \subset SU(N)^d$, from Eqs. (56)–(58') we deduce that v_0 being a stationary solution, then $T(u)v_0T^{-1}(u)$, $(u \in Q)$ is also stationary because an arbitrary Q-transformations does not change the trace in $\varphi(U_0,V_0)$. This property could be used in order to simplify the solution v_0 and is compatible with Eqs. (63), (64), and (65).

The Eqs. (63) get the following symmetric algebraic form

$$f_{mnp}^{(D)}q_m^{(a)}h_nc_p = 0, \quad d_{mnp}^{(D)}q_m^{(a)}h_nd_p = 0 \tag{66}$$

where $\{q_m^{(a)}\}$, $\{h_n\}$, $\{c_p\}$, and $\{d_p\}$ are the coordinates in the basis $\{\Lambda_j^{(D)}, \Lambda_j^{(D)'}\}$ of the matrices $Q_a^{(D)'}$, ξ, C and D', (D), respectively and $f_{mnp}^{(D)}$ and $d_{mnp}^{(D)}$ are components proportional to those of the usual anti-symmetric and symmetric tensors.[5]

4. THE $(N,N^*) \oplus (N^*N)$ MODEL OF $U(N) \times U(N)$

We now give explicitly the extremum solutions for the $T_{1R} \equiv (N,N^*) \oplus (N^*,N)$ models of $U(N) \times U(N)$ and $SU(N) \times SU(N)$ with N and N^* the mutually contragradient fundamental representations of minimal dimensions, given up to equivalence by the maps

$$t(U) = U, \quad t^*(U) = U^*, \quad U \in U(N) \text{ or } SU(N). \tag{67}$$

In the case when the unitary groups are $U(N)$, the Eqs. (39), (58) become

$$v'(U,V) = T_1(U,V)v = UvV^{+}, \tag{68}$$

$$\varphi(U,V) = \mathrm{Tr}(UvV^{+}\xi + \xi Vv^{+}U^{+}), \quad U, \ V \in U(N), \tag{69}$$

where $v \in (N,N^*) + (N^*,N)$ is a complex $N \times N$ matrix.

In what follows it is shown that whenever $\varphi(U_0,V_0) =$ extremum $\varphi(U,V)$, then (i) both of the matrices $U_0vV_0^{+}\xi$ and $\xi U_0vV_0^{+}$ are positive (negative) definite for the maximum (minimum) of $\varphi(U,V)$; (ii) $\max\varphi(U,V) = m_ih_i$ and $\min\varphi(U,V) = -m_ih_i$ where $\{m_i\}$ and $\{h_i\}_{i=1:N}$ are all ordered (say in nonincreasing order) eigenvalues of the unique positive definite square roots of vv^{+} and $\xi\xi^{+}$, respectively.

In order to prove these properties we begin by studying the maximum of the function

$$f(\alpha) \equiv \mathrm{Re\,Tr}v\,U(\alpha) = \mathrm{Tr}(v\,U(\alpha) + U^{+}(\alpha)v^{+}) = \mathrm{Re\,Tr}\,U(\alpha)v, \quad U(\alpha) \in U(N). \tag{70}$$

We simply obtain the necessary extremum conditions

$$\mathrm{Tr}Q_a^{(N)'}(v_0 - v_0^{+}) = 0, \tag{71}$$

where

$$f(\alpha_0) = \max f(\alpha), \quad v_0 \equiv vU(\alpha_0) \equiv vU_0 \tag{72}$$

and

$$\{Q_a^{(N)'}\}_{a=1:N^2}$$

are the generators of $U(N)$ group. As $\{Q_a^{(N)'}\}$ is a basis in $\mathscr{M}^{N^2}$, from Eq. (71) we obtain that v_0 is Hermitian.

Furthermore, v_0 is positive definite and as $v_0^2 = vv^{+}$ we conclude that $v_0 = vU_0 = |v|$ where $|v|$ is the unique positive square root of vv^{+}. Let $v = |v|\mathscr{U}_0$ be the polar decomposition of v, where $\mathscr{U}_0 \in U(N)$. Then

$$f(\mathscr{U}_0^{+}) = \mathrm{Tr}|v| = \sum_i m_i, \tag{73}$$

where $\{m_i\}$ are the eigenvalues of $|v|$. It is easy to see that

$$\mathrm{Re\,Tr}\,|v|U \leqslant \mathrm{Tr}|v|, \quad U \in U(N),$$

the equality holding iff $U = \mathbb{1}$. Therefore, $U_0 = \mathscr{U}_0^{+}$. Obviously if $\mathrm{Re\,Tr}\,vU'_0 = \min f(U)$, then $vU'_0 = -|v|$, $U'_0 = -U_0$ because $f(-U) = -f(U)$ and $\min f(U) = -\max(-f(U)) = -\max f(U)$. Let us consider now the function in Eq. (69) and take U_0,V_0 such that

$$\varphi(U_0,V_0) = \max\varphi(U,V), \quad U,V \in U(N). \tag{74}$$

As

$$\mathrm{Re\,Tr}\,U_0vV_0^{+}\xi W = \mathrm{Re\,Tr}\,WU_0vV_0^{+}\xi \leqslant \varphi(U_0,V_0), \quad W \in U(N)$$

and

$$\mathrm{Re\,Tr}\,U_0vV^{+}\xi = \mathrm{Re\,Tr}\xi U_0vV^{+}$$

$\leqslant \varphi(U_0, V_0), \quad V \in \mathrm{U}(N),$

we conclude that $U_0 v V_0^+ \xi$ and $\xi U_0 v V_0^+$ respectively, are Hermitian positive definite matrices. (We will suppose v and ξ nonsingular matrices but, by using some well known spectral theorems about the normal operators,[8] the following conclusions may be obtained in the general case). Now let U_3 be such that $U_0 v V_0^+ U_3 = |U_0 v V_0^+|$. By introducing the notations $A \equiv |U_0 v V_0^+|$ and $B \equiv U_3^+ \xi$ we see that $AB = U_0 v V_0^+ \xi$ and $BA = U_3^+ \xi U_0 v V_0^+ U_3$ are Hermitian positive definite matrices and as A is Hermitian positive we deduce that A and B commute. On the other hand $(A, |v|)$ and $(B, |\xi|)$ respectively have the same eigenvalues with the same multiplicities, therefore there is a permutation P so that $\{m_i h_{P(i)}\}$ are the eigenvalues of AB where $\{m_i\}$ and $\{h_i\}$ are the ordered (say in nonincreasing order) eigenvalues of $|v|$ and $|\xi|$ respectively. Let U_4 be the (unitary) operator defined by $U_4 \zeta_i = \zeta_{P(i)}$ where $\{\zeta_i\}$ are the eigenvectors for both A and B, $A\zeta_i = m_i \zeta_i$ and $B\zeta_i = h_{P(i)}\zeta_i$. Then we have

$$A U_4 B U_4^{-1} \zeta_{P(i)} = m_{P(i)} h_{P(i)} \zeta_{P(i)}$$

and as

$$\mathrm{Tr} A U_4 B U_4^{-1} = \mathrm{Tr} U_4^{-1} U_0 v V_0^+ U_3 U_4 U_3^+ \xi = m_i h_i,$$

$$\mathrm{Tr} AB = m_i h_{P(i)},$$

we conclude that the second sum assumes its possible value, namely, $m_i h_i$ only if the sequences $\{m_i h_i\}$ and $\{m_i h_{P(i)}\}$ are permutation of each other. Hence the maximality, (minimality), of $\varphi(U_0, V_0)$ implies that $\{m_i h_i\}$, $(\{-m_i h_i\})$, is an enumeration of the eigenvalues of $U_0 v V_0^+ \xi$. It is possible to identify immediately U_0 and V_0 if in the polar representation[6] of v and ξ we diagonalize the occurring positive definite matrices so that

$$v = U_1 v_d V_1, \quad \xi = U_2 \xi_d V_2, \tag{75}$$

where v_d and ξ_d are diagonal matrices with ordered both sequences $\{h_i\}$ and $\{m_i\}$ as diagonals. By taking this decomposition in Eq. (69) we observe that

$$\varphi(U,V) = \mathrm{Re\,Tr}\; U U_1 v_d V_1 V^+ U_2 \xi_d V_2 \tag{76}$$

attains its maximum for

$$U = U_0 = V_2^+ U_1^+, \quad V = V_0 = U_2 V_1. \tag{77}$$

Therefore, a solution maximizing $\varphi(U,V)$ is

$$v_0 = U_0 v V_0^+ = V_2^+ U_1^+ v V_1^+ U_2^+. \tag{78}$$

It is easy to check that this v_0 is satisfying the equations

$$[C, \xi] = 0, \tag{79a}$$

$$\{D', \xi\} = 0, \tag{79b}$$

which are exactly the Eqs. (63) or (64) [(65)] as in this case $\{Q_a^{(N)'}\}$ span the whole space $\mathscr{M}^{N^2}$ hence $\mathscr{A}_1 = \mathscr{A}$, $\mathscr{A}_2 = 0$. As Q in $G_\xi = Q \square (Z_2(P) \times Z_2(C))$ is a subgroup of $\mathrm{U}(N)^d$, at least $\langle w, u_{N^2} w \rangle \equiv \xi_{N^2}$ in Eqs. (48)–(50) [with $u_{N^2} = F_R(\Lambda_{N^2} \sim 1)$ a $\mathrm{U}(N)^d$-singlet] being different from zero, we suppose that the eigenvalues $\{H_i\}$ of ξ satisfy the relations

$$H_i \neq 0, \quad H_i + H_j \neq 0, \tag{80}$$

the last equations excluding the accidental relations (symmetrices), for the parameters of ξ (ξ). Then Eq. (79b) implies $D' = 0$, hence v_0 is Hermitian. From Eq. (79a) we deduce

$$v_0 = C = c_{0j_\xi} \Lambda_{j_\xi}, \quad \mathscr{H}'_{B0} = c_{0j_\xi} u_{j_\xi}, \tag{81}$$

where $\{\Lambda_{j_\xi}\} \subset \{\Lambda_j^{(N)}\}_{j=1:N^2}$ are those (Hermitian) matrices (40) commuting with ξ and $\{u_{j_\xi}\}$ are the corresponding (even) operators. ($\{\Lambda_{j_\xi}\}$ inclose all the generators of Q in G_ξ —the stabilizer of vacuum.)

If $Q = \mathrm{U}(q) \subset \mathrm{U}(N)^d$, a useful basis in $\mathscr{M}^{N^2}$ [or in the $u(N)$ algebra] is the following sequence of the Hermitian matrices Λ_j (which may be identified with $Q_j^{(N)'}$), $j = 1:N^2$:

$$\Lambda_{N^2} \equiv \sqrt{2/N}\, 1,$$

$$\Lambda_{m^2-1} \equiv \sqrt{2/m(m-1)}\ \mathrm{diag}\,\{\overset{m-1}{1,1,\ldots,1} - (m-1), 0,0,\ldots,0\}, \quad m = 2:N, \tag{82b}$$

$$\Lambda_{m^2+2p} \equiv E_{p+1}^{m+1} + E_{m+1}^{p+1}, \quad \Lambda_{m^2+2p+1} \equiv -i\,(E_{p+1}^{m+1} - E_{m+1}^{p+1}), \tag{82c}$$

$p = 0:(m-1), \quad m = 1:N,$

where

$$(E_m^n)_{kl} = \delta_{mk}\delta_{nl}, \quad m \neq n, \quad m,n = 1:N. \tag{82d}$$

These matrices generalize the Gell–Mann matrices for the U(3) case for each unitary subgroup occuring in the canonical chain

$$\mathrm{U}(2) \subset \mathrm{U}(3) \cdots \subset \mathrm{U}(N-1) \subset \mathrm{U}(N)$$

and they have the following properties: (i) $\{\Lambda_j\}_{j=1:q^2-1}$ is an o.n.b. of the $\mathrm{su}(q) \subset \mathrm{su}(N)$ algebra; (ii) $\{\Lambda_{N^2}; \Lambda_{m^2-1};, m = q+1:N\}$ are singlets of $\mathrm{SU}(q) \subset \mathrm{SU}(N)^d$; (iii) the set (82b) is a basis of a Cartan subalgebra $\mathscr{C}$ of $\mathrm{su}(N)$.

If $Q = \mathrm{U}(q)$ and use the advantage of the freedom of a $\mathrm{U}(q)$ transform which does not change $\varphi(U_0, V_0)$ but it diagonalizes v_0, we get

$$v_0 = C = c_{0m^2-1}\Lambda_{m^2-1} + c_{0N^2}\Lambda_{N^2}, \quad m = 1:N,$$

$$\mathscr{H}'_{B0} = c_{0m^2-1} u_{m^2-1} + c_{0N^2} u_{N^2}.$$

The previous analysis of the $\mathrm{U}(N) \times \mathrm{U}(N)$ case is giving (i) the way of getting the transform (U_0, V_0) selecting the extremum solution v_0, $(\mathscr{H}'_{B0})$ on the $\mathrm{U}(N) \times \mathrm{U}(N)$ orbit of v, $(\mathscr{H}_B)$; (ii) the structure of the extremum solution, depending on properties (symmetry) of ξ; (iii) the extremum constraints eliminate the CP-invariance breaking terms even if such terms was contained in the initial $\mathscr{H}_B$.

5. THE $(N, N^*) \oplus (N^*, N)$ MODEL OF $\mathrm{SU}(N) \times \mathrm{SU}(N)$

Now let us look for the extremum of $\varphi(U,V)$ in Eq. (60) on the $\mathrm{SU}(N) \times \mathrm{SU}(N)$ orbit of v, $(\mathscr{H}_B)$, with $U, V \in \mathrm{SU}(N)$.

As the generators $\{Q_a^{(N)'}\}_{a=1:N^2-1}$ of $\mathrm{SU}(N)$ are traceless, it comes out that the subspace $\mathscr{A}_2$ in Eq. (64) is generated by $\Lambda_{N^2}^{(N)'} \equiv i\sqrt{2/N}\; 1$. Hence we have

$$[C, \xi] = 0, \tag{83a}$$

$$\{D', \xi\} = i\mu 1. \tag{83b}$$

With the assumptions in Eq. (80), from Eq. (83b) we deduce that now D' has the symmetry of ξ is $\mu \neq 0$ and $D' = 0$ if $\mu = 0$. Therefore, the CP-symmetry breaking part of $\mathscr{H}'_{B0}$ has the vacuum (w) symmetry (in $\epsilon = 0$ limit) or it is absent.

Indeed, let U be a (unitary) transform such that $\xi' = U\xi U^{-1}$ is diagonal and let us define $D'' = UD'U^{-1}$. The Eq. (83b) may be written as

$$D''_{ij}(H_i + H_j) = i\mu\delta_{ij} \tag{84}$$

and by Eqs. (80) we obtain

$$D''_{ij} = D_i\delta_{ij}, \quad D_i = \frac{i}{2}\mu H_i^{-1}, \quad D'' = \frac{i}{2}\mu\xi^{-1}. \tag{84'}$$

On the other hand, by Q-symmetry of ξ, U may be chosen such that ξ' has the form of diagonal blocks corresponding to the IR's of $Q \subset \mathrm{SU}(N)^d$ contained in the IR (N) of $\mathrm{SU}(N)^d$; the eigenvalues of ξ in each multiplet are equal. Hence D'' has the same spectral structure of diagonal blocks with equal eigenvalues inside a Q-multiplet. The transform reducing ξ is also reducing D', hence D' (or D) has the same Q-symmetry as ξ.

Equation (83a) is the same as Eq. (79a) hence C is given by Eq. (81) and the extremum solution is

$$v_0 = c_{0j_\xi}\Lambda_{j_\xi} + d_{0j_\lambda}\Lambda_{j_\lambda}, \quad \mathscr{H}'_{B0} = c_{0j_\xi}u_{j_\xi} + d_{0j_\lambda}v_{j_\lambda}, \tag{85}$$

where $\Lambda_{j_\lambda} = -i\Lambda'_{j_\lambda}$ are the Q invariant matrices which appear in ξ, v_{j_λ} being the (odd) operators associated with Λ'_{j_λ}.

If $Q = \mathrm{SU}(q) \subset \mathrm{SU}(N)^d$ and we take the advantage of the freedom of a SU(q)-transform, then in the basis (82) we get

$$\mathscr{H}'_{B0} = c_{0N^2}u_{N^2} + d_{0N^2}v_{N^2} + \sum_{m=q+1}^{N}(c_{0m^2-1}u_{m^2-1} + d_{0m^2-1}v_{m^2-1} + \sum_{m=2}^{q}c_{0m^2-1}u_{m^2-1}, \tag{86}$$

where u_{N^2} and v_{N^2} are the only $\mathrm{SU}(N)^d$ singlets in the (real) representation space $(N,N^*) + (N^*,N)$.

We now have to determine the parameters of $\mathscr{H}'_{B0}$, functions of the parameters in $\mathscr{H}_B$ and ξ. We start by the remark that D and ξ are commuting so that they may be simulataneously diagonalized. From Eq. (83a) we deduce that the same transform bringing ξ and D to the structure of diagonal blocks is bringing C to a structure of blocks corresponding to the IR's of Q contained in the IR (N) of $\mathrm{SU}(N)^d$. Therefore, we get

$$[C,\xi] = [D,\xi] = [C,D] = [v_0,\xi] = 0, \qquad v_0 = C + \frac{i}{2}\mu\,\xi^{-1}. \tag{87}$$

[This time the matrix C, organized in blocks, can be diagonalized by a Q-transform which leaves $\varphi(U_0,V_0)$ unchanged.]

In virtue of Eqs. (87) we deduce

$$\varphi(U_0,V_0) = \mathrm{Re\,Tr}\, v_0\,\xi = \mathrm{Tr}\, C\xi = C_iH_i, \tag{88}$$

where C_i and H_i are the (real) eigenvalues of C and ξ respectively. For given v and ξ, the $(N+1)$ parameters C_i and μ should be determined and ordered to make C_iH_i and an extremum of $\varphi(U,V)$.

We see that vv^+, v^+v, $v_0v_0^+$, and $v_0^+v_0$ (where $v_0 = U_0vV_0^+$) have the same eigenvalues[8] $M_i = m_i^2$ (with the same multiplicities), hence

$$M_i = C_i^2 + \tfrac{1}{4}\mu^2H_i^{-2}, \quad C_i = \pm(M_i - \tfrac{1}{4}\mu^2H_i^{-2})^{1/2}. \tag{89}$$

Also we have

$$\det v_0 = \det v. \tag{90}$$

Hence the parameters $\{C_i,\mu\}$ satisfy Eqs. (89), (90); $\mu \neq 0$ iff $M_i - \frac{1}{4}\mu^2H_i^{-2} > 0$ for all i. If we enumerate H_i and M_i with their moduli in nonincreasing order, the same is true for C_i in Eqs. (89).

There are many possible cases. The "absolute" extrema of C_i and H_i are obtained for

$$\mu = 0, \quad C_i^2 = M_i \tag{91a}$$

$$\mathrm{sgn}C_i = \mathrm{sgn}H_i \quad \text{when } C_iH_i\text{-maximum}; \tag{91b}$$

$$\mathrm{sgn}C_i = -\mathrm{sgn}H_i \quad \text{when } C_iH_i\text{-minimum}. \tag{91c}$$

These extrema are obtained if Im Det$v = 0$ and if there exist the matrices $\mathscr{U}$ (Selected as the simplest possible), U_0 and V_0 so that

$$v_0 = U_0vV_0^+ = \mathscr{U}v'\mathscr{U}^{-1}, \quad [\mathscr{U},\xi] = 0, \quad U_0,V_0 \in \mathrm{SU}(N), \tag{92}$$

where v' are diagonal matrices with the diagonal elements satisfying (91a) and (91b) or (91a), (91c).

If $\mu \neq 0$ and there exist $\mathscr{U}$, U_0, and V_0 to carry v in

$$v_0 = U_0vV_0^+ = \mathscr{U}(C' + (i/2)\mu\xi^{-1})\mathscr{U}^{-1}; \quad [\mathscr{U},\xi] = 0, U_0,V_0 \in \mathrm{SU}(N), \tag{93}$$

where C' are diagonal matrices with the diagonal elements satisfying (89) and (91a) or (89) and (91b), then $\varphi(U,V)$ in Eq. (69) attains its extrema with U_0, V_0, v_0 given in Eq. (93). Obviously, if such a v_0 is inaccessible by $\mathrm{SU}(N)\times\mathrm{SU}(N)$ transforms, some C_i in Eqs. (91b) or (91c) should be modified so that C_iH_i in Eq. (88) suffers minimal variations with respect to the previous extrema and the new extremum to become accessible by the $\mathrm{SU}(N)\times\mathrm{SU}(N)$ transforms.

6. CONCLUSIONS

In this section we enumerate some well-known results of theories based on unitary models. We have already given a natural and straightforward explanation of these models, taking as tool an extremum principle on groups.

1. First of all we remark that the Hamiltonians in the usual SU(3)×SU(3)[1,5,9,10] and SU(4)×SU(4)[11-13] models with a SU(q)-symmetric vacuum ($q = 2,3,4$) belong to the class of stationary solutions having the form given in Eq. (86) with $c_{m^2-1} = 0$ ($m = 3{:}q$) and all $d_j = 0$, the same remark holds for the U(3)×U(3)[4] and U(4)×U(4)[14] models and partially for the gauge SU(4)×SU(4)[15] and SU(16)×SU(16)[16] models.

Information concerning the parameters are extracted from the particle mass spectrum in the first order of perturbation theory and from decay amplitudes data.[17]

There are some effects that suggest taking into account the last terms in $\mathscr{H}'_{B0}$ in Eq. (86).

Among these we recall the "nonelectromagnetic" violation of the isospin by the c_3u_3 term in the chiral SU(3)×SU(3); an alternative motivation for this term may be a nonzero amplitude for the medium-strong $\eta\to3\pi$ decay.[3,18]

2. An application of the extremum condition would

lead, by the above method to the extrmum domain for c_i and ξ_i in the $N=3$ models. Now $\mathscr{H}_B$ is some "vector" in the $(3,3^*)\oplus(3^*,3)$ representation space[9] and the vacuum is $SU(3)^d$ or SU(2)-symmetric.

For a $\epsilon\mathscr{H}_B$ and a vacuum SU(2)-symmetric, with the usual notation we have[1,9]

$$\epsilon\mathscr{H}_B = c_0 u_0 + c_8 u_8, \quad (u_0 \equiv u_9),$$

and the only nonzero vacuum expectations values are ξ_0 and ξ_8.

Therefore, in the Gell-Mann basis [see Eqs. (82)] we have

$$v = c_0\lambda_0 + c_8\lambda_8 \equiv \tfrac{2}{3} c_0 \operatorname{diag}(M_1, M_2, M_3),$$

$$\xi = \xi_0\lambda_0 + \xi_8\lambda_8 \equiv \tfrac{2}{3}\xi_0 \operatorname{diag}(H_1, H_2, H_3),$$

where

$$M_1 = M_2 = 1 + a, \quad M_3 = 1 - 2a, \quad a \equiv c_8/\sqrt{2}c_0,$$

$$H_1 = H_2 + 1 + b, \quad H_3 = 1 - 2b, \quad b \equiv \xi_8/\sqrt{2}\xi_0.$$

If we define S as

$$S \equiv 2|M_1||H_1| + |M_3||H_3|$$

we observe that for the $U(3)\times U(3)$ case we obtain

(I) $\operatorname{Min}\varphi(U.V) = -\gamma S$ if $\gamma > 0$,

(II) $\operatorname{Min}\varphi(U,V) = \gamma S$ if $\gamma < 0$, $\gamma \equiv -\tfrac{2}{3} c_0\xi_0$,

with the following orderings of the eigenvalues $|M_i|$ and $|H_i|$ of matrices $|v|$ and $|\xi|$:

(i) $|M_1| \geqslant |M_3|$, $\quad |H_1| \geqslant |H_3|$,

(ii) $|M_1| \leqslant |M_3|$, $\quad |H_1| \leqslant |H_3|$.

Looking for (v, ξ) models with $\operatorname{Re}\operatorname{Tr} v\xi = \langle w, \mathscr{H}_B w\rangle$ having minimum value we remark that M_i and H_i should have the same algebraic signs in the first and opposite ones in the second case. These constraints define the extremum domains for the parameters γ, a, and b. For example, the orderings (i) with $H_i, M_i > 0$ and $M_1, H_1 > 0$, $M_3, H_3 < 0$ respectively, corresponding to some situations of the first case, determine the limitations $0 < a,b < \tfrac{1}{2}$ and $\tfrac{1}{2} < a,b < 2$, respectively.

Bu enumerating all possible constraints, the Okubo and Mathur[4] domains for a and b simply result from previous discussion.

On the other hand, given v and ξ with the orderings (i) or (ii), according to the results obtained in Sec. 4, the extremum solution v_0 on the $U(3)\times U(3)$ orbit of v has the form

$$v_0 = \varepsilon\sqrt{\tfrac{2}{3}}\, c_0 \operatorname{diag}(|M_1| \operatorname{sgn} H_1, |M_1| \operatorname{sgn} H_1, |M_3| \operatorname{sgn} H_3,$$

where $\varepsilon = +1$ if $\gamma < 0$ and $\varepsilon = -1$ if $\gamma > 0$. Therefore, $\epsilon\mathscr{H}'_{B0} = c_{00}u_0 + c_{08}u_8$ has the same SU(2)-symmetry as that of $\epsilon\mathscr{H}_B$.

However for other orderings of $|M_i|$ and $|H_i|$ the initial SU(2)-symmetry of the Hamiltonian is broken. For example if $|M_1| > |M_2|$ and $|H_1| < |H_2|$ the extremum solution in the case $\gamma > 0$ is

$$v_0 = \sqrt{\tfrac{2}{3}}\, c_0 \operatorname{diag}(|M_2| \operatorname{sgn} H_1, |M_1| \operatorname{sgn} H_1, |M_1| \operatorname{sgn} H_2$$

so that $\epsilon\mathscr{H}'_{B0} = c_{00}u_0 + c_{08}u_8 + c_{03}u_3$.

In all these cases the CP-symmetry violating terms in the extremum solutions are absent.

Let us consider now the extremum solutions on the $SU(3)\times SU(3)$ orbit of $\epsilon\mathscr{H}_B$.

In order to compare our results with those from the literature[1] we will use the parameters

$$\delta = (1+a)(1-2a)^{-1}, \quad \eta = (1+b)(1-2b)^{-1}$$

which allow us to put

$$v = 3\sqrt{\tfrac{2}{3}}\, c_0(1+2\delta)^{-1}\mathscr{M}, \quad \mathscr{M} \equiv \operatorname{diag}(\delta,\delta,1),$$

$$v = 3\sqrt{\tfrac{2}{3}}\, \xi_0(1+2\eta)^{-1}\mathscr{X}, \quad \mathscr{X} \equiv \operatorname{diag}(\eta,\eta,1).$$

Depending on the sign of $9\gamma(1+2\delta)^{-1}(1+2\eta)^{-1}$ the maximum or minimum of the function $\operatorname{Re}\operatorname{Tr} U\mathscr{M}V^{+}\mathscr{X} \sim \varphi(U,V)$ should be reached. The Eqs. (87), (89), and (90) straightforwardly give the extremum solution

$$\mathscr{M}_0 = \mathscr{C} + \frac{i}{2}\mu\mathscr{X}^{-1},$$

where

(a) $\mu = \pm 2\eta\delta(1 - \tfrac{1}{4}\eta^2\delta^2)^{1/2} \neq 0$ if $\eta^2\delta^2 < 4$,

$\mathscr{C} = \operatorname{diag}(c_1, c_2, c_3)$, $\quad c_1 = c_2 = -\tfrac{1}{2}\eta\delta^2$,

$c_3 = \tfrac{1}{2}\eta^2\delta^2 - 1$,

(b) $\mu = 0$ if $\eta^2\delta^2 \geqslant 4$, $\quad |c_1| = |c_2| = |\delta|$, $|c_3| = 1$.

In the last case we chose c_1, c_2, and c_3 so that $\det\mathscr{M}_0 = \det\mathscr{M} = \delta^2$. The discussion is similar to the previous one and we do not repeat it here.

According to Eq. (86) in the case (a), a CP-invariance violating part occurs in the Hamiltonian

$$\epsilon\mathscr{H}'_{B0} = c_{00}u_0 + c_{08}u_8 + d_{00}v_0 + d_{08}v_8.$$

the particular choice $\eta = 1$ [corresponding to a $SU(d)^d$-symmetric vacuum] gives the Dashen's[1] solutions with a $SU(3)^d$-singlet as CP-symmetry violating term ($d_{08} = 0$).

For the $SU(4)\times SU(4)$[11-13] models with an SU(4) or SU(3)-symmetric vacuum, the general method given above can be immediately adapted.

3. A general feature of all previous introduced models and an interesting consequence of the extremum constraint is that it may infer a CP-violating part which has the vacuum symmetry in the $\epsilon = 0$ limit. In fact this result is a pure mathematical consequence of an extremum principle in the presence of the outer automorphisms identified with C and P transformations. In $SU(3)\times SU(3)$ and $SU(4)\times SU(4)$ models this part does not contribute to the pseudoscalar meson masses but has a nonzero contribution to some decay amplitudes.[3,13]

However, in the usual models such a CP symmetry violating part implies some difficulties. It is well known that the very small upper limit on the electrical dipole moment of the neutron represents a very serious difficulty with the order of of magnitude for any theory in which the CP-violating part conserves hypercharge.[3,18]

4. It is easy to see that if there is a spontaneous breaking of CP-symmetry, that is, there exists an operator $W \in SU(N)\times SU(N)$ (which in general does not leave the vacuum unchanged) such that $\mathscr{H}'_{B0}$ is invariant under the extended CP-operation defined by $W(\mathrm{CP})W^{-1}$, then $\mathscr{H}'_{B0}$ belongs to the orbit of an $\mathscr{H}_B$ invariant under the usual CP-transformations defined in Eq. (45) and W carries $\mathscr{H}_B$ in $\mathscr{H}'_{B0}$. (Obviously the CP-symmetry violating part has the vacuum symmetry in the $\epsilon = 0$ limit.)

In this case even if all d_j in $\mathscr{H}_B$ are zero we obtain d_{0j} not equal to zero in $\mathscr{H}'_{B0}$. However the parameters in $\mathscr{H}'_{B0}$ are not independent; they satisfy the relation

$$\mathrm{Im}\,\det v_0 = \mathrm{Im}\,\det v = 0.$$

This implies that the order of magnitude of the CP-violating effects are directly related with that of the parameters c_{0j} in $\mathscr{H}'_{B0}$.

For the $N = q = 3$ model, the last equation implies the relation $d^2_{00} = \frac{3}{2}(2c^2_{00} - c^2_{08} - c^2_{03})$ and one can conclude that, to the extent that perturbation theory about chiral limit is adequate for order of magnitude estimations, the mechanism of spontaneous CP-violation in the $SU(3)\times SU(3)$ models give absurdly large effects.[3]

Similar relations and conclusions immediately follow for the $N = 4$, $q = 4,3$ models.

5. Even though the problem studied here is not formulated as an extremum problem for some $SU(N)\times SU(N)$ invariant functions, the approach of Michel and Radicati[5] occurs in the $\epsilon = 0$ limit.

The arguments which extremizes such invariant functions must satisfy some simple algebraic equations in the symmetric algebras and belong to the critical orbits. The algebraic technique of Michel and Radicati had been used for determining the Hamiltonians $\mathscr{H}_B$ associated with the "directions of breaking" of the symmetries $SU(3)\times SU(3)$[5,10] and $SU(4)\times SU(4)$.[12,13] A further analysis of the same problem from the point of view of orbit structure will be published in a forthcoming paper.

ACKNOWLEDGMENTS

One of the authors (O.G) would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

[1] R. Dashen, Phys. Rev. D **3**, 1979 (1971).
[2] J. Nuyts, Phys. Rev. Lett. **26**, 1604 (1971); **27**, 361 (E) (1971).
[3] M.A.B. Bég, *Phys. Rev. D* **4**, 3810 (1971); Schechter and Y. Ueda, Phys. Rev. D **5**, 2821 (1972).
[4] S. Okubo and V.S. Mathur, Phys. Rev. D **1**, 2046 (1970); V.S. Mathur and S. Okubo, Phys. Rev. D **1**, 3468 (1970).
[5] L. Michel and L. Radicati, Ann. Phys. (N.Y.) **66**, 758 (1971); L. Michel, Lecture Notes in Physics, No. 6, 1969, Batelle Recontres, Seattle (Springer-Verlag, Berlin, 1970).
[6] M. Hamermesh, *Group Theory* (Addison-Wesley, Reading, Mass., 1964).
[7] E.P. Wigner, *Group Theory* (Academic, New York, 1959).
[8] E. Hewitt and K.A. Ross, *Abstract Harmonic Analysis* (Springer-Verlag, New York, 1970), Vols. I, II.
[9] M. Gell-Mann, R. Oakes, and B. Renner, Phys. Rev. **175**, 2195 (1969); S L. Glashow and S. Weinberg, Phys. Rev. Lett. **20**, 224 (1968).
[10] A. Chakrabarti and C. Darzens, Ann. Phys. (N.Y.) **69**, 193 (1972); C. Darzens, Ann. Phys. (N.Y.) **76**, 236 (1973).
[11] D. Amati, H. Bacry, J. Nuyts, and J. Prentki, Nuovo Cimento **34**, 1732 (1964); B.J. Bjorken and S.L. Glashow, Phys. Lett. **11**, 255 (1964); S.L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D **2**, 1285 (1970); Z. Maki, T. Maskawa, and I. Umemura, Prog. Theor. Phys. (Kyoto) **47**, 1682 (1971); H. Georgi and S.L. Glashow, Phys. Rev. D **6**, 2977; J.D. Bjorken and C.H. Llewellyn-Smith, Phys. Rev. D **7**, 887 (1973); P. Dittner and S. Eliezer, Phys. Rev. D **8**, 1929 (1973); P. Dittner, S. Eliezer, and T.K. Kuo, Phys. Lett. **30**, 1274 (1973); D.A. Dicus and V.S. Mathur, Phys. Rev. D **9**, 1003 (1974); S.C. Prasad, Phys. Rev. D **9**, 1017 (1974).
[12] R.F. Mott, Nucl. Phys. B **76**, 301 (1974).
[13] R.J. Oakes and P. Sorba, FERMILAB THY Pub. 77/78 (1977).
[14] M. Singer, Phys. Rev. D **16**, 2304 (1977); Nuovo Cimento A **42**, 25 (1977).
[15] J. Pati and A. Salam, Phys. Rev. D **8**, 1240 (1973).
[16] H. Fritzsch and P. Minkowski, Ann. Phys. **93**, 193 (1975).
[17] M.K. Gaillard, J. Rosner, and B.W. Lee, Rev. Mod. Phys. **47**, 2777 (1975).
[18] M.A.B. Bég, Comm. Nucl. Part. Phys. **4**, 182 (1970).

Differentiable manifolds and the principle of virtual work in continuum mechanics

Marcelo Epstein and Reuven Segev
Department of Mechanical Engineering, The University of Calgary, Calgary, Alberta, Canada

A material body is conceived in terms of a global and a local model, the compatibility of which is shown to imply a generalization of the principle of virtual work. The classical notion of stress appears as a particular case of the force associated with the local model when an affine connection is specified in the physical manifold.

1. INTRODUCTION

The last two decades have witnessed a revival of classical particle mechanics[1,2] which can be characterized by the emphasis placed on the geometrical structure underlying the description of physical phenomena. The purpose of this paper is to present the foundations for a similar approach, in terms of differentiable manifolds, in continuum mechanics. Since our intention is to explore the geometrical structure of the basic variables, in particular the idea of stress, we limit ourselves to statics, leaving a more embracing treatment for future work.

Unlike the case of particle mechanics, where the idea of a configuration manifold is a direct reflection of the intuitive notion of the degrees of freedom of the system, the situation in continuum mechanics is more involved in two senses. Firstly, the configuration manifold of the body conceived as a collection of material particles, is infinite dimensional. Secondly, and more important, in order to introduce the structural properties of the medium, it becomes necessary to view the body not merely as a collection of points, but also as a collection of neighborhoods. Thus, the material body is actually described not just by one, but by two models which we call the global and the local models, respectively. Mathematically speaking the local model is based on the tangent bundle of the original body, conceived in the global model as a collection of particles. Once the models are defined, each with its own infinite dimensional configuration space, the concepts of virtual displacements and forces for each model are introduced in a natural way as elements of the tangent and cotangent bundles of their respective configuration spaces.

A key notion in this approach is that of the compatibility between both models. As it is shown in Sec. 4, such compatibility entails, in addition to the geometric compatibility, the statical compatibility of forces, which turns out to be the analog of the principle of virtual work in the traditional formulation.

The concept of force associated with the local model, as defined in Sec. 3, results in a generalization of the idea of stress. To explore under which circumstances such a force is indeed derivable from a stress in the traditional sense, Sec. 5 presents the important particular case in which the physical space is endowed with an affine connection (whose meaning should be clarified for any particular problem being treated). In such a case, the induced decomposition of the double tangent bundle into horizontal and vertical components, allows the expression of the local force in terms of linear automorphisms of the tangent spaces. This results in the explicit expression of the principle of virtual work in terms of an equilibrium equation as in the conventional treatment.

2. THE GLOBAL MODEL

In this work we assume that S, the physical space in which physical events take place, is a fixed m-dimensional differentiable manifold.

A body B can be regarded as an n-dimensional differentiable manifold whose points are referred to as "material points". We will restrict our attention to bodies satisfying the additional conditions:

(B1) B is orientable;

(B2) B can be covered by an atlas consisting of only one chart;

(B3) $n \leqslant m$.

Bodies manifest themselves through their configurations in the physical space. A configuration κ is an immersion

$$\kappa : B \to S. \tag{1}$$

Obviously $\kappa(B)$, the image of B in S, satisfies B1, B2, and B3.

For a given body B, the set

$$Q = \{\kappa : B \to S\}, \tag{2}$$

of all its configurations, when appropriately given the structure of an infinite dimensional manifold, is called the configuration space.

Consider now the "tangent bundle" TQ of the configuration space. A typical element $\delta\kappa \in TQ$ is a map[3,4]

$$\delta\kappa : B \to TS, \tag{3}$$

defining a configuration κ by

$$\tau_S \circ \delta\kappa = \kappa, \tag{4}$$

where τ_S denotes the natural tangent bundle projection. [The composition $\delta\kappa \circ \kappa^{-1}$ defines a vector field on $\kappa(B)$.] Traditionally, $\delta\kappa$ is referred to as "virtual displacement".

Consider an element f in the "cotangent bundle" T^*Q of the configuration space. The evaluation of f at a virtual displacement $\delta\kappa$ is known as the "virtual work of the force f".

The foregoing formulation, which will be referred to as the Global Model, is analogous to its counterpart in the mechanics of particles, with the difference that in our case the configuration space is infinite dimensional.

 0022-2488/80/051243-03$01.00

3. THE LOCAL MODEL

The structural properties of the body are usually represented by a relation which expresses the behavior of a material point as determined by its neighboring points. In order that such a relation can be given, a different description of the configuration and the virtual displacements of the body is necessary. In a formulation of this kind the body will be regarded as a collection of neighborhoods, and notions such as configuration, virtual displacements and forces will accordingly apply to this collection. A neighborhood of a material point will be modelled mathematically by the tangent space at this point. This model of a body as a tangent bundle will be referred to as the local model.

A local configuration is an immersion

$$\chi: TB \to TS, \text{ such that } \tau_S \circ \chi = \kappa \circ \tau_B \text{ for some } \kappa, \tag{5}$$

and the local configuration manifold, R, is defined similarly to the global configuration manifold as

$$R = \{\chi: TB \to TS\}. \tag{6}$$

A local virtual displacement will be an element $\delta\chi \in TR$ and, as in the previous model, it can be identified with the map

$$\delta\chi: TB \to TTS. \tag{7}$$

Similarly, a local force is an element $\sigma \in T^*R$ and the virtual work is the evaluation of σ at $\delta\chi$. Note that the local force σ is a generalization of the classical notion of stress.

4. COMPATIBILITY OF THE MODELS

In this section we formulate the conditions that the local configuration, local virtual displacements, and local force have to satisfy in order that they be compatible with their counterparts in the global model.

We say that a local configuration χ is compatible with a given global configuration κ if

$$\chi = T\kappa, \tag{8}$$

where $T\kappa$ denotes the tangent map to κ.

In order to define compatibility of virtual displacements the following should be noted:

(i) If the local configuration χ is compatible with the global configuration κ, we have

$$\tau_{TS}\delta\chi = T\tau_S(T\delta\kappa), \quad T\tau_S(\delta\chi) = \tau_{TS}(T\delta\kappa), \tag{9}$$

where

$$\tau_{TS}: TTS \to TS, \tag{10}$$

is the natural tangent bundle projection of TTS,

$$T\delta\kappa: TB \to TTS, \tag{11}$$

is the tangent map to τ_S and

$$T\tau_S: TTS \to TS, \tag{12}$$

is the tangent map to the natural projection τ_S;

(ii) If $(\mathbf{x}, \mathbf{a}, \mathbf{b}, \mathbf{c})$ is the representation of $u \in TTS$ under the chart α, we can define the canonical involution on TTS

$$\omega: TTS \to TTS, \tag{13}$$

by defining its local representative ω_α with respect to the manifold chart α as

$$\omega_\alpha(\mathbf{x},\mathbf{a},\mathbf{b},\mathbf{c}) = (\mathbf{x},\mathbf{b},\mathbf{a},\mathbf{c}). \tag{14}$$

It can be shown,[5] that the canonical involution is independent of the manifold chart.

With these preliminaries in mind, we say that the local virtual displacement $\delta\chi$ is compatible with a given global virtual displacement $\delta\kappa$ if

$$\delta\chi = \omega \circ T\delta\kappa. \tag{15}$$

$$\begin{tikzcd}
B \arrow[r, "\kappa"] \arrow[dr, "\delta\kappa"] & S \\
TB \arrow[u, "\tau_B"] \arrow[r, "\chi"] \arrow[dr, "\delta\chi"'] & TS \arrow[u, "\tau_S"'] \\
& TTS \arrow[u, "\tau_{TS}"']
\end{tikzcd}$$

Now that the tangent vectors to the respective configuration manifolds are comparable, the compatibility of the covectors can be established using the result of their evaluation on vectors. Thus, we say that a local force σ is compatible with the global force f if

$$\sigma(\delta\chi) = f(\delta\kappa), \tag{16}$$

for every compatible pair $\delta\chi$, $\delta\kappa$.

This condition is a generalization of what is classically referred to as the principle of virtual work.

5. EXAMPLE

In the particular case when an affine connection ∇ is specified in the physical manifold S, it will be shown that the local force σ can be represented by the classical local stress.[6] We will also give the conditions under which the local stress is compatible with a given global force of a special traditional type. In the sequel, we assume $m = n$.

A vector $u \in T_v TS$ is called vertical if it is tangent to the submanifold $T_{\tau_S v} S$. It can be proven that u is vertical iff

$$T\tau_S u = 0, \tag{17}$$

and we can define the vertical subspace V_v of $T_v TS$ as

$$V_v = \{u \in T_v TS;\ T\tau_S u = 0\}. \tag{18}$$

Given a connection on S, a unique horizontal distribution h can be determined. For $v \in TS$, $h(v) \in H_v$ is a unique vector where $H_v \subset T_v TS$ is the horizontal subspace of $T_v TS$. Thus, once we have a horizontal distribution, the tangent space $T_v TS$ is the direct sum of H_v and V_v and each vector can be decomposed uniquely into horizontal and vertical components. A map

$$\nu: T_v TS \to T(T_{\tau_S v} S), \tag{19}$$

can be defined that will give a unique vertical vector in $T_v TS$, and thus, a vector tangent to $T_{\tau_S v} S$, by

$$\nu(u) = u - h(T\tau_S u). \tag{20}$$

As $\nu(u)$ is a tangent vector to the vector space $T_{\tau_S v} S$ we can identify it with a vector

$$i \circ \nu(u) \in T_{\tau_S v} S, \tag{21}$$

where i is the canonical isomorphism between a vector space and its tangent space at a point.

With these preliminary ideas we can define the map

$$\Delta\chi: TB \to TS, \tag{22}$$

by

$$\Delta\chi = i \circ \nu \circ \delta\chi. \tag{23}$$

The following results should be noted:

(i) One can easily prove that for compatible virtual displacements

$$\Delta\chi(A) = i \circ v \circ \delta\chi(A) = \nabla_{T\kappa(A)}\delta\kappa, \text{ for all } A \in TB; \quad (24)$$

(ii) The affine connection $\nabla_X Y$ in a manifold is linear in X by definition; thus, we can define a linear differential operator

$$\mathbf{d}Y: TM \to TM, \quad (25)$$

by

$$\mathbf{d}Y(X) = \nabla_X Y; \quad (26)$$

(iii) Using (i) and (ii) we can write

$$\Delta\chi \equiv i \circ v \circ \delta\chi = \mathbf{d}\delta\kappa \circ T\kappa. \quad (27)$$

Once $\Delta\chi$ is defined so as to represent the local virtual displacement, we can use the fact that

$$\Delta\chi_x : T_x B \to T_{\kappa(x)} S, \quad (28)$$

$\Delta\chi_x$ being the restriction of $\Delta\chi$ to $T_x B$, is a linear transformation, in order to define the appropriate covector. A natural choice is the restriction $\bar{\sigma}_x$ to $T_{\kappa(x)}S$ of a linear map

$$\bar{\sigma}: \chi(TB) \to TB, \quad (29)$$

satisfying

$$\tau_B \circ \bar{\sigma}(a) = \kappa^{-1} \circ \tau_S(a), \text{ for all } a \in \chi(TB), \quad (30)$$

which is known as the local stress (1st Piola–Kirchhoff stress). We can now define the local force by

$$\sigma(\delta\chi) = \int_{\kappa(B)} \langle \bar{\sigma}_x, \Delta\chi_x \rangle \theta_S, \quad (31)$$

where θ_S is a volume element in S. Using Eq. (29) we have

$$\sigma(\delta\chi) = \int_{\kappa(B)} \langle \bar{\sigma}_x, (\mathbf{d}\delta\kappa \circ T\kappa)_x \rangle \theta_S = \int_{\kappa(B)} \langle T\kappa \circ \bar{\sigma}_x, \mathbf{d}\delta\kappa_x \rangle \theta_S, \quad (32)$$

The composition $T\kappa \circ \bar{\sigma}$ will be denoted by s and will be referred to as the Cauchy stress.

Let $(T^{k^*}S, S, \pi_S^k)$ denote the bundle of k-forms on S. We can associate a map

$$\hat{s}: TS \to T^{(m-1)^*}S, \quad (33)$$

satisfying

$$\pi_S^{m-1} \circ \hat{s}(a) = \tau_S(a), \quad \text{for all } a \in TS, \quad (34)$$

with each s, by the relation

$$\hat{s} \circ \delta\kappa = (s \circ \delta\kappa) \lrcorner\, \theta_S, \quad (35)$$

Define now the divergence operator

$$\operatorname{div}\hat{s}: TS \to T^{m^*}S, \quad (36)$$

by

$$\operatorname{div}\hat{s} \circ \delta\kappa \equiv d \circ \hat{s} \circ \delta\kappa - \langle s, \mathbf{d}\delta\kappa \rangle \theta_S, \quad (37)$$

where d denotes exterior differentiation. This definition is valid as all terms are proportional to the volume element in S.

Using the above definition, Eq. (35) and (37), we have

$$\sigma(\delta\chi) = \int_{\kappa(B)} \langle s_x, \mathbf{d}\delta\kappa_x \rangle \theta_S = \int_{\kappa(B)} (d \circ \hat{s} \circ \delta\kappa - \operatorname{div}\hat{s} \circ \delta\kappa). \quad (38)$$

With the compatibility condition for forces, Eq. (16), and Stokes' theorem we obtain

$$f(\delta\kappa) = \int_{\partial\kappa(B)} \hat{s} \circ \delta\kappa - \int_{\kappa(B)} \operatorname{div}\hat{s} \circ \delta\kappa. \quad (39)$$

We assume now, as is usually done, that the global force functional f can be represented in terms of two given mappings

$$\hat{b}: \chi(TB) \to T^{m^*}S, \quad (40)$$

$$\hat{t}: \chi(T\partial B) \to T^{(m-1)^*}S, \quad (41)$$

satisfying

$$\pi_S^m \circ \hat{b} = \tau_S, \quad (42)$$

$$\pi_S^{m-1} \circ \hat{t} = \tau_S, \quad (43)$$

as

$$f(\delta\kappa) = \int_{\kappa(B)} \hat{b} \circ \delta\kappa + \int_{\partial\kappa(B)} \hat{t} \circ \delta\kappa. \quad (44)$$

As Eqs. (41) and (46) hold for every virtual displacement we can write

$$\operatorname{div}\hat{s} + \hat{b} = 0, \quad \text{on } \kappa(B), \quad (45)$$

and

$$\hat{t} = \hat{s}, \quad \text{on } \partial\kappa(B). \quad (46)$$

The last two equations are customarily known as the equations of equilibrium.

ACKNOWLEDGMENT

This work has been supported by the Natural Sciences and Engineering Research Council of Canada, through Operating Grant NSERC-A4662, and by the Killam Foundation, through a predoctoral scholarship to the second author.

[1] R. Abraham and J.E. Marsden *Foundations of Mechanics*, 2nd Ed. (Benjamin/Cummings, Massachusetts, 1978).

[2] V.I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York, 1978).

[3] S. Sternberg, *Lectures on Differential Geometry* (Prentice-Hall, Englewood Cliffs, New Jersey, 1964).

[4] R. Hermann *Differential Geometry and the Calculus of Variations*, (Academic, New York, 1968).

[5] R. Abraham and J. Robbin, *Transversal Mappings and Flows* (Benjamin, New York, 1967).

[6] C. Truesdell and R. Toupin, "The Classical Field Theories", in *Handbuch der Physik*, Vol. III/1, edited by S. Flugge (Springer-Verlag, Berlin, 1960).

The growth and decay of weak waves in relativistic flows of dissociating gases

Rishi Ram, H. N. Singh, and Madhukar Gaur
Department of Applied Sciences, Institute of Technology, Banaras Hindu University, Varanasi-221005 India

(Received 15 Dcember 1978; accepted for publication 1 June 1979)

The propagation of a weak wave in a relativistic flow of a dissociating gas has been studied. The velocity of propagation of a relativistic weak wave has been determined. The fundamental growth equation governing the growth and decay of the wave has been obtained and solved. The relativistic results have been shown in full agreement with earlier results of classical gas dynamics. The problem of breakdown of weak discontinuities has also been solved. The critical time t_c is determined when the breakdown of the wave and the consequent formation of a shock wave occur due to nonlinear steepening. It is concluded that there exists a critical amplitude of the wave such that all compressive waves with an initial amplitude greater than the critical one will break down after a finite time t_c and a shock-type discontinuity will be formed, while an initial amplitude less than the critical one will result in a decay of the wave. On the other hand, an expansion wave will always decay and will ultimately be damped out. The global behavior of the wave amplitude has also been studied. It is concluded that the dissociative character of the gas is to increase the critical time. The relativistic and dissociative effects on the global behavior of weak discontinuities have also been discussed.

I. INTRODUCTION

A great deal of attention has been focused towards the relativistic study of weak discontinuities. Eckart[1] and Taub[2] presented the theoretical foundations of relativistic shock waves. The relativistic theory of propagation of weak waves in a perfect gas has been treated by Saini[3], Coburn[4], Kanwal[5] and McCarthy[6]. Grot and Eringen[7] formulated a general theory of relativistic continuum mechanics. The main aim of this paper is to develop a relativistic theory for the growth and decay of weak waves propagating in a simple dissociating gas of Lighthill's model.[8] In the temperature range where dissociation is important; the contribution of energy from electronic excitation and ionization are both assumed negligible. In a dissociating diatomic gas the state of the reacting mixture is uniquely described by three independent parameters such as the pressure p, the temperature T and the degree of dissociation α. We assume that the molecular effects leading to viscosity, diffusion and heat conduction are negligible. A simple dissociating gas is defined as a mixture resulting from a dissociation reaction in a symmetrical diatomic gas A_2, each A_2 molecule being made up from $2A_1$ atoms. The reaction is

$$A_2 + X \underset{k_r}{\overset{k_f}{\rightleftarrows}} 2A_1 + X,$$

where the species X can be either A_2 or A_1 and k_f and k_r are the reaction rate constants for the forward and reverse reactions.

II. BASIC PRELIMINARIES

The notation used in this paper is, with a few minor exceptions, identical with that employed by Grot and Eringen.[7]

Let X^k be the rectangular coordinates of a material point in a three dimensional space. The motion of a material body can be described by a new set of coordinates x^k given by

$$x^k = x^k(X^i, x^4), \quad x^4 = ct,$$

where t is the time and c is the constant velocity of light in vacuum. Let us introduce the concept of an Einstein–Riemann space V_4 by four coordinates $x^\alpha = (x^k, x^4)$ with a metric

$$ds^2 = \Gamma_{\alpha\beta} dx^\alpha dx^\beta.$$

The metric $\Gamma_{\alpha\beta}$ has constant components given by

$$\Gamma^{\alpha\beta} = \Gamma_{\alpha\beta}, \quad \Gamma^{ij} = \Gamma_{ij} = \delta_{ij}, \quad \Gamma^{44} = \Gamma_{44} = -1.$$

The world velocity can be expressed as

$$U^\alpha(x^\beta) = \beta(v^k/c, 1), \quad U^\alpha U_\alpha = -1, \tag{2.1}$$

where

$$v^k = c\frac{\partial x^k}{\partial x^4}, \quad \beta = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}. \tag{2.2}$$

In general, the range of a Latin index is 1, 2, 3 and that of a Greek index is 1, 2, 3, 4 unless stated otherwise. A dummy index will usually imply summation.

The invariant derivative of any function $\phi(x^\alpha)$ can be expressed in the form

$$D\phi = \left(\frac{\beta}{c}\right)\left(\frac{\partial\phi}{\partial t} + v^i\phi_{,i}\right) = U^\alpha\phi_{,\alpha}. \tag{2.3}$$

The equations governing the flow of an ideal dissociating gas[8] are

$$(\rho U^\alpha)_{,\alpha} = 0 \quad \text{(equation of continuity)}, \tag{2.4}$$

$$T^{\alpha\beta}_{,\beta} = 0 \quad \text{(equation of energy–momentum balance)}, \tag{2.5}$$

$$D\alpha = (4\beta/c)\{\rho D_0^2 k_r(1+\alpha)/R^2 T_d^2\} \times \{\rho_d(1-\alpha)\exp(-T_d/T) - \rho\alpha^2\}$$

(rate equation for dissociation), (2.6)

 0022-2488/80/051246-05$01.00

where

$$T^{\alpha\beta} = \omega U^\alpha U^\beta + pS^{\alpha\beta},$$
$$S^{\alpha\beta} = U^\alpha U^\beta + \delta^{\alpha\beta},$$
$$\omega = \rho c^2\{1 + p/(\gamma_e - 1)\rho c^2 + \alpha D_0/c^2\},$$
$$h = (4+\alpha)RT + \alpha D_0, \tag{2.7}$$
$$p = (1+\alpha)RT. \tag{2.8}$$

Here $T^{\alpha\beta}, \rho, D_0, \gamma_e, R, \rho_d$ and T_d respectively represent the energy–momentum tensor, the proper material density per unit volume in the instantaneous rest frame, the dissociation energy per unit mass, the effective heat exponent for a dissociating gas, the gas constant for the mixture, the characteristic density and temperature for dissociation. Although ρ_d is a function of T, it has been seen that the variation of ρ_d over the temperature range from 1000 °K to 7000 °K is very slight. Hence for all practical purposes ρ_d can be regarded as a constant, as it leads to negligible errors. The specific internal energy e consists of translational, rotational and vibrational energies, i.e.,

$$e = e_{tr} + e_{rot} + e_{vib} = (3/2)RT + RT + RT.$$

The ideal dissociating gas is defined by Lighthill[8] as having the vibrational mode of its diatomic components always "half excited," whereas the translational and rotational modes are fully excited." It follows that

$$e = 3RT, \quad c_v = 3R, \quad \gamma = (c_v + R)/c_v = 4/3.$$

Thus for an ideal dissociating model of a diatomic gas, the specific internal energy per unit mass can be written in the form[9]

$$e = p/(\gamma_e - 1)\rho, \quad \gamma_e = (4+\alpha)/3,$$

when $\alpha = 0$, all the results reduce to those for a perfect gas with $\gamma = 4/3$. In this case the rest specific internal energy is

$$\omega = \rho c^2 + 3p,$$

where $3p$ is the specific internal energy per unit volume for a perfect fluid with $\gamma = \gamma_e = 4/3$. This case corresponds to a photon gas at very high temperature in cosmological models.[10] For a nondissociating gas $\alpha = 0$ and $D_0 = 0$ and hence the rate Eq. (2.6) is satisfied identically. The two special cases for nondissociating and nonrelativistic limits will be discussed in Sec. VII of this paper.

From (2.4) and (2.5) we get

$$\rho\sigma c^2 DU^\alpha + S^{\alpha\beta}p_{,\beta} = 0, \tag{2.9}$$
$$(\rho\sigma c^2 U^\beta)_{,\beta} - U^\beta p_{,\beta} = 0, \tag{2.10}$$

where

$$\sigma = 1 + h/c^2. \tag{2.11}$$

Equations (2.9) and (2.10), respectively, represent the conservation of momentum and energy in relativistic flows with dissociation. In view of (2.3) and (2.4), Eq. (2.10) can be expressed in the form

$$\rho Dh - Dp = 0. \tag{2.12}$$

Applying (2.7) and (2.8) in (2.12) we get

$$Dp - \gamma_e(p/\rho)D\rho + F(p,\rho,\alpha) = 0, \tag{2.13}$$

where

$$F(p,\rho,\alpha) = (\beta/c)\{4\rho D_0^2 k_r/3R^2T_d^2\}\{\rho D_0(1+\alpha)^2 - 3p\} \times\{\rho_d(1-\alpha)\exp(-T_d/T) - \rho\alpha^2\}.$$

III. COMPATIBILITY CONDITIONS ON A TIMELIKE HYPERSURFACE

Let $\Sigma(x^\mu)$ be a regular surface in the V_4 space with parametric equations

$$x^\mu = \psi^\mu(b^1, b^2, b^3), \tag{3.1}$$

where b^1, b^2 and b^3 are parametric coordinates of the surface. The vectors $x^\alpha_{;\tau}$, where semicolon denotes covariant differentiation with respect to b^τ, are tangential to $\Sigma(x^\mu)$. The surface $\Sigma(x^\mu)$ is called a timelike hypersurface, if N_α is a spacelike vector, i.e.,

$$N_\alpha N^\alpha = 1,$$

where N_α are the components of the unit normal vector to the surface. The timelike hypersurface $\Sigma(x^\mu)$ may be regarded as a surface $S(t)$ in space–time for which the parametric equations are

$$x^4 = ct, \quad x^i = x^i(b^1, b^2, x^4). \tag{3.2}$$

If n_i are the components of the unit space normal to $S(t)$ and G is its speed of propagation, then we can write

$$N^\alpha = \bar\beta\{n^i, G/c\}, \quad N_\alpha = \bar\beta\{n_i, -G/c\}, \tag{3.3}$$

where

$$\bar\beta = (1 - G^2/c^2)^{-1/2}.$$

Let H be the region of the Einstein–Riemann space V_4 which is divided by the timelike hypersurface $\Sigma(x^\mu)$ into two regions H_1 and H_2. Let any flow parameter Z with its first and second derivatives be continuous in $H_1 + \Sigma$ and $H_2 + \Sigma$, but suffers a discontinuity in its first and second derivatives across $\Sigma(x^\mu)$. Such a discontinuity is called a "weak discontinuity" or "weak wave." If $[Z]$ denotes the jump in Z across $\Sigma(x^\mu)$, the geometrical compatibility conditions to be satisfied across $\Sigma(x^\mu)$ are[11]

$$[Z_{,\alpha}] = CN_\alpha, \tag{3.4}$$
$$[Z_{,\alpha\beta}] = \bar C N_\alpha N_\beta + 2N_{(\alpha}x^\tau_{\beta)}C_{;\tau} - Cb_{\tau\phi}x^\tau_{(\alpha}x^\phi_{\beta)}, \tag{3.5}$$

where

$$C = [Z_{,\alpha}]N^\alpha, \bar C = [Z_{,\alpha\beta}]N^\alpha N^\beta,$$
$$b_{\tau\phi} = N_\alpha x^\alpha_{;\tau\phi}, \quad x^\tau_\beta = \Gamma_{\alpha\beta}a^{\tau\phi}x^\alpha_{;\phi},$$
$$M_{(\alpha\beta)} = \tfrac{1}{2}(M_{\alpha\beta} + M_{\beta\alpha}), \quad a_{\alpha\beta} = \Gamma_{\tau\phi}x^\tau_{;\alpha}x^\phi_{;\beta}.$$

IV. VELOCITY OF PROPAGATION OF WEAK WAVES

Using (3.4) and the identity

$$a^{\tau\phi}x^\alpha_\tau x^\beta_\phi = \Gamma^{\alpha\beta} - N^\alpha N^\beta,$$

we get

$$[DZ] = V[Z_{,\beta}N^\beta] + \delta[Z], \tag{4.1}$$

where

$$\delta[Z] = U^\mu[Z_{,\mu}] - VN^\mu[Z_{,\mu}]. \quad (4.2)$$

Taking jumps in (2.4), (2.6), (2.9), and (2.13) and making use of (4.1) we get

$$V\nu + \rho\lambda^\alpha N_\alpha = 0, \quad (4.3)$$

$$V\epsilon = 0, \quad (4.4)$$

$$\rho\sigma V\lambda^\alpha + \frac{1}{c^2}\mu N_\beta S^{\alpha\beta} = 0, \quad (4.5)$$

$$V\mu - \gamma_e \frac{p}{\rho} V\nu = 0, \quad (4.6)$$

where

$$\lambda^\alpha = [U^\alpha_{,\beta}]N^\beta, \quad \nu = [\rho_{,\beta}]N^\beta, \quad \mu = [p_{,\beta}]N^\beta,$$
$$V = U^\alpha N_\alpha, \quad \epsilon = [\alpha_{,\beta}]N^\beta. \quad (4.7)$$

From the above set of equations, we can deduce

$$\mu = a_e^2\nu = -a_e^2\rho\lambda/V, \quad \lambda = \lambda^\alpha N_\alpha, \quad (4.8)$$

where

$$a_e^2 = \gamma_e P/\rho$$

is the effective velocity of sound. Multiplying (4.5) by N_α and using (4.3), we get

$$V^2 = a_e^2/c^2(\sigma - a_e^2/c^2). \quad (4.9)$$

Using (2.1) and (3.3) we get

$$V = -\beta\bar{\beta}G_0/c, \quad (4.10)$$

where $G_0 = G - v^i n_i$ is the local speed of propagation of the surface $S(t)$ in space–time, which coincides with G in the instantaneous rest frame. In view of (3.3), (4.2), and (4.10) we get in the local instantaneous rest frame

$$c\delta[Z] = \bar{\beta}^2 \frac{\delta}{\delta t}[Z], \quad (4.11)$$

where $\delta/\delta t$ is Thomas' delta derivative.[12] From (4.9) and (4.10) we get

$$G_0^2 = a_e^2/\beta^2\bar{\beta}^2(\sigma - a_e^2/c^2). \quad (4.12)$$

In an instantaneous rest frame, (4.12) assumes the form

$$G_0^2 = a_e^2/\sigma. \quad (4.13)$$

The velocity of propagation G_0 given by (4.12) in an instantaneous rest frame is in full agreement with earlier results of Ram and Gaur[13] and McCarthy[6] in particular cases.

If the medium is in a uniform state of rest ahead of the wave front and if the motion is studied in the rest frame of this uniform state, the speed of propagation G_0 is a constant.

V. THE GROWTH EQUATION

In this section we shall derive a fundamental growth equation which will govern the growth and decay of a weak discontinuity during its course of propagation. The medium ahead of the wave is assumed uniform and at rest.

Now we define the amplitude $b(t)$ of the wave $\Sigma(x^\mu)$ by the relation

$$b = c\lambda = c\lambda^\alpha N_\alpha = c\lambda^\alpha \overset{*}{N}_\alpha, \quad (5.1)$$

where $\overset{*}{N}{}^\alpha = S^{\alpha\beta}N_\beta$ are the spacelike components of N^α.

Differentiating (2.4), (2.9) and (2.13) with respect to x^β and taking jumps across $\Sigma(x^\mu)$ with the help of (3.4) and (3.5) we get

$$\rho\left\{\sigma + a_e^2\left(\frac{1+V^2}{V^2c^2}\right)\right\}\delta(\lambda) + \rho V\left\{\sigma - a_e^2\left(\frac{1+V^2}{V^2c^2}\right)\right\}$$
$$\times\bar{\lambda}^\alpha N_\alpha - \rho a_e^2\left(\frac{1+V^2}{Vc^2}\right)x^\tau_\alpha \lambda^\alpha_{;\tau} - \rho\sigma\lambda^\alpha\delta(N_\alpha)$$
$$+ \frac{a_e^2}{Vc^2}\lambda N_\alpha S^{\alpha\gamma}\delta\left(\frac{\rho N_\gamma}{V}\right)$$
$$+\rho\lambda^2\left\{(\gamma_e - 1)a_e^2\left(\frac{1+V^2}{V^2c^2}\right) + 2a_e^2\left(\frac{1+V^2}{V^2c^2}\right)\right.$$
$$\left.-\rho\left(\frac{\gamma_e-1}{1+\alpha}\right)\frac{a_e^2}{c^2} - \frac{2a_e^2}{c^2}\right\} + \rho\lambda(La_e^2 + M)$$
$$\times\left(\frac{1+V^2}{V^2c^2}\right) - \frac{F}{V}\frac{\beta^3}{c^5}v_k\lambda^k N_\gamma N_\alpha S^{\alpha\gamma} = 0, \quad (5.2)$$

where

$$L = \frac{\beta}{c}\frac{4\rho D_0^2 K_f}{3R^2T_d^2}\left[\{\rho D_0(1+\alpha)^2 - 3p\}\right.$$
$$\times\left\{\rho_d(1-\alpha)\exp\left(\frac{-T_d}{T}\right)\frac{T_d}{PT}\right\}$$
$$\left.-3\left\{\rho_d(1-\alpha)\exp\left(\frac{-T_d}{T}\right) - \rho\alpha^2\right\}\right],$$

$$M = \frac{\beta}{c}\frac{4D_0^2K_r}{3R^2T_d^2}\left[\left[\left\{\rho_d(1-\alpha)\exp\left(\frac{-T_d}{T}\right) - \rho\alpha^2\right\}\right.\right.$$
$$\times\{2\rho D_0(1+\alpha)^2 - 3p\} - \rho\alpha^2\{\rho D_0(1+\alpha)^2 - 3p\}$$
$$-\{\rho D_0(1+\alpha)^2 - 3p\}$$
$$\left.\left.\times\left\{\rho_d(1-\alpha)\exp\left(\frac{-T_d}{T}\right)\frac{T_d}{T}\right\}\right]\right].$$

In view of (4.9), the coefficient of $\bar{\lambda}^\alpha N_\alpha$ in (5.2) vanishes and, therefore, we get the following equation to be satisfied by λ:

$$\rho\left(2\sigma - \frac{a_e^2}{c^2}\right)\delta(\lambda) - \frac{\rho a_e^2}{Vc^2}\left\{x^\tau_\alpha \overset{*}{N}{}^\alpha\lambda_{;\tau} + \lambda\overset{*}{N}{}^\alpha_{;\alpha}\right\}$$
$$-\rho\sigma\lambda(1+V^2)^{-1}\overset{*}{N}{}^\alpha\delta(N_\alpha) + \frac{a_e^2}{Vc^2}\lambda N_\alpha S^{\alpha\gamma}\delta\left(\frac{\rho N_\gamma}{V}\right)$$
$$+\rho\lambda^2\left\{(\gamma_e+1)\sigma - \left(\frac{3(\gamma_e-1)}{1+\alpha}+2\right)\frac{a_e^2}{c^2}\right\} + \rho\lambda\frac{\beta}{c}$$
$$\times(La_e^2 + M)\frac{\sigma}{a_e^2} - \frac{F}{V}\frac{\beta^3}{c^5}v_k\lambda^k(1+V^2) = 0, \quad (5.3)$$

which is the required growth equation governing the global behavior of the amplitude $b(=c\lambda)$ of a relativistic wave in dissociating gases. In a local instantaneous rest frame for which $\overset{*}{N}{}^\alpha = (1+V^2)^{1/2}(n^i, 0)$, Eq. (5.3) takes on a particularly simple form

$$A\frac{\delta b}{\delta t} - (\Omega - E)b + Bb^2 = 0, \quad (5.4)$$

where

$$A = G_0(2\sigma - a_e^2/c^2)/2a_e^2(1 - G_0^2/c^2),$$
$$B = (G_0/2a_e^2)[(\gamma_e + 1)\sigma - \{3(\gamma_e - 1)/(1+\alpha) + 2\}a_e^2/c^2,$$
$$E = \sigma(La_e^2 + M)G_0/2a_e^4,$$

and

$$\Omega = -\frac{1}{2}\frac{\partial n^i}{\partial x^i}.$$

Here Ω is the mean curvature of the propagating surface $S(t)$ in space–time.

VI. LOCAL AND GLOBAL BEHAVIOR OF THE WAVE

In this section we shall study the local and global behavior of a weak discontinuity in an instantaneous rest frame. If S denotes the distance traversed by the wave along its normal trajectory in time t, we have

$$\frac{\delta S}{\delta t} = G_0, \tag{6.1}$$

which provides us a relation $S = G_0 t$, where G_0 is the constant speed of the wave front propagating in a uniform state ahead of it in the rest frame of this uniform state.

The mean curvature Ω of the wave surface $S(t)$ is a function of distance S and has been calculated in the form[14]

$$\Omega = \frac{\Omega_0 - K_0 S}{1 - 2\Omega_0 S + K_0 S^2}, \tag{6.2}$$

where Ω_0 and K_0 are the values of the mean and Gaussian curvatures of the initial wave front.

Using (6.1) and (6.2) in (5.4) and integrating we get

$$b(S) = b_0\phi(S)\left\{1 + \frac{Bb_0}{AG_0}\int_0^S \phi(S)\,dS\right\}^{-1}, \tag{6.3}$$

where

$$\phi(S) = e^{-ES/G_0}\{(1 - k_1 S)(k - k_2 S)\}^{-1/2AG_0}.$$

Here b_0 is the initial wave amplitude at time $t = 0$ and k_1, k_2 are the principal curvatures of the initial wave front.

Now we consider the case of a diverging wave for which k_1 and k_2 are negative. When $b_0 > 0$, the equation (6.3) shows that the amplitude $b(t)$ will continuously decrease as t increases and tends to zero as $t \to \infty$. Under these conditions a weak wave will decay and will be damped out ultimately.

When $b_0 < 0$ (the case of a compressive wave), there exists a critical value b_c of $|b_0|$ given by

$$b_c = \left\{\frac{A}{B}\int_0^\infty \phi(t)\,dt\right\}^{-1} \tag{6.4}$$

such that

(i) when $|b_0| < b_c$, $\lim_{t\to\infty} b(t) = 0$,

(ii) when $|b_0| > b_c$, there exists a finite critical time t_c given by

$$\int_0^{t_c} \phi(t)\,dt = A/B\,|b_0| \tag{6.5}$$

such that

$$\lim_{t\to t_c} |b(t)| = \infty.$$

In this case there shall occur a breakdown of a weak wave at time $t = t_c$ and consequently a shock wave will be formed at the cusp of intersecting characteristics.

In order to study the effects of dissociation on the behavior of a weak wave, we observe that

$$\frac{db_c}{dE} = \frac{A}{Bb_c^2}\int_0^\infty t\phi(t)\,dt > 0, \tag{6.6}$$

$$\frac{dt_c}{dE} = \frac{1}{\phi(t_c)}\int_0^{t_c} t\phi(t)\,dt > 0. \tag{6.7}$$

The Eqs. (6.6) and (6.7) show that the dissociation effects will delay the process of shock formation and thus it has a stabilizing effect on the weak wave propagation.

VII. SPECIAL CASES

For a nondissociating perfect gas model ($\alpha = 0, D_0 = 0$ the solution (6.3) for the wave amplitude $b(s)$ reduces to

$$b = b_0(1 - 2\Omega_0 S + k_0 S^2)^{-c_1} \times\left\{1 + c_2 b_0\int_0^S (1 - 2\Omega_0 S + k_0 S^2)^{-c_1}\,dS\right\}^{-1}, \tag{7.1}$$

where

$$c_1 = (1 + 2\tau)(2 + 5\tau)^{-1},$$
$$c_2 = a^{-1}(\tfrac{7}{3} + 4\tau)(1 + 2\tau)(2 + 5\tau)^{-1}(1 + 3\tau)^{-1/2}.$$

Here $\tau = a^2/c^2$ is the dimensionless parameter of relativistic effects and varies from zero to one where $a = (4p/3\rho)^{1/2}$ is the sound speed for a perfect gas with $\gamma = 4/3$.

For a plane wave ($\Omega = 0$), the solution (7.1) assumes a simpler form,

$$b = b_0\{1 + c_2 b_0 S\}^{-1}, \tag{7.2}$$

which agrees with the result of McCarthy.[6]

The solution (7.2) suggests that the weak expansion waves ($b_0 > 0$) will decay faster under relativistic effects, since c_2 lies between $1.17a^{-1}$ and $1.457a^{-1}$. For compressive waves ($b_0 < 0$), the critical time t_c for the shock formation is given by

$$t_c = (1 + 3\tau)(2 + 5\tau)/|b_0|(7/3 + 4\tau)(1 + 2\tau),$$

which shows that t_c increases with τ. When $\tau = 0$ for nonrelativist case, $c_1 = \frac{1}{2}$ and $c_2 = 7/6a$. This case corresponds to a perfect gas model with $\gamma = 4/3$ in classical gas dynamics.

For nonrelativistic case of ideal dissociating gas model, the solution (6.3) for a plane wave reduces to

$$b(s) = b_0 e^{-ES}\{1 + (1 - e^{-ES})(\gamma_e + 1)b_0/2Ea_e\}^{-1}, \tag{7.3}$$

which suggests that $b(S)$ decreases monotonically to zero for $b_0 > 0$. On the other hand, there exists a critical amplitude $b_c = 2Ea_e(\gamma_e + 1)^{-1}$ such that

(i) when $|b_0| < b_c$, $\lim_{s\to\infty} b(s) = 0$,
(ii) when $|b_0| > b_c$, $\lim_{s\to s_c} b(s) = \infty$,

where

$$S_c = a_c t_c = \log\{1 - b_c/|b_0|\}^{-1/E}.$$

This shows that under dissociation effects all compressive weak waves will not terminate into shock waves, whereas for a nondissociating perfect gas model all weak compressive waves terminate into shock waves. The effects of dissociation play a stabilizing role in the sense that they delay the shock formation by increasing t_c.

[1]C. Eckart, Phys. Rev. **58**, 919 (1940).

[2]A.H. Taub, Phys. Rev. **74**, 328 (1948).

[3]G.S. Saini, Proc. R. Soc. **A 260**, 61–78 (1961).

[4]N. Coburn, J. Math. Mech. **10**, 361–92 (1961).

[5]R.P. Kanwal, J. Math. Mech. **15**, 379–92 (1966).

[6]M.F. McCarthy, Int. J. Eng. Sci. **6**, 209–16 (1969).

[7]R.A. Grot and A.C. Eringen, Int. J. Eng. Sci. **4**, 611–70 (1966).

[8]M.J. Lighthill, J. Fluid Mech. **2**, 1 (1957).

[9]J.F. Clarke, J. Fluid. Mech. **7**, 577 (1960).

[10]N.R. Sibgatullin and O.Yu Dinariev, Sov. Phys. JETP **46**, 840 (1977).

[11]C. Truesdell and R.A. Toupin, The classical Field Theories, in *Handbuch der physik*, **111/1** (Springer, New York and Berlin, 1960).

[12]T.Y. Thomas, J. Math. Mech. **6**, 311 (1957).

[13]R. Ram and M. Gaur, Acta Physica Hungarica, Tomas **42**, (2), 85–94 (1976).

[14]T.Y. Yhomas, *Concepts from Tensor analysis and Differential Geometry*, (Academic, New York 1963) 2nd edition.

On the scattering of plane electromagnetic waves off cylindrically confined cold plasmas with overdense and steep densities

Gregory A. Kriegsmann
Department of Mathematics and Statistics, University of Nebraska at Lincoln, Lincoln, Nebraska 68588

(Received 27 March 1979; accepted for publication 17 May 1979)

In this paper we study the scattering of plane electromagnetic waves off a cylindrically confined cold plasma. The plasma density is taken to be overdense and very steep. This causes the cutoff radius, r_0, to be within a fraction of a wavelength from the cylindrical boundary. For simplicity, we assume two types of incident polarization. In both cases scalar second-order elliptic partial differential equations describe the fields. These problems are studied in the asymptotic limit $a\omega/c \to \infty$ with $0 < 1 - r_0 = O(c/a\omega)$. Here a is the radius of the cylinder, ω is the frequency of the incident radiation, and c is the velocity of light in free space. We develop an asymptotic technique which reduces the partial equations to ordinary differential equations within the plasma. Our method is a blend of geometrical optics and boundary layer techniques. Outside the plasma we use straight geometrical optics to describe the scattered field.

1. INTRODUCTION

The conversion of electromagnetic energy into kinetic energy is a major factor in the laser fusion concept.[1] As a first step to understanding this process, one linearizes the pertinent equations, neglects ionic motion, and assumes a cold plasma. The ensuing equations give rise to a linear scattering problem. This problem has been extensively studied when the plasma is planar.[2,3] Recently, the case of a spherical plasma target has received considerable attention.[1,4] The interest in this geometrical configuration arises from the fact that the plasma pellet is initially spherical in shape.

The scattering of plane waves off cylindrically confined plasmas has likewise received a large amount of attention.[5,6] The reasons for considering this geometry are twofold: Cylindrical plasmas can be made in the laboratory and the cylinder can be thought of as a compromise between the spherical and planar shapes.

We took the later view and studied[6] the scattering problem for an overdense plasma column of radius a. In that work the plasma density was assumed to be quadratic in the radial variable $r' = ra$. (here r is the dimensionless variable.) Moreover, the frequency of the incident plane wave (ω) was fixed to insure the vanishing of the refractive index on the cylinder $r = r_0 < 1$. We applied the method of geometrical optics to approximate both the scattered field and the field within the plasma in the limit as $(a\omega/c) \to \infty$ (where c is the velocity of light in free space). These results become invalid when $0 < 1 - r_0 \sim O(c/\omega a)$. This condition is satisfied when the density profile is sufficiently steep. Such profiles occur when an infrared laser initially irradiates a plasma target.[1]

In this paper we shall assume that the profile is steep enough to give $1 - r_0 = O(c/\omega a)$. We shall again fix ω and seek asymptotic approximations to the fields in the limit as $(a\omega/c) \to \infty$. To simplify our analysis, we shall consider only two types of incident polarization. The first choice is to orient the incident electric field parallel to the axes of the cylinder (the E problem). The second choice is to orient the incident magnetic field parallel to the axis of the cylinder (the B problem). The mathematical formulation of these problems gives rise to scalar second-order elliptic partial differential equations which describe the appropriate fields.

In this work we develop an asymptotic scheme which exploits the smallness of $1 - r_0$ and the largeness of $(a\omega/c)$. Our method reduces the partial differential equations into ordinary differential equations within the plasma. It is basically a blend of geometrical optics[7] and boundary layer techniques.[8] Outside the plasma the scattered wave is approximated using straight geometrical optics.

The first part of this paper is concerned with the E problem. Sections 2 and 3 are devoted to the calculation of the electric field, both interior and exterior to the plasma. We find that the magnitude of the scattering cross section is identical to the magnitude of the cross section for a metal cylinder of radius a. The second part is concerned with the B problem. In Sec. 4 we compute the magnetic field in the entire plane. Our calculations show that the amplitude of the scattering cross section is the same as in the E problem. The phases are different. Section 5 is concerned with the effect of damping on the cross section. We find that its magnitude $|A|$ is reduced by an amount proportional to the energy absorbed. Moreover, our numerical computations show that $|A|$ has a minimum at a certain angle. A maximum amount of energy is absorbed by the plasma at this angle. These results agree with those given in Ref. 9.

2. FORMULATION OF THE E PROBLEM

A high frequency plane electromagnetic wave impinges upon a cylindrically confined "cold" plasma of infinite extent and scatters from it. This cylinder is circular in cross section with radius a. We further assume that the incident radiation is polarized with the electric field parallel to the axis of the cylinder. It follows from this assumption and the equations governing the plasma[10] that the time harmonic electric field $u\exp(-i\omega t)$ remains in this direction and satisfies the Helmholtz equation

$$\nabla^2 u + k^2 n u = 0. \tag{2.1}$$

 0022-2488/80/051251-05$01.00

In this equation, $k = k'a$ and the index of refraction n is given by[10]

$$n = 1 - 4\pi e^2 N(ar)/m\omega^2, \tag{2.2}$$

where ω is the frequency of the incident plane wave, e is the charge of an electron with mass m, r is the radial variable, and $N(ar)$ is the charge density. Implicit in Eq. (2.2) are the assumptions that ω is large enough to neglect ionic motion and the density is dependent only upon r.

We now assume that the plasma is overdense, is confined to the region $0 \leqslant r \leqslant 1$, and is very steep. These assumptions lead us to make the following hypotheses about $n(r)$:

(H1) $n(r) = 1,\ r \geqslant 1,$

(H2) $n(r_0) = 0,\quad 0 < 1 - r_0 = O(1/k),$

(H3) $n'(r_0) = mk,\quad m = O(1),$

(H4) $n^{(l)}(r_0) = o(k^l).$

These conditions are met when an infrared laser initially irradiates an overdense plasma target.[1]

To complete the mathematical statement of this problem, we must impose further conditions. First we demand that u and ∇u are continuous everywhere. Secondly, the scattered field must satisfy the radiation condition. Finally, we choose the x axis to be parallel to the incident wave vector $\mathbf{k}'$ and the z axis to be that of the cylinder.

We shall now suppose that $k \gg 1$, which corresponds to the physical situation mentioned above. Thus, we seek an asymptotic approximation of u as $k \to \infty$. At first this seems to be a natural setting for the method of geometrical optics. However, the cutoff radius is a fraction of a wavelength away from the boundary of the plasma (H2). Thus, geometrical optics cannot be used directly to determine an asymptotic approximation of u within the plasma.

3. AN ASYMPTOTIC METHOD

The largeness of k and smallness of $1 - r_0$ will now be exploited to change Eq. (2.1) into an ordinary differential equation. The field within the plasma is assumed to be of the form

$$u(r,\theta,k) = e^{ik\cos\theta}[\phi(\bar{r},\theta) + O(1/k)], \tag{3.1}$$

as $k \to \infty$, where the boundary layer variable $\bar{r}$ is defined by[8]

$$\bar{r} = km(r - r_0), \tag{3.2}$$

and θ is the polar angle. The index of refraction within the plasma is expressed in terms of $\bar{r}$ as

$$n(r) = \bar{r} + o(1), \tag{3.3}$$

as $k \to \infty$. This follows from Eqs. (H2)–(H4) and from expanding n in a Taylor series about $r = r_0$. Upon inserting Eqs. (3.1)–(3.3) into Eq. (2.1), it is found that ϕ satisfies

$$\frac{d^2\phi}{d\bar{r}^2} + \gamma^2(\bar{r} - \sin^2\theta)\phi = 0, \tag{3.4}$$

where $\gamma = 1/m$. This is Airy's equation; it arises when a plane wave impinges upon a plasma slab[3] with a linear refractive index.

Outside the plasma column ($r > 1$) the field is assumed to be of the form

$$u = e^{ikx} + e^{ik\psi(x,y)}[A(x,y) + O(1/k)] \tag{3.5}$$

as $k \to \infty$. The first term is the incident wave and the second is the scattered field. Upon inserting Eq. (3.5) into (2.1) and equating the coefficients of like powers of k to zero, it is found that ψ and A satisfy

$$\nabla\psi\cdot\nabla\psi = 1 \quad \text{(eikonal equation)}, \tag{3.6}$$

$$2\nabla A\cdot\nabla\psi + A\nabla^2\psi = 0 \quad \text{(transport equation)}. \tag{3.7}$$

Thus, the scattered field will be approximated by the method of geometrical optics.[7]

Now to compute u we must specify boundary conditions for Eq. (3.4) and initial data for Eqs. (3.6) and (3.7). For a fixed $r < r_0$ it follows from Eq. (3.2) that $\bar{r} \to -\infty$ as $k \to \infty$. Thus, the limit of ϕ as $\bar{r} \to -\infty$ must be specified. Since Eq. (3.4) has one solution which grows exponentially and another that decays exponentially[11] as $\bar{r} \to -\infty$, the assumption of a bounded solution within the plasma implies

$$\lim_{\bar{r} \to -\infty} \phi(\bar{r}) = 0. \tag{3.8}$$

From Eqs. (3.1) and (3.5) and the assumption that u and $\partial u/\partial r$ are continuous at $r = 1$, it follows that

$$\psi(\cos\theta,\sin\theta) = \cos\theta, \tag{3.9}$$

$$1 + A(\cos\theta,\sin\theta) = \phi(1,\theta), \tag{3.10}$$

$$i\cos\theta + i\frac{\partial\psi}{\partial r}(\cos\theta,\sin\theta)A(\cos\theta,\sin\theta) = m\frac{\partial\phi}{\partial\bar{r}}(1,\theta). \tag{3.11}$$

In deriving Eq. (3.11) a term of order $O(1/k)$ was neglected and $\bar{r}$ was set equal to 1 at the plasma boundary. The later approximation follows from Eqs. (H1) and (3.3) and introduces an error of $O(1/k)$; we have consistently neglected terms of this order.

From Eqs. (3.6) and (3.9) it follows that

$$\frac{\partial\psi}{\partial r}(\cos\theta,\sin\theta) = |\cos\theta|. \tag{3.12}$$

When this result is inserted into Eq. (3.11) we find from Eqs. (3.10) and (3.11) that

$$\frac{\partial\phi}{\partial\bar{r}}(1,\theta) - i\gamma|\cos\theta|\phi(1,\theta) = i\gamma(\cos\theta - |\cos\theta|), \tag{3.13}$$

$$A(\cos\theta,\sin\theta) = \phi(1,\theta) - 1. \tag{3.14}$$

Equations (3.4), (3.8), and (3.13) are readily solved to give

$$\phi(r,\theta) = \frac{2i\gamma\cos\theta A_i(-\xi)}{i\gamma\cos\theta A_i(-\xi_1) - \gamma^{2/3}A_i'(-\xi_1)}, \tag{3.15}$$

where $\xi = \gamma^{2/3}(\bar{r} - \sin^2\theta)$, $\xi_1 = \gamma^{2/3}\cos^2\theta$, A_i is the Airy function, and θ is in the interval $[\pi/2, 3\pi/2]$. When $|\theta| \leqslant \pi/2$, the boundary condition (3.13) becomes homogeneous; we prove in Appendix A that $\phi(r,\theta)$ is identically zero under this condition.

Combining Eqs. (3.14) and (3.15) gives the initial value of the scattered wave as

$$A(\cos\theta,\sin\theta) = A_0(\theta) = \frac{i\gamma\cos\theta A_i(-\xi_1) + \gamma^{2/3}A_i'(-\xi_1)}{i\gamma\cos\theta A_i(-\xi_1) - \gamma^{2/3}A_i'(-\xi_1)}, \tag{3.16}$$

for $\pi/2 \leqslant \theta \leqslant 3\pi/2$ and

$$A(\cos\theta,\sin\theta) \doteq A_0(\theta) = -1, \tag{3.17}$$

for $|\theta| \leqslant \pi/2$.

Equations (3.9), (3.16), and (3.17) are the initial data required to solve the eikonal and transport equations. These equations are easily solved by making the following observation: The initial phase given by Eq. (3.9) is the same phase that would occur if a plane wave impinged upon a metal cylinder of unit radius. Thus, the reflected rays and phase of our plasma problem are identical to those of the irradiated metal cylinder. Since the rays determine the expansion ratio[7] which is proportional to any solution of Eq. (3.7), our amplitude is given by

$$A(x,y) = -A_0(\theta)A_c(x,y), \tag{3.18}$$

where $A_c(x,y)$ is the amplitude for the metal cylinder problem. For completeness we compute ψ, the rays, and A_c in Appendix B. We state here the far field result

$$u(x,y) \simeq e^{ikx} + S(\Phi)(e^{ikr}/r^{1/2}) \tag{3.19}$$

as $r \to \infty$, where $S(\Phi)$ is defined by

$$S(\Phi) = +\left(\tfrac{1}{2}\sin\frac{\Phi}{2}\right)^{1/2} A_0\left(\pi/2 + \frac{\Phi}{2}\right) \times \exp\left(-2ik\sin\frac{\Phi}{2}\right), \tag{3.20}$$

and Φ is the polar angle of the far field point.

The far field result becomes inaccurate as Φ approaches zero; the observation point then lies in the shadow region which is devoid of scattered rays. It is easy to deduce from Eqs. (3.9) and (3.17) that $u = 0$ in this region. From Eqs. (3.16) and (3.20) it follows that the magnitude of $S(\Phi)$ is given by

$$|S(\Phi)| = [\tfrac{1}{2}\sin(\Phi/2)]^{1/2}, \tag{3.21}$$

which is the same as the result for a metal cylinder. However, the phase is different. We have evaluated Eq. (3.16) numerically for $\gamma = 1,2$ and plotted the phase of $A_0[(\pi/2) + (\Phi/2)]$ for these cases in Fig. 1.

4. FORMULATION AND RESULTS FOR THE B PROBLEM

In this section we shall examine the scattering problem

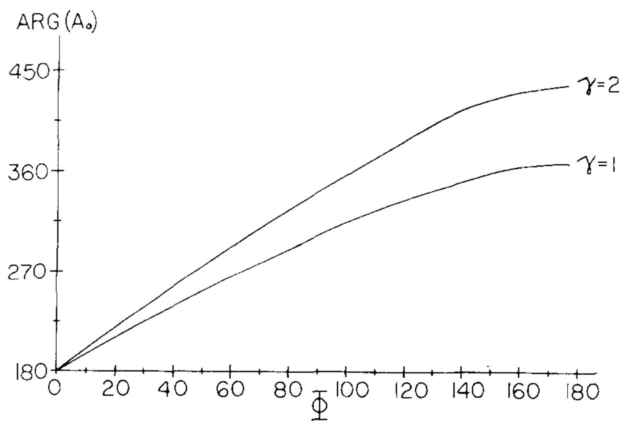

FIG. 1. The phase of $A_0[(\pi/2) + (\Phi/2)]$ for the E problem with $\gamma = 1,2$, where Φ is the polar angle of the far field point.

for a cylindrically confined cold plasma with a different incident polarization. We consider the case where the incident magnetic field is parallel to the cylinder's axis. From the cold plasma equations[10] we find that the time harmonic magnetic field $u\,e^{-i\omega t}$ remains in this direction and satisfies

$$\nabla^2 u - \frac{\nabla u \cdot \nabla n}{n} + k^2 n u = 0. \tag{4.1}$$

The coordinate axes, boundary conditions, and hypotheses on n are the same as in the E problem.

We again assume that u is given by Eq. (3.1) within the plasma and by Eq. (3.5) outside the plasma. When Eqs. (3.1) and (3.5) are inserted into Eq. (4.1), we find, as before, that ψ and A satisfy Eqs. (3.6) and (3.7), respectively. However, ϕ must now satisfy

$$\frac{d^2\phi}{d\bar{r}^2} - \frac{1}{r}\frac{d\phi}{d\bar{r}} + \gamma^2(\bar{r} - \sin^2\theta)\phi = 0. \tag{4.2}$$

The boundary conditions on ϕ, ψ, and A are the same as before.

Equation (4.2) arises when an electromagnetic plane wave (of a particular polarization) impinges upon a cold plasma slab.[3,12] It has been approximately solved by Denisov[12] using asymptotic techniques when $\gamma \gg 1$. His boundary conditions were Eq. (3.8) and $\phi(1,\theta) = 1$. Freidberg, Mitchell, Morse, and Rudsinski[9] have solved Eq. (4.2) numerically and presented the results for $\gamma = 1,10$. They prescribed ϕ at $\bar{r} = 1$ and used Eq. (3.8). To the best of our knowledge Eq. (4.2) cannot be solved in terms of tabulated functions.

For our particular application, γ is $O(1)$ and $\phi(1,\theta)$ is unknown. Once $\phi(1,\theta)$ is obtained, the initial scattering amplitude is given by Eq. (3.14). We cannot ascertain this information from the results of Denisov and Freidberg. However, we can deduce two important facts from Eqs. (4.2), (3.8), and (3.13). When $|\theta| \leqslant \pi/2$, the boundary condition (3.13) becomes homogeneous and $\phi(\bar{r},\theta) \equiv 0$ (see Appendix C). Thus, for this angular range Eq. (3.14) gives

$$A_0(\theta) = -1, \quad |\theta| \leqslant \pi/2. \tag{4.3}$$

We also prove in Appendix C that

$$|A_0(\theta)| = 1, \quad \pi/2 \leqslant \theta \leqslant \pi, \tag{4.4}$$

or equivalently $|\phi(1,\theta) - 1| = 1$. The phase of $A_0(\theta)$ and $\phi(1,\theta)$ must be determined from a numerical solution of Eq. (4.2).

We have numerically solved Eq. (4.2) using the finite difference method.[13] To perform the calculations we chose a step size of 1.0/10.5 and replaced Eq. (3.8) by $\phi(-9,\theta) = 0$. To test the sensitivity of our results to this approximate boundary condition we replaced -9.0 by -15.0, kept the step size fixed, and reran our program. We found that the numerical values changed insignificantly.

We will not present the totality of our numerical results here. It is sufficient to state that our results do satisfy Eq. (4.4) and give the phase of $A_0(\theta)$. Since the scattered field is computed as before and is given by Eq. (3.20), the phase of $S(\Phi)$ is given by

$$\mathrm{ph}[S(\Phi)] = \mathrm{ph}\left[A_0\left(\frac{\Phi}{2} + \frac{\pi}{2}\right)\right] - 2k\sin\frac{\Phi}{2}. \tag{4.5}$$

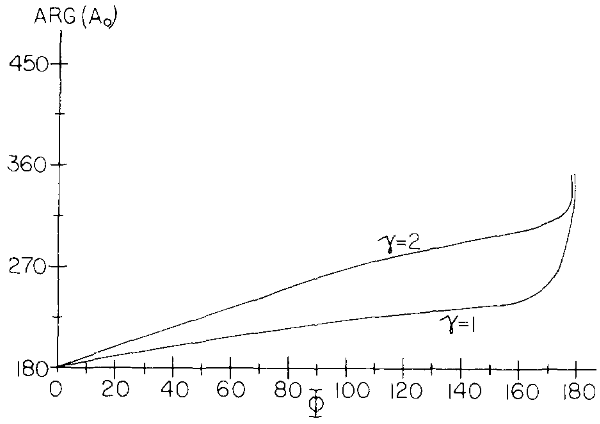

FIG. 2. The phase of $A_0[(\pi/2)+(\Phi/2)]$ for the B problem with $\gamma=1,2$, where Φ is the polar angle of the far field point.

The phase of $A_0[(\pi/2)+(\Phi/2)]$ is plotted in Fig. 2 for $\gamma=1$ and 2. The amplitude of $S(\Phi)$ is again given by Eq. (3.21); this follows from Eq. (4.4). Thus, the cross section $S(\Phi)$ of the B problem differs only in phase from the cross section of the E problem.

5. THE EFFECT OF DAMPING

In this section we briefly discuss the effect of damping on the cross section $S(\Phi)$ of the B problem. The addition of a small amount of damping relaxes Eq. (H2) and gives

$$n(r_0)=i\delta,\quad 0<1-r_0<O(1/k), \tag{H2$'$}$$

where $\delta>0$. When this is incorporated into our asymptotic scheme, we find that ϕ satisfies Eq. (4.2) with $\bar r$ replaced by $\bar r+i\delta$. In Appendix C we show that the modulus of $A_0(\theta)$ is given by

$$|A_0(\theta)|^2=1-\frac{\delta}{\gamma|\cos\theta|}(\|\phi'\|^2+\gamma^2\sin^2\theta\,\|\phi\|^2 z),$$
$$\pi/2<\theta<\pi, \tag{5.1}$$

where

$$\|f\|^2=\int_{-\infty}^{1}\frac{|f(r)|^2}{r^2+\delta^2}\,dr.$$

When $|\theta|\leqslant\pi/2$, we find as before that $A_0=-1$.

Since the cross section has $A_0[(\pi/2)+(\Phi/2)]$ as a factor, the effect of damping diminishes the amplitude of the scattered wave. The bracketed term in Eq. (5.1) is the energy absorbed per unit of angle. We have plotted $|A_0[(\pi/2)+(\Phi/2)]|$ for the cases $\delta=0.05$ ($\gamma=1,2$) and $\delta=0.10$ ($\gamma=1,2$) in Figs. 3 and 4. It is interesting to note that $|A_0|$ has a minimum at an angle which depends upon δ and γ. This angle corresponds to a maximum amount of absorption and agrees with the results of Freidberg *et al.*

APPENDIX A

In this section we prove that Eq. (3.4) is subject to Eq. (3.8) and

$$\frac{d\phi}{d\bar r}(1,\theta)-i\gamma\cos\theta\phi(1,\theta)=0,\quad |\theta|\leqslant\frac{\pi}{2} \tag{A1}$$

has only the trivial solution $\phi(\bar r,\theta)=0$. The general solution of Eq. (3.4) subject to Eq. (3.8) is given by

$$\phi=c(\theta)A_i(-\xi), \tag{A2}$$

where $\xi=\gamma^{2/3}(\bar r-\sin^2\theta)$ and A_i is the Airy function. When this is inserted into Eq. (A1) we obtain

$$c(\theta)[+\gamma^{2/3}A_i'(-\xi_i)+i\gamma A_i(-\xi_1)]=0, \tag{A3}$$

where $\xi_1=\gamma^{2/3}\cos^2\theta$. Since A_i', A_i, and γ are real, $c(\theta)$ equals zero and $\phi(\bar r,\theta)=0$. From Eq. (3.14) it follows that $A_0(\theta)=-1$ for $|\theta|\leqslant\pi/2$.

APPENDIX B

Consider a plane scalar wave impinging upon a metal cylinder and scattering from it. The geometrical optics approximation to the scattered wave is given by

$$u(x,y,k)=[A_c(x,y)+O(1/k)]e^{ik\psi(x,y)}, \tag{B1}$$

where the amplitude A_c satisfies Eq. (3.7) and the phase ψ satisfies Eq. (3.6). Since the total field vanishes on the cylinder, $A_c(\cos\theta,\sin\theta)=-1$ and $\psi=\cos\theta$, where $(\cos\theta,\sin\theta)$ is the intersection point of the incident ray and the cylinder. From Eq. (3.6) and the initial data it follows that the scattered rays satisfy the law of reflection. Thus, they are given by

$$X=-\sigma\cos2\theta+\cos\theta,$$
$$Y=-\sigma\sin2\theta+\sin\theta, \tag{B2}$$

where σ is the arclength. These rays form a one parameter family of straight lines.

The eikonal and transport equations are readily solved to give

$$\psi=\sigma+\cos\theta, \tag{B3}$$

$$A_c=-\left[\frac{J(0,\theta)}{J(\sigma,\theta)}\right]^{1/2}, \tag{B4}$$

where $J(\sigma,\theta)$ is the Jacobian of the ray map $(\sigma,\theta)\mapsto(x,y)$ given in Eq. (B2). It is easily found to be

$$J=4\sigma-2\cos\theta. \tag{B5}$$

Now in the far field $\sigma\gg1$. From Eqs. (B2)–(B5) we deduce that

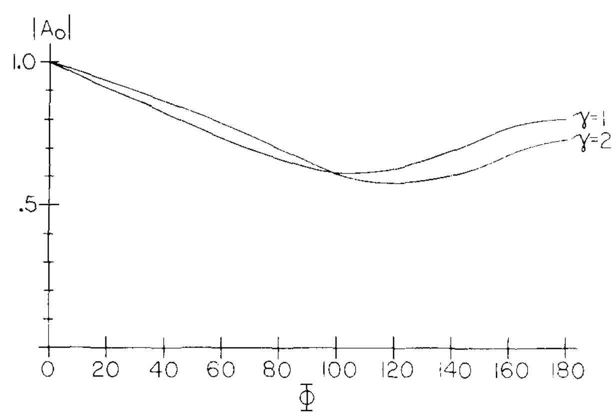

FIG. 3. The amplitude of $A_0[(\pi/2)+(\Phi/2)]$ for the B problem with $\gamma=1,2$ and $\delta=0.10$, where δ is the damping and Φ is the polar angle of the far field point.

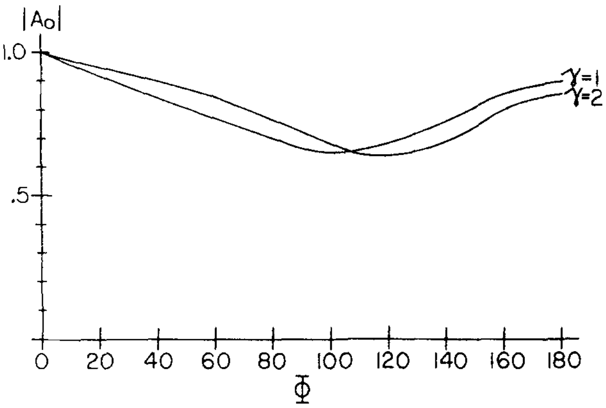

FIG. 4. The amplitude of $A_0[(\pi/2)+(\Phi/2)]$ for the B problem with $\gamma = 1,2$ and $\delta = 0.05$, where δ is the damping and Φ is the polar angle of the far field point.

$$\sigma \simeq r + \cos\theta,\quad J \simeq 4R + 2\cos\theta,\quad \psi \simeq r + 2\cos\theta,$$
$$A \simeq (-2\cos\theta)^{1/2},\quad \Phi \simeq 2\theta - \pi, \tag{B6}$$

where r and Φ are the polar coordinates of the far field point. when Eq. (B6) is inserted into Eq. (B1) we obtain the far field result

$$u \simeq \left(\tfrac{1}{2}\sin\frac{\Phi}{2}\right)^{1/2} \exp\left(-2ik\sin\frac{\Phi}{2}\right)\frac{e^{ikr}}{\sqrt{r}}, \tag{B7}$$

as $r \to \infty$.

APPENDIX C

Let $\phi(\bar r,\theta)$ be the solution of

$$L\phi = \frac{d^2\phi}{d\bar r^2} - \frac{1}{\bar r + i\delta}\frac{d\phi}{d\bar r} + \gamma^2(\bar r + i\delta - \sin^2\theta)\phi = 0, \tag{C1}$$

which satisfies Eqs. (3.8) and (3.13). Then the complex conjugate of ϕ, i.e., ϕ^*, satisfies

$$L^*\phi^* = \frac{d^2\phi^*}{d\bar r^2} - \frac{1}{\bar r - i\delta}\frac{d\phi^*}{d\bar r} + \gamma^2(\bar r - i\delta - \sin^2\theta)\phi^* = 0 \tag{C2}$$

subject to Eq. (3.8) and

$$\frac{d\phi^*}{d\bar r}(1,\theta) + i\gamma|\cos\theta|\phi^*(1,\theta) = -i\gamma(\cos\theta - |\cos\theta|). \tag{C3}$$

Using integration by parts and Eq. (3.8), we find that

$$0 = \int_{-\infty}^{1}(\phi^*L\phi - \phi L^*\phi^*)d\bar r = \frac{\phi^*(1,\theta)\phi'(1,\theta)}{1+i\delta} - \frac{\phi(1,\theta)\phi^{*\prime}(1,\theta)}{1-i\delta} + 2i\delta\|\phi'\|^2 + 2i\delta\gamma^2\sin^2\theta\|\phi\|^2, \tag{C4}$$

where

$$\|f\|^2 = \int_{-\infty}^{1}\frac{ff^*\,d\bar r}{\delta^2 + \bar r^2}.$$

Inserting Eqs. (3.13) and (C3) into Eq. (C4) and approximating $1+\delta^2$ by 1 gives

$$0 = \gamma|\cos\theta|\cdot|\phi(1,\theta)|^2 + \frac{\gamma}{2}(\cos\theta - |\cos\theta|) \times[\phi^*(1,\theta) + \phi(1,\theta)] + \delta\|\phi'\|^2 + \delta\gamma^2\sin^2\theta\|\phi\|^2. \tag{C5}$$

Now, when $|\theta| \leqslant \pi/2$, Eq. (C5) becomes

$$\gamma\cos\theta|\phi(1,\theta)|^2 + \delta\|\phi'\|^2 + \delta\gamma^2\sin^2\theta\|\phi\|^2 = 0. \tag{C6}$$

Since each term is positive, we deduce that $\phi(\bar r,\theta) = 0$ for $-\infty < \bar r \leqslant 1$. When $\pi/2 < \theta \leqslant \pi$ Eq. (C5) gives

$$|\phi(1,\theta) - 1|^2 = 1 - \frac{\delta}{\gamma|\cos\theta|}\{\|\phi'\|^2 + \gamma^2\sin^2\theta\|\phi\|^2\}. \tag{C7}$$

This result is equivalent to Eq. (5.1). Setting $\delta = 0$ gives Eq. (4.4).

[1]K.A. Brueckner and S. Jorna, Rev. Mod. Phys. **46**, 325 (1974).
[2]F.F. Chen and R.B. White, "Amplification of Electromagnetic Waves in Overdense Plasma," UCLA Plasma Physics Group Report 126 (1972).
[3]V.L. Ginzburg, *Propagation of Electromagnetic Waves in Plasma* (Gordon and Breach, New York, 1961).
[4]J.J. Thomson, C.E. Max, J. Erkkila, and J.E. Tull, Phys. Rev. Lett. **37**, 1052 (1976).
[5]N.G. Alexopoulos, J. Opt. Soc. Am. **9**, 1088 (1972).
[6]G.A. Kriegsmann, J. Math. Phys. **17**, 112 (1976).
[7]J.B. Keller and B.D. Seckler, J. Acoust. Soc. Am. **31**, 192 (1958).
[8]J.D. Cole, *Perturbation Methods in Applied Mathematics* (Blaisdell, Waltham, 1968).
[9]J.P. Freidberg, R.W. Mitchell, R.L. Morse, and L.I. Rudinski, Phys. Rev. Lett. **28**, 795 (1972).
[10]J. Brandstatter, *An Introduction to Waves, Rays, and Radiation in Plasma Media* (McGraw-Hill, New York, 1963).
[11]E.T. Copson, *Asymptotic Expansions* (Cambridge University, Cambridge, 1967).
[12]N.G. Denisov, Sov. Phys.-JETP **4**, 544 (1957).
[13]E. Isaacson and H.B. Keller, *Analysis of Numerical Methods* (Wiley, New York, 1966).

Roots of the modal equation for em wave propagation in a tropospheric duct

R. H. Ott

U.S. Department of Commerce, National Telecommunications and Information Administration, Institute for Telecommunication Sciences, Boulder, Colorado 80303

(Received 5 September 1979; accepted for publication 21 December 1979)

The solution of the modal equation for electromagnetic wave propagation in a tropospheric duct, having a trilinear refractivity profile, shows the existence of two sets of roots in the complex propagation plane. One of these is of the Whispering Gallery type with very low attenuation and the other is of the creeping wave variety with relatively high attenuation. When the contrast in refractive index is small (the so-called linear atmosphere) the creeping wave modes dominate. A large contrast in refractive index causes the Whispering Gallery modes to dominate (surface or elevated ducts). Several examples of the altitude charts for the modal equation for propagation at 200 MHz through a medium having various constrasts in the refractive index give some physical insight into these two classes of modes. In addition, asymptotic solutions for the Whispering Gallery Modes are derived and used as initial choices in Newton's method for the solution of the modal equation.

1. INTRODUCTION

A number of investigators have studied the problem of electromagnetic wave propagation in surface-based or elevated ducts. Much of the analysis discussed in this paper has its origin in the treatment of em wave propagation in a laterally nonuniform troposphere by Cho and Wait.[1] The fundamental problem in em propagation in a tropospheric duct is the solution of the modal equation, and this paper addresses this question. The general problem of finding the zeros of an analytic function of a complex variable or the poles of a function which is analytic except at isolated singularities, is not easy, especially when numerical results are required. The motivation for the altitude charts for the modal equation is to give some comprehension of the topology of a complicated complex function of a complex variable. Goodhart and Pappert[2] have used the winding number technique to locate the zeros of the modal equation and have successfully applied the technique to the case of ducting produced by a strong elevated layer [40 N unit deficit at an elevation of 200 m above the earth $N = (n-1)\times 10^6$, n = retractive index] for frequencies ranging from 65 MHz to 3.3 GHz. One obvious deficiency in the winding number technique occurs if there exists an equal number of poles and zeros in the search region of the complex plane. Actually, a single pole in the search area may cause problems regardless of the number of zeros present. What is required in the winding number technique is that all the desired roots lie inside of the selected boundary (search region).

At lower frequencies, the ionosphere becomes the region where possible ducting can occur and is generally anisotropic unless a zero magnetic field is assumed. The problem of solving the modal equation at these frequencies was investigated by Budden[3] (assuming perfectly conducting ground) and by Wait[4,5] assuming the anisotropy as a perturbation to the TE and TM modes for the isotropic case but accounting fully for the finite ground conductivity). In the chapter entitled "Characteristics of the Modes for V.L.F. Propagation", Wait discusses some of the problems related to finding the roots of the modal equation. In particular, in his discussion of the "earth detached mode," Wait notes that, if the magnitude of the complex number $(z_0 - z)$ is not much larger than unity, an additional complex function needs to be included in the modal equation. Here z_0 is related to the height of the lower boundary of the ionosphere and z is the complex number describing the topology of the mode equation. This additional feature gives rise to the Whispering Gallery effect for sound which was discovered by Lord Rayleigh[6] at the base of the dome of St. Paul's cathedral. In a similar fashion, we show, by numerical experiment, the existence of a set of roots in the complex z plane along a curve or a "trough of zeros" which has asymptotes tending toward the rays $\arg(z) = -\pi/3$ and $\arg(z) = -\pi$. Here $z \approx (\nu - ka)/(ka/2)^{1/3}$ where ν is the normalized complex wave number and ka is the circumference of the earth in wavelengths. A more general definition is given below but for the purposes of discussion here the attenuation rate of the modes is approximately $-\mathrm{Im}(\nu)/a$ n m^{-1} where a is the earth radius. The seperation of the zeros into two sets is not distinct, especially for small values of the refractive index contrast. Whether or not modes in each set are necessary for the field calculation is dependent upon (among other things) the particular geometry involved (transmitter-receiver location and layer configuration).

We also give asymptotic solutions for the roots of the modal equation along the ray $\arg(z) \simeq -\pi$. The asymptotic solutions for the roots along the ray $\arg(z) \simeq -\pi$ do provide good initial choices for Newton's method for solving the exact modal equation.

2. THE MODAL EQUATION FOR A PIECEWISE LINEAR ATMOSPHERE

Requiring the field in two dimensions and its normal derivative to satisfy the boundary conditions at $r = a_0, a_1,$

 0022-2488/80/051256-11$01.00

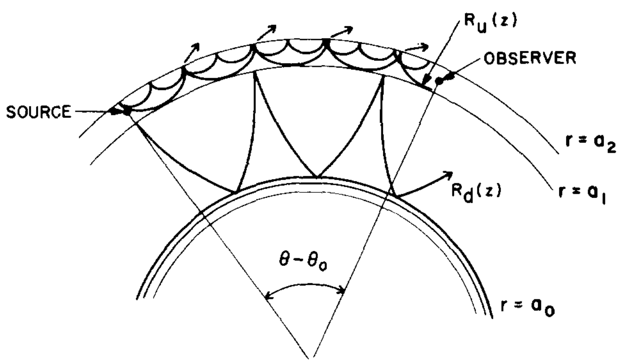

FIG. 1. Boundary conditions on field and its normal derivative yield reflection coefficients R_u and R_d at $r = a_1, a_2$ and $r = a_0, a_1$, respectively.

and a_2 in Fig. 1 yields the following integral for the field[1]

$$\oint_c \frac{\exp(-ixz)F(z)}{f(z)-1}\,dz, \tag{1}$$

where $x \propto (\theta - \theta_0)$, $F(z)$ is the product of the height-gain functions[1] for the source and observer and the contour c encloses the poles which are solution to the modal equation $f(z) = 1$ where

$$f(z) = \frac{w_1(z)}{w_2(z)}\,\frac{w_2(z + D_0x_0)}{w_1(z + D_0x_0)}\,R_u(z)R_d(z). \tag{2}$$

$w_1(z)$ and $w_2(z)$ are two linearly independent solutions to $w'' - zw = 0$ and $R_u(z)$ and $R_d(z)$ are reflection coefficients referring to the level $r = a_1$ as depicted in Fig. 1. They are defined by

$$R_u(z) = \frac{w_1'(z)/w_1(z) - x_1(z)}{-w_2'(z)/w_2(z) + x_1(z)}$$

and

$$R_d(z) = -\frac{w_2'(z + D_0x_0)/w_2(z + D_0x_0) - q}{w_1'(z + D_0x_0)/w_1(z + D_0x_0) - q},$$

where the following parameters appear:

$$x_1(z) = -\frac{D_1}{D_0}\,\frac{v'(R_1z) - \beta(z)w_1'(R_1z)}{v(R_1z) - \beta(z)w_1(R_1z)},$$

$$\begin{aligned}\beta(z) = {} & [D_1v'(R_1z - D_1x_2)W_1(R_2z + x_D) \\ & + D_2v(R_1z - D_1x_2)w_1'(R_2z + x_D)] \\ & \times [D_1w_1'(R_1z - D_1x_2)w_1(R_2z + x_D) \\ & + D_2w_1(R_1z - D_1x_2)w_1'(R_2z + x_D)]^{-1},\end{aligned} \tag{3}$$

$$\begin{aligned}
D_0 &= (1 - a_1S_1)^{1/3},\\
D_1 &= (a_1S_2 - 1)^{1/3},\\
D_2 &= (1 - a_1S_3)^{1/3},\\
R_1 &= D_0^2/D_1^2,\\
R_2 &= D_0^2/D_2^2,\\
x_2 &= k_1(a_1 - a_2)/(k_1a_1/2)^{1/3},\\
x_0 &= k_1(a_1 - a_0)/(k_1a_1/2)^{1/2},\\
x_D &= \Delta nk_1a_1/(k_1a_1/2)^{1/3}D_2^2 + D_2x_2,\\
q &= -i(k_1a_1/2)^{1/3}\Delta/D_0,\\
\Delta &= (\eta - 1)^{1/2}/\eta \quad \text{(for vertical polarization)},\\
&= (\eta - 1)^{1/2} \quad \text{(for horizontal polarization)},\\
\eta &= e_r - i\sigma/\omega\epsilon_0,
\end{aligned}$$

and $w_1(z) = \sqrt{\pi}[Bi(z) - iAi(z)]$, $w_2(z) = \sqrt{\pi}[Bi(z) + iAi(z)]$, $v(z) = \sqrt{\pi}Ai(z)$ and where $\exp(i\omega t)$ time dependence is assumed, σ is the ground conductivity in S/m, and ϵ_r is the dielectric constant. The trilinear model for the refractive index and parameters defined in Eq. (3) is shown in Fig. 2. In Eq. (3) k_1 is the value of the wave number at the lower boundary of the duct in Fig. 2; i.e., $k_1 = k(a_1)$ and Δn is the refractive index contrast shown in Fig. 2.

We now study some of the general features of the modal equation, given by Eq. (2). Typical values for σ and ϵ_r at UHF/VHF frequencies; e.g., at 200 MHz, are $\sigma = 0.005$ S/m and $\epsilon_r = 10$, and we find $q \simeq 1.57 - i78.48$ for vertical polarization and $q \simeq 10^5$ for horizontal polarization.

Now we may observe that for the above values of q

$$\lim_{z\to 0} R_d(z) = \lim_{|z|\to 0} -\frac{-i|z|^{1/2} + i|q|}{i|z|^{1/2} + i|q|} = \exp(i\pi)$$

and then the modal equation becomes [here $|q| \gg |z + X_0D_0|^{1/2}$ so $R_d(z) = \exp(i\pi)$]

$$\frac{w_1(z)}{w_2(z)}\,\frac{w_2(z + D_0x_0)}{w_1(z + D_0z_0)}\,R_u(z) = -1. \tag{4}$$

Three types of modes are generated from the modal equation (2): planar waveguide (or intralayer) modes, Whispering Gallery or earth detached) modes and creeping wave modes (sometimes referred to as earth diffraction modes). These modes can be identified by their zero locations [i.e., solutions to the modal equation (2)] in the complex z plane shown in Fig. 3. However, often there is no clear distinction between the mode types. A subset of the Whispering Gallery modes is identified in Fig. 3 as trapped and we discuss those shortly.

In Fig. 4, a loci of the roots of the modal equation are shown with $f(z)$ as defined in Eq. (2), as the refractive index contrast, Δn, varies from 0.01 to 30. As Δn increases in Fig.

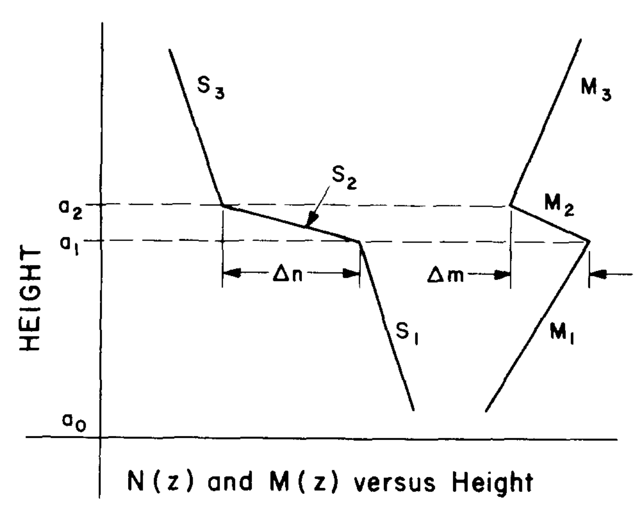

FIG. 2. The trilinear refractive index profile, N and the modified profile M.

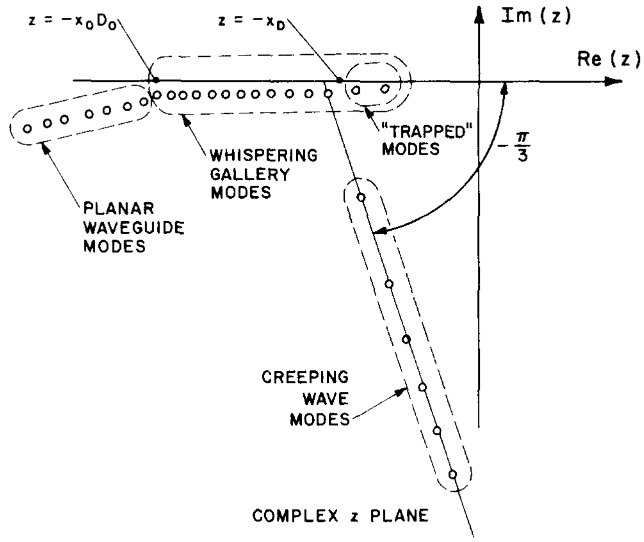

FIG. 3. Roots of modal equation in the complex-z plane showing locations of various types of modes.

4, the Whispering Gallery modes move in from $z \sim -i\infty$ toward the ray $\arg(z) = -\pi$ and the creeping wave modes recede from the ray $\arg(z) = -\pi/3$. In Fig. 4, the planar waveguide modes for $\Delta n = 25$ and 30 were omitted for clarity.

Figure 5 gives a plane wave representation of the modes in Fig. 4 for the case $\Delta n = 30$. The plot to the right of the ray diagram is the modified refractive index versus height defined as[4]

$$m(z) = n(z) + z/a_0. \tag{5}$$

In Fig. 6 we show a ray trace for a refractive index contrast of 30 N units located 1 km above the earth's surface. The caus-

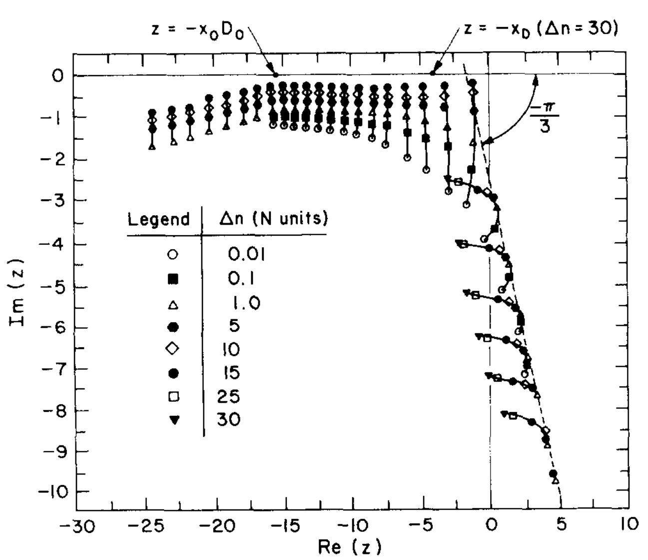

FIG. 4. Loci of $f(z) = 1$ for various values of the refractive index contrast, Δn. Here frequency = 200 MHz, $a_0 = 6368$ km, $a_1 = a_2 = 6369$ km, $\sigma = 0.005$ S/m; $\epsilon_r = 10$, and $s_1 = s_3 = 40$ N units/km. ($s_2 = -\infty$).

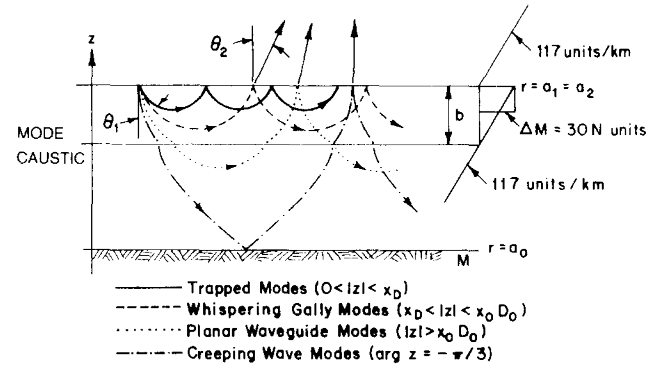

FIG. 5. Ray picture interpretation of an elevated duct together with modified index profile. The mode caustic corresponds to a turning point in the partial differential equation.

tic along the top of the duct is clearly seen. From the familiar relationship among the width of the guide, the angle of inclination of the allowed modes (m) and the wavelength we have, from Fig. 5, for the "track width"

$$b\cos\theta = m\lambda/2. \tag{6}$$

From Snell's law, defining θ_c as the value of θ_1 in Fig. 5 for which $\theta_2 = \pi/2$ we have

$$\sin\theta_c = n_1/n_2 = 1 - \Delta n, \tag{7}$$

and for $\Delta n = 30$ N units we have $\theta_c \approx 89.5562°$. Suppose $f = 200$ MHz, $\alpha_0 = 6378$ km, $a_1 = a_2 = 6379$ km, the slopes of the modified refractive index above and below the abrupt change are 117 N units/km and the refractive index contrast in 30 N units; then $b \approx 264$ m in Fig. 5. Substituting $b = 264$ m, $\lambda = 1.5$ m, and $\theta = 89.5562°$ into Eq. (6) yields $m \approx 2.73$ or an integer value of 2 or 3. This means that two to three modes are trapped in the duct for this example. It turns out that the parameter x_D in Eq. (3) has the value $x_D \approx 4.11$ for this example. This point is plotted in Fig. (4) and we find that two modes lie to the right of the point $\mathrm{Re}(z) = -x_D = -4.111$. Therefore modes with values of m satisfying

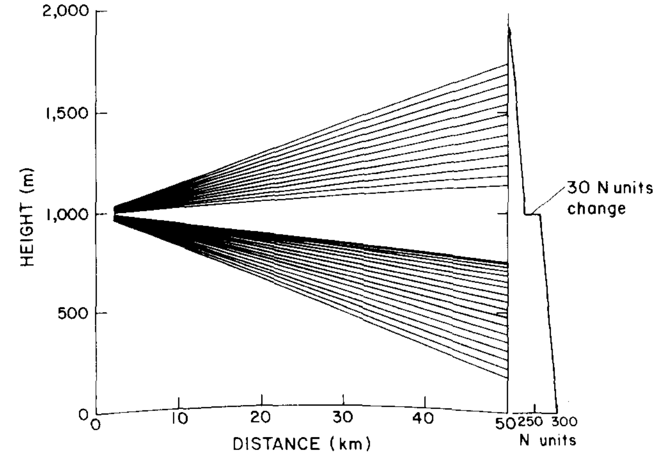

FIG. 6. Ray trace showing trapped modes for an elevated tropospheric duct.

$\mathrm{Re}(z) > -x_D$ will in general be "trapped" in the duct.

The remaining modes in Fig. 4 are interpreted as follows. Those with roots near the ray $\arg(z) \approx -\pi$ and having real parts less than $-x_D$ represent modes "leaking" out of the topside of the duct. From Fig. 4 these modes have much larger attenuation than the trapped or interlayer modes as expected. Those roots near the ray $\arg(z) = -\pi/3$ represent the "creeping" or earth diffracted modes which are present under "standard" atmospheric conditions. The planar waveguide modes have larger attenuation than the Whispering Gallery modes because they shed energy out of the lower boundary of the duct in Fig. 5.

In Fig. 4 as $\Delta n \to 0$, the refractive index profile reduces to the so-called standard or linear atmosphere and the roots to the modal equation lie along the ray $\arg(z) = -\pi/3$.

In order to gain some measure of the relative importance of the zeros along the entire arc, consider the poles in the corresponding reciprocal of the modal equation; i.e., $1/[f(z)-1]_{z=z_0}$. These will be simple poles whose residue is

$$a_{-1} = (z-z_0)\frac{1}{f(z)-1}\bigg|_{z=z_0} = \frac{1}{f'(z_0)}. \tag{8}$$

In Fig. 7 we plot level curves for $|a_{-1}|$ in Eq. (8) for the case $\Delta n = 10$. The residues along $\arg(z) = -\pi/3$ decrease more rapidly than those along $\arg(z) = -\pi$. As an example, consider Table I showing the location of the first zero along $\arg(z) \simeq -\pi/3$ and several of the zeros along $\arg(z) \simeq -\pi$ together with their residues as defined in Eq. (8). Equation (8) ignores the residue of $F(z)$ since it is near unity.

From Table I, we find that out as far as $z \simeq -15 + i0$ the residues of the zeros along $\arg(z) \simeq -\pi$ are slightly larger than the first along $\arg(z) \simeq -\pi/3$; however, the first few zeros along $\arg(z) \simeq -\pi/3$ cannot be neglected. The imaginary part of the root is related to the attenuation of each mode as

$$\left.\begin{array}{l}\text{field}\\ \text{for}\\ m\text{th}\\ \text{root}\end{array}\right\} \propto \exp[-iz_m(k_1a_1/2)^{1/3}(\theta-\theta_0)/D_0^2]/(\theta-\theta_0)^{1/2}, \tag{9}$$

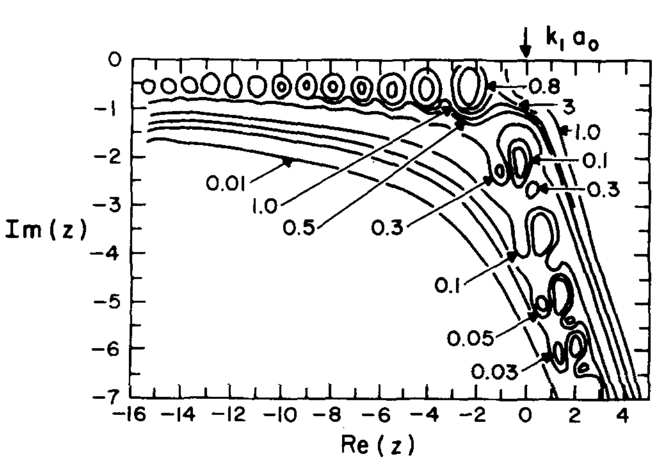

FIG. 7. Plot of the level curves of the residue of $1/f'(z)$ for $\Delta n = 10$.

TABLE I. Comparison of residues for the case $\Delta n = 10$.

z_m	$\lvert a_{-1}\rvert$	Comment
$-0.2 - i2.7$	0.12	(first zero along ray $\arg(z) \simeq -\pi/3$)
$-1.288 - i.378$	0.73	(first zero along ray $\arg(z) \simeq -\pi$)
$-3.2 - i0.5$	0.47	(2nd zero along ray $\arg(z) \simeq -\pi$)
$-4.8 - i0$	0.37	(3rd zero along ray $\arg(z) \simeq -\pi$)
$-6.2 - i0$	0.31	(4th zero along ray $\arg(z) \simeq -\pi$)
...		
$-15 - i0$	0.19	

where $\theta - \theta_0$ is the angular separation of the source and observer, and we find the roots along the ray $\arg(z) = -\pi/3$ decaying in an exponential fashion, for large values of $\mathrm{Im}(z_m)$. The values for z_m in Table I were obtained using the same parameters as used in Fig. 4.

3. ASYMPTOTIC SOLUTION FOR THE MODAL EQUATION FOR THE ROOTS ALONG $\arg(z) = -\pi$

Since those roots along the ray $\arg(z) = -\pi/3$ are usually highly attenuated, we confine attention to the asymptotic solution to the case for those roots along the ray $\arg(z) = -\pi$. There are several cases with subcases, and we present the final asymptotic formulas and their various ranges and requirements for validity in this section and the derivations are given in the Appendix.

In all cases, $|D_1x_2| \ll |z|$

(i) $a_1 \neq a_2$,

$0 < |z| \ll x_0D_0$,

$$-x = \xi \simeq [(3\pi/2)\widehat{m}]^{2/3} + D_0x_2, \tag{10}$$

$$y = \frac{1}{2\xi^{1/2} - D_0x_2\xi^{-1/2}} \times \ln\left(\frac{(D_0^3 + D_1^3)}{8D_0^3\xi^{3/2}}\left|\exp(-i2D_0x_2\xi^{1/2}) - \frac{1}{(1-X_D/R_2\xi)^{3/2}}\right|\right). \tag{11}$$

(ii) $a_1 \neq a_2$,

$$z \simeq -x_0D_0, \tag{12}$$

$$x = \mathrm{Re}(s - iR/2)^2,$$

$$s = \begin{cases} A^{1/3} + B^{1/3}\exp(i2\pi/3), & \Delta < 0, \\ A^{1/3} + B^{1/3}\exp(-i4\pi/4), & \Delta > 0. \end{cases} \tag{13}$$

$$\begin{aligned} A &= -q/2 + (q^2/4 + p^3/27)^{1/2} = -q/2 + \sqrt{\Delta} \\ &= (i/2)[\tfrac{5}{4}R^3 - \tfrac{3}{2}D_0x_2R - \tfrac{3}{2}\pi(m-\tfrac{3}{8})] \\ &\quad + [\tfrac{3}{8}R^6 - \tfrac{15}{16}D_0x_0R^4 - \tfrac{15}{16}\pi R^3(m-\tfrac{3}{8}) \\ &\quad + \tfrac{9}{16}D_0^2x_2^2R^2 + \tfrac{9}{2}D_0x_2R(m-\tfrac{3}{8}) \\ &\quad + \tfrac{9}{4}\pi^2(m-\tfrac{3}{8})^2]^{1/2} \\ &= i(\tfrac{1}{2}[\tfrac{5}{4}R^3 - \tfrac{3}{2}D_0x_2R - \tfrac{3}{2}\pi(m-\tfrac{3}{8})] \\ &\quad \times[-\tfrac{3}{8}R^6 + \tfrac{15}{16}D_0x_2R^4 + \tfrac{15}{16}\pi R^3(m-\tfrac{3}{8}) \\ &\quad - \tfrac{9}{16}D_0^2x_2^2R^2 - \tfrac{9}{2}\pi D_0x_2R(m-\tfrac{3}{8}) \\ &\quad - \tfrac{9}{4}\pi^2(m-\tfrac{3}{8})^2]^{1/2}). \end{aligned} \tag{14}$$

$$\begin{aligned}B = &-q/2 - \sqrt{\Delta}\\ = &i(\tfrac{1}{2}[\tfrac{5}{4}R^3 - \tfrac{3}{2}D_0x_2R - \tfrac{3}{2}\pi(m - \tfrac{3}{8})]\\ &- [-\tfrac{3}{8}R^6 + \tfrac{15}{16}D_0x_2R^4 + \tfrac{15}{16}\pi R^3(m - \tfrac{3}{8})\\ &- \tfrac{9}{16}D_0^2x_2^2R^2 - \tfrac{9}{2}\pi D_0x_2R(m - \tfrac{3}{8})\\ &- \tfrac{9}{4}\pi^2(m - \tfrac{3}{8})^2]^{1/2}). \end{aligned} \tag{15}$$

$$y = \frac{-1}{-2R + D_0x_2/R + \xi/R}\ln|R_u(z)|, \quad x > -x_0D_0, \tag{16}$$

$$y = \frac{-1}{-2R + 2\xi^{1/2} + D_0x_2/R - \xi/R}\ln|R_u(z)|, \quad x < -x_0D_0, \tag{17}$$

where

$$|R_u(z)| = \frac{(D_0^3 + D_1^3)}{8D_0^3\xi^{3/2}}\left|\exp(-i2D_0x_2\xi^{1/2}) - \frac{1}{(1 - x_D/R_2\xi)^{3/2}}\right|.$$

(iii) $a_1 \neq a_2$,

$|z| \gg x_0D_0$,

$$x = \frac{-\pi^2(m - \frac{3}{8})^2}{(R^2 - D_0x_2)^2} - \frac{x_0D_0}{2}, \tag{18}$$

$$y = \frac{\xi^{1/2}}{\frac{2}{3}(R^2 - \frac{3}{2}D_0x_2)}\ln|R_u(z)|. \tag{19}$$

(iv) $a_1 = a_2$,

$|z| \ll x_0D_0$,

$$x \approx \begin{cases}[\tfrac{3}{2}\pi(m - \tfrac{3}{4})]^{2/3}, & x_D < |z| < x_0D_0\\ [\tfrac{3}{2}\pi(m - \tfrac{3}{4})]^{2/3} + 1/x_D, & |z| < x_D,\end{cases} \tag{20}$$

$$y = \frac{1}{2\xi^{1/2}}\ln|R_u(\xi)|, \tag{21}$$

where

$$|R_u(\xi)| = \left|\frac{\xi^{1/2} + (\xi - x_D/R_2)^{1/2}}{\xi^{1/2} - (\xi - x_D/R_2)^{1/2}}\right|.$$

(v) $a_1 = a_2$,

$|z| \simeq x_0D_0$,

$$x = \mathrm{Re}(s - iR/2)^2, \tag{22}$$

$$S = \begin{cases}A^{1/3} + B^{1/3}\exp(i2\pi/3), & \Delta < 0,\\ A^{1/3} + B^{1/3}\exp(-i4\pi/3), & \Delta > 0,\end{cases} \tag{23}$$

$$\Delta = -\tfrac{3}{8}R^6 + \tfrac{15}{16}\pi(m - \tfrac{1}{2}) - \tfrac{9}{4}\pi^2(m - \tfrac{1}{2})^2, \tag{24}$$

$$\begin{aligned}A = &(i/2)[\tfrac{5}{4}R^3 - \tfrac{3}{2}\pi(m - \tfrac{1}{2})]\\ &+ i[-\tfrac{3}{8}R^6 + \tfrac{15}{16}\pi R^3(m - \tfrac{1}{2}) - \tfrac{9}{4}\pi^2(m - \tfrac{1}{2})^2]^{1/2},\end{aligned} \tag{25}$$

$$\begin{aligned}B = &(i/2)[\tfrac{5}{4}R^3 - \tfrac{3}{2}\pi(m - \tfrac{1}{2})]\\ &- i[-\tfrac{3}{8}R^6 + \tfrac{15}{16}\pi R^3(m - \tfrac{1}{2}) - \tfrac{9}{4}\pi^2(m - \tfrac{1}{2})^2]^{1/2},\end{aligned} \tag{26}$$

$$q = i[\tfrac{5}{4}R^3 - \tfrac{3}{2}\pi(m - \tfrac{1}{2})], \tag{27}$$

$$p = (3/4)R^2, \tag{28}$$

$$y = \begin{cases}\dfrac{1}{2R - \xi/R}\ln|R_u(\xi)|, & x > -x_0D_0,\\[2ex] \dfrac{1}{2R - 2\xi^{1/2} + \xi/R}\ln|R_u(\xi)|, & x < -x_0D_0.\end{cases} \tag{29}$$

(vi) $a_1 = a_2$,

$|z| \gg x_0D_0$,

$$x = -\pi^2\frac{(m - \frac{1}{2})^2}{R^4} - \frac{R^2}{2}, \qquad y = \frac{3\xi^{1/2}}{2R^2}\ln|R_u(\xi)|. \tag{30}$$

No difficulties were encountered in obtaining the exact solutions to the model eqıuation using the asymptotic values given in this section as starting values in Newton's method. Section 4 gives two examples.

4. EXAMPLES FOR THE SOLUTIONS OF THE MODAL EQUATION

We seek those values of z satisfying

$$g(z) = f(z) - 1 = 0.$$

Taylor's Theorem gives

$$g(z) \simeq g(z_0) + g'(z_0)\cdot(z - z_0),$$

where z_0 is an initial guess. Even though $g(z)$ is a complex function of a complex variable, $g'(z)$, is analogus to the tangent plane at the point z of the modulus of $g(z)$; i.e.,

$$g(z_0) + (\partial g/\partial z)\epsilon = 0.$$

So

$$\epsilon = -(\partial g/\partial z)^{-1}g(z_0),$$

and

$$z = z_0 - [(\partial g/\partial z)^{-1}g(z_0)].$$

Now

$$[g] = [u\ iv],$$

$$[z] = [x\ iy],$$

so

$$\frac{\partial g}{\partial z} \triangleq \begin{bmatrix}\dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y}\\[2ex] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y}\end{bmatrix},$$

and

$$\left[\left(\frac{\partial g}{\partial z}\right)^{-1}\right] = \begin{pmatrix}\dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x}\\[2ex] \dfrac{\partial v}{\partial y} & \dfrac{\partial u}{\partial y}\end{pmatrix}\Bigg/ |g'|^2,$$

where

$$|g'|^2 = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 v}{\partial x^2}.$$

Also, we have found that using Newton's method with an under relaxation factor reduces the problem of missing a

TABLE II.

Mode	Initial Roots (Asymptotic Values)		Exact Roots	
1	−0.970586	− i0.053274	− 1.44303	− i0.00255008
2	−2.88307	+ i0.137520	−3.049145	− i0.08141006
3	−4.54859	− i0.102077	−4.450727	− i0.242440
4	−5.832742	− i0.298220	−5.757889	− i0.339668
5	−7.002429	− i0.381305	−6.955289	− i0.394370
6	−8.090054	− i0.432289	−8.06518	− i0.431049
7	−9.114720	− i0.470009	−9.10695	− i0.459193
8	−10.08888	− i0.501487	−10.09399	− i0.482932
9	−11.02125	− i0.530113	−11.03564	− i0.504370
10	−11.91817	− i0.557828	−11.9385	− i0.524723
11	−12.7845	− i0.585763	−12.8074	− i0.54479
12	−13.6240	− i0.614034	−13.6455	− i0.565876
13	−14.4397	− i0.640233	−14.4564	− i0.594980
14	−15.2572	− i0.642014	−15.2593	− i0.651949
15	−16.02004	− i0.674579	−16.1002	− i0.743285
16	−16.9616	− i0.813693	−17.01186	− i0.841797
17	−18.09499	− i0.874790	−18.0067	− i0.922549
18	−19.4365	− i0.921807	−19.09532	− i0.982184
19	−20.12839	− i0.862878	−20.28137	− i1.030368
20	−21.46563	− i0.975490	−21.56115	− i1.07594
21	−22.87278	− i1.08212	−22.9294	− i1.123358
22	−24.34986	− i1.186676	−24.38195	− i1.17443
23	−25.89685	− i1.29154	−25.91588	− i1.229904
24	−27.51377	− i1.398376	−27.52915	− i1.290147
25	−29.20061	− i1.50847	−29.2203	− i1.35543
26	−30.95737	− i1.622930	−30.98831	− i1.426050

root of the modal equation. With the under relaxation factor we are essentially increasing the "slope" of the tangent plane.

In Table II, we show the initial choices for the first 26 roots of the modal equation. These were obtained from the asymptotic solutions using the methods in Sec. 3. The roots are compared with the roots of the modal equation obtained using Newtons method. The relaxation factor used was 0.5. The parameters for this example are also given in Fig. 8.

Table III shows the initial choices for the first 27 roots of the modal equation obtained from the asymptotic solutions of Sec. 3 for the case of a duct with zero thickness. The roots are compared with the roots of the exact modal equation obtained using Newton's method, with an under relaxation factor of 0.5. The parameters for this example are given in Fig. 9.

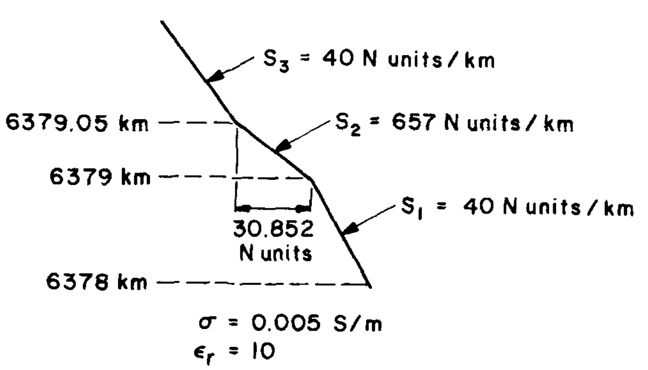

FIG. 8. Refractive index profile and parameters for example of numerical solution of modal equation.

5. CONCLUDING REMARKS

Altitude charts of the *exact* modal equation for a trilinear duct show the existence of roots in the complex-z plane (wave number plane) with $\arg(z) \simeq -\pi$ and *another* with $\arg(z) \simeq -\pi/3$. The set of roots along the ray $\arg(z) \simeq -\pi/3$ are present if the duct thickness and refractive index contrast go to zero. These roots are tracked in the complex-z plane as the refractive index contrast Δn is varied.

The asymptotic solutions of the modal equation are close enough to the exact values that they may be used to obtain first order results for the field strength.

ACKNOWLEDGMENTS

The author acknowledges the significant contribution made by Dr. R.K. Rosich in developing the computer algorithm for plotting the altitude charts for the roots to the modal equation. We also acknowledge Mr. Ray Wasky of the Fire Control Analysis Group at the Avionics Lab of Wright–Patterson Air Force Base for supporting and closely following the analysis presented in this paper. The author also wishes to thank friends, Dr. George Hufford, Dr. J.R. Wait, and Ms. Katherine Warner for many stimulating discussions and suggestions.

APPENDIX: DERIVATION OF ASYMPTOTIC FORMULAS FOR MODAL EQUATION FOR INITIAL CHOICES IN NEWTONS METHOD

The cases follow:

(i) $a_1 \neq a_2$,

$|z| \ll x_0 D_0$.

Then, after a great deal of algebra, one can show from

TABLE III.

Mode	Initial Roots (Asymptotic Values)		Exact Roots	
1	−1.351972	− i0.0	−1.78353	− i0.0221027
2	−3.49814	− $i1\times10^{-12}$	−3.43243	− i0.067132
3	−4.82632	− i0.167417	−4.86113	− i0.192396
4	−6.167129	− i0.255254	−6.17718	− i0.258259
5	−7.374853	− i0.287547	−7.37916	− i0.288324
6	−8.490507	− i0.303459	−8.492867	− i0.303756
7	−9.53705	− i0.312187	−9.53856	− i0.312329
8	−10.52898	− i0.317178	−10.53006	− i0.317256
9	−11.476163	− i0.3200175	−11.477008	− i0.3200653
10	−12.38514	− i0.321534	−12.38644	− i0.321579
11	−13.26305	− i0.322196	−13.26379	− i0.322504
12	−14.11223	− i0.3222897	−14.11523	− i0.325462
13	−14.93659	− i0.321994	−14.95858	− i0.339864
14	−15.98099	− i0.320709	−15.8378	− i0.375504
15	−17.04457	− i0.424976	−16.7984	− i0.426206
16	−18.08915	− i0.482316	−17.85648	− i0.48157357
17	−19.11588	− i0.5284115	−19.0107	− i0.5378257
18	−20.12590	− i0.5698519	−20.256913	− i0.594516
19	−21.19259	− i0.617207	−21.591406	− i0.651856
20	−22.657253	− i0.716872	−23.011605	− i0.7100549
21	−24.199006	− i0.815103	−24.5156	− i0.769248
22	−25.817846	− i0.912831	−26.102105	− i0.8295117
23	−27.513773	− i0.010607	−27.7700	− i0.8908809
24	−29.286789	− i0.1087758	−29.5185	− i0.983365
25	−31.13689	− i0.207558	−31.347116	− i1.01696
26	−33.06408	− i0.307089	−33.25524	− i1.081661
27	−35.068361	− i1.40747	−35.2425	− i1.147438

Eq. (3) from $\mathrm{Re}(z)\ll 0$, and $|D_1x_2|\ll|z|$ that

$$1+2i\beta(z)\simeq\frac{1}{8}\frac{(D_1^3+D_0^3)}{D_0^3}\frac{1}{(z+x_D/R_2)^{3/2}}\times\frac{w_2(R_1z-D_1x_2)}{w_1(R_1z-D_1x_2)} \tag{A1}$$

and from Eq. (3)

$$x_1(z)\sim-\sqrt{z}\left(1-\frac{D_1^3}{4D_0^3z^{3/2}}+\frac{(D_1^3+D_2^3)}{4D_0^3(z+x_D/R_2)^{3/2}}\times\exp\left[\tfrac{4}{3}(R_1z)^{3/2}+\tfrac{4}{3}(R_1z-D_1x_2)^{3/2}\right]\right). \tag{A2}$$

Then $R_u(z)$

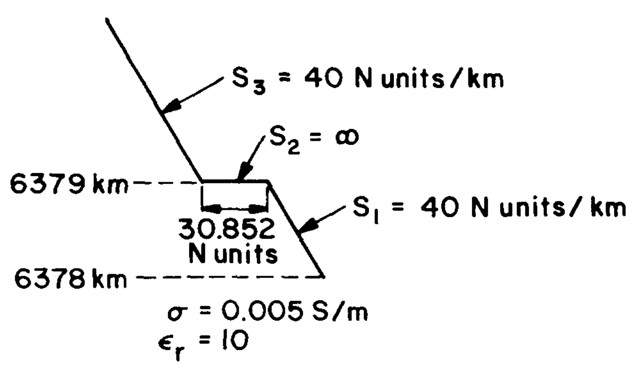

FIG. 9. Refractive index profile and parameters for numerical example.

$$\sim\frac{(D_1^3+D_0^3)}{8D_0^3z^{3/2}}\left(1-\frac{(D_1^3+D_2^3)}{(D_1^3+D_0^3)}\frac{1}{(1+x_D/R_2z)^{3/2}}\times\exp\{-\tfrac{4}{3}[(R_1z)^{3/2}-(R_1z-D_1x_2)^{3/2}]\}\right).$$

Since $D_1x_2\simeq0$ we have

$$R_u(z)\simeq\frac{(D_0^3+D_1^3)}{8D_0^3z^{3/2}}\times\left(1-\frac{(D_1^3+D_2^3)}{(D_0^3+D_1^3)}\frac{\exp(-2\,D_0x_2z^{1/2})}{(1+x_D/2R_2)^{3/2}}\right). \tag{A3}$$

It can be shown that

$$\lim_{z\to0}\arg[R_u(z)]=-\pi,$$

and

$$\lim_{z\to x_0D_0}\arg[R_u(z)+2D_0x_2z^{1/2}]=-\tfrac{3}{4}\pi.$$

From Eq. (4) $\exp(u+iv)=-1$, where $\exp(u+iv)=w_1(z)R_u(z)/w_2(z)$ and

$$R_u(z)=\exp(i\arg R_u(z)+\ln|R_u(z)|). \tag{A4}$$

We now approximate $\arg R_u(z)$ as

$$\arg R_u(z)\simeq\begin{cases}-2D_0x_2z^{1/2}-\pi, & 0<|z|<x_D,\\ -2D_0x_2z^{1/2}-3\pi/4, & x_D<|z|<x_0D_0.\end{cases} \tag{A5}$$

In (A4) above we have used $w_2(z+x_0D_0)/w_1(z+x_0D_0)\simeq-1$ for $|z|\ll x_0D_0$.

Let $z=[\exp(-i\pi)]\xi$, with $\xi>0$ and real, then for the modulus of $R_u(z)$

$$\ln|R_u(z)| = \ln\left(\frac{(D_0^3+D_1^3)}{8D_0^3\xi^{3/2}}\left|\exp(-i2D_0x_2\xi^{1/2}) - \frac{1}{(1-x_D/R_2\xi)^{3/2}}\right|\right). \quad (A6)$$

Now we will use

$$z^{3/2}=(x+iy)^{3/2}\simeq x^{3/2}+\tfrac{3}{2}iyx^{1/2}, \quad y>0$$

and, since $x=[\exp(-i\pi)]\xi$, $x^{3/2}=i\xi^{3/2}$, $x^{1/2}=-i\xi^{1/2}$,

$$z^{3/2}\simeq i\xi^{3/2}+\tfrac{3}{2}y\xi^{1/2}.$$

Also

$$z^{1/2}\simeq -i\xi^{1/2}-(y/2)\xi^{1/2}.$$

Using

$$\frac{w_1(z)}{w_2(z)}\sim\exp(-\tfrac{4}{3}z^{3/2}+i\pi/2), \quad -\pi<arg(z)<\pi/3, \quad (A7)$$

we find from $\exp(u+iv)=1$, that

$$\exp(iv)=\exp\left(-i\tfrac{4}{3}\xi^{3/2}+i\pi/2+i\begin{Bmatrix}-\pi+2D_0x_2\xi^{1/2}\\ -\tfrac{3}{4}\pi+2D_0x_2\xi^{1/2}\end{Bmatrix} \pm i2\pi(m-1)\right)=1, \quad m=1,2\cdots \quad (A8)$$

and from the real part of $u+iv=0$

$$u=-2y\xi^{1/2}+D_0x_2y\xi^{1/2}+\ln|R_u(z)|=0. \quad (A9)$$

From (A8), for large ξ,

$$-i\tfrac{4}{3}\xi^{3/2}\pm i2\pi(m-1)=0$$

and the upper sign is needed. The modal equation becomes

$$-i\tfrac{4}{3}\xi^{3/2}+i2D_0x_2\xi^{1/2}+i2\pi\hat{m}=0$$

or

$$\xi^{3/2}-\tfrac{3}{2}D_0x_2\xi^{1/2}-\tfrac{3}{2}\pi\hat{m}=0, \quad (A10)$$

where

$$\hat{m}=\begin{cases}m-\tfrac{1}{2}, & 0<|z|<x_D,\\ m-\tfrac{3}{8}, & x_D<|z|<x_0D_0.\end{cases}$$

Let $z=\xi^{1/2}$, then (A10) becomes

$$x^3-\tfrac{3}{2}D_0x_2z-\tfrac{3}{2}\pi\hat{m}=0,$$

a cubic equation with in general three roots. Since ξ is real

$$\xi=z^2=\{[\tfrac{3}{4}\pi\hat{m}+(\tfrac{9}{16}\pi^2\hat{m}^2-D_0^3x_2^3/8)^{1/2}]^{1/3}+[\tfrac{3}{4}\pi\hat{m}-(\tfrac{9}{16}\pi^2\hat{m}^2-D_0^3x_2^3/8)^{1/2}]^{1/3}\}^2. \quad (A11)$$

Since $9\pi^2\hat{m}^2 16\gg D_0^3x_2^3/8$,

$$-x=\xi\simeq[(3\pi/2)\hat{m}]^{2/3}+D_0x_2 \quad (A12)$$

and from (A9)

$$y=\frac{1}{2\xi^{1/2}-D_0x_2\xi^{-1/2}}\times\ln\left(\frac{(D_0^3+D_1^3)}{8D_0^3\xi^{3/2}}\left|\exp(-i2D_0x_2\xi^{1/2})-\frac{1}{(1-x_D/R_2\xi)^{3/2}}\right|\right), \quad (A13)$$

the desired solution for this case.

Now consider

(ii) $a_1\neq a_2$,

$z\simeq -x_0D_0$.

We now push the asymptotics near the point $z\simeq -x_0D_0-i$ to write

$$\frac{w_2(z+D_0x_0)}{w_1(z+D_0x_0)}\sim\exp[\tfrac{4}{3}(z+D_0x_0)^{3/2}-i\pi/2], \quad -\pi<\arg(z)<\pi/3. \quad (A14)$$

We can partially justify this step by comparing the asymptotics in this case with the previous case and also with the exact results and we find the approximation to be better than expected.

The reflection coefficient changes phase by $\pi/2$ rad at the caustic $z=-x_0D_0$ [i.e., $\exp(2D_0x_2z^{1/2})$ changes argument for $z\gtrless -D_0x_0$] so, in place of (A5) we have

$$\arg R_u(z)\simeq -\mathrm{Im}(2D_0x_0z^{1/2})+\pi/4. \quad (A15)$$

Substituting (A15), (A14), and (A7) into

$$\exp(u+iv)=\frac{w_1(z)}{w_2(z)}\frac{w_2(z+D_0x_0)}{w_1(z+D_0x_0)}R_u(z), \quad (A16)$$

where $|R_u(z)|$ is given in (A6) and

$$\exp(iv)=\exp[-\tfrac{4}{3}z^{3/2}+i\pi/2+\tfrac{4}{3}(z+D_0x_0)^{3/2}-i\pi/2-2D_0x_0z^{1/2}+i\pi/4+i2\pi(m-1)+i\pi] = -1, \quad m=1,2,\cdots,$$

or, equivalently,

$$-\tfrac{4}{3}z^{3/2}+\tfrac{4}{3}(z+D_0x_0)^{3/2}-2D_0x_2z^{1/2} = -i2\pi(m-\tfrac{3}{8}), \quad m=1,2,\cdots. \quad (A17)$$

Let,

$$z=\Delta z-x_0D_0=x+iy-x_0D_0, \quad |\Delta z|\ll x_0D_0,$$

$$(z+D_0x_0)^{3/2}=\Delta z^{3/2}, \quad \arg(\Delta z)\simeq -\pi \text{ or } 0,$$

$$z^{3/2}=(\Delta z-x_0D_0)^{3/2}=(-x_0D_0)^{3/2}(1-\Delta z/x_0D_0)^{3/2}\simeq i(x_0D_0)^{3/2}-i\tfrac{3}{2}(x_0D_0)^{1/2}\Delta z+i\tfrac{3}{8}\Delta z^2/(x_0D_0)^{1/2},$$

where $(-x_0D_0)^{3/2}=i(x_0D_0)^{3/2}$, and $(-x_0D_0)^{1/2}=-i(x_0D_0)^{1/2}$. Let $R=(x_0D_0)^{1/2}$, then

$$z^{3/2}=iR^3-i\tfrac{3}{2}R\Delta z+i\tfrac{3}{8}\Delta z^2/R,$$

$$z^{1/2}=(-x_0D_0)^{1/2}(1-\Delta z/x_0D_0)^{1/2}\simeq -iR+(i/2)\Delta z/R,$$

and (A17) becomes

$$\Delta z^{3/2}+\tfrac{3}{2}iR\Delta z-iR^3+(3i/2)D_0x_0R+i\tfrac{3}{2}\pi(m-\tfrac{3}{8})=0. \quad (A18)$$

Let $\omega=\Delta z^{1/2}$, then (A18) becomes

$$\omega^3+\tfrac{3}{2}iR\omega^2-iR^3+\tfrac{3}{2}iD_0x_2R+i\tfrac{3}{2}\pi(m-\tfrac{3}{8})=0. \quad (A19)$$

Let $\omega=s-iR/2$, then (A19) becomes

$$s^3+\tfrac{3}{4}R^2s-i\tfrac{5}{4}R^3+i\tfrac{3}{2}D_0x_2R+i\tfrac{3}{2}\pi(m-\tfrac{3}{8})=0. \quad (A20)$$

Then define

$$q=-i[\tfrac{5}{4}R^3-\tfrac{3}{2}D_0x_2R-\tfrac{3}{2}\pi(m-\tfrac{3}{8})], \quad (A21)$$

$$p=\tfrac{3}{4}R^2,$$

$$
\begin{aligned}
A &= -q/2 + (q^2/4 + p^3/27)^{1/2} = -q/2 + \Delta^{1/2} \\
&= (i/2)[\tfrac{5}{4}R^3 - \tfrac{3}{2}D_0x_2R - \tfrac{3}{2}\pi(m-\tfrac{3}{8})] \\
&\quad + [\tfrac{3}{8}R^6 - \tfrac{15}{16}D_0x_0R^4 - \tfrac{15}{16}\pi R^3(m-\tfrac{3}{8}) \\
&\quad + \tfrac{9}{16}D_0^2x_2^2R^2 + \tfrac{9}{2}\pi D_0x_2R(m-\tfrac{3}{8}) \\
&\quad + \tfrac{9}{4}\pi^2(m-\tfrac{3}{8})^2]^{1/2} \qquad \text{(A22)} \\
&= i\{\tfrac{1}{2}[\tfrac{5}{4}R^3 - \tfrac{3}{2}D_0x_2R - \tfrac{3}{2}\pi(m-\tfrac{3}{8})] \\
&\quad + [-\tfrac{3}{8}R^6 + \tfrac{15}{16}D_0x_2R^4 + \tfrac{15}{16}\pi R^3(m-\tfrac{3}{8}) \\
&\quad - \tfrac{9}{16}D_0^2x_2^2R^2 - \tfrac{9}{2}\pi D_0x_2R(m-\tfrac{3}{8}) \\
&\quad - \tfrac{9}{4}\pi^2(m-\tfrac{3}{8})^2]^{1/2}\}. \qquad \text{(A23)}
\end{aligned}
$$

$$
\begin{aligned}
B &= -q/2 - \sqrt{\Delta} \\
&= i\{\tfrac{1}{2}[\tfrac{5}{4}R^3 - \tfrac{3}{2}D_0x_2R - \tfrac{3}{2}\pi(m-\tfrac{3}{8})] \\
&\quad - [-\tfrac{3}{8}R^6 + \tfrac{15}{16}D_0x_2R^4 + \tfrac{15}{16}\pi R^3(m-\tfrac{3}{8}) \\
&\quad - \tfrac{9}{16}D_0^2x_2^2R^2 - \tfrac{9}{2}\pi D_0x_2R(m-\tfrac{3}{8}) \\
&\quad - \tfrac{9}{4}\pi^2(m-\tfrac{3}{8})^2]^{1/2}\}. \qquad \text{(A24)}
\end{aligned}
$$

The root we need is

$$
s = \begin{cases} A^{1/3} + B^{1/3}\exp(i2\pi/3), & \Delta < 0, \\ A^{1/3} + B^{1/3}\exp(-i4\pi/3), & \Delta > 0, \end{cases} \qquad \text{(A25)}
$$

to ensure the cross product terms in

$$
[\exp(i\pi/6)(\sigma/2 + i\sqrt{\Delta}) + \exp(i5\pi/6)(\sigma/2 - i\sqrt{\Delta})]^2,
$$
$$
\sigma = \tfrac{5}{4}R^3 - \tfrac{3}{2}D_0x_2R - \tfrac{3}{2}\pi(m-\tfrac{3}{8})
$$

and

$$
[\exp(i\pi/6)(\sigma/2 + \sqrt{\Delta}) + \exp(-i\pi/6)(\sigma/2 - \sqrt{\Delta})]^2
$$

equal $-p/3$ and the sum of the squared terms equal $-q$. Then our solution is

$$
x = \mathrm{Re}(s - iR/2)^2. \qquad \text{(A26)}
$$

We take as transition from case (i) and (ii) when $\xi = R^2 = x_0D_0$ ($z = -x_0D_0$). Then, from (A10),

$$
m_1 \triangleq \tfrac{3}{8} + (2/3\pi)(R^3 - \tfrac{3}{2}D_0x_2R), \qquad \text{(A27)}
$$

and the transition from case (ii) to case (iii) when $\Delta z = -x_0D_0/2$ or $z = -3x_0D_0/2$. Then, from (A17)

$$
m_2 \triangleq \tfrac{3}{8} + (1/2\pi)[R^3(\sqrt{6} - 2/3) - D_0x_2R]. \qquad \text{(A28)}
$$

From (A16) $u = 0$ gives the imaginary part of the root for this case. With $\xi > 0$, $z = x + iy - x_0D_0$ and

$$
x = \begin{cases} \exp(-i\pi\xi), & x < -x_0D_0, \\ \xi, & x > -x_0D_0, \end{cases}
$$
$$
x^{3/2} = \begin{cases} i\xi^{3/2}, & x < -x_0D_0, \\ \xi^{3/2}, & x > -x_0D_0, \end{cases}
$$
$$
x^{1/2} = \begin{cases} -i\xi^{1/2}, & x < -x_0D_0, \\ \xi^{1/2}, & x > -x_0D_0, \end{cases}
$$
$$
\mathrm{Im}(\Delta z)^2 = 2xy,
$$
$$
\mathrm{Im}(\Delta z^{3/2}) \simeq 0,
$$
$$
\begin{aligned}
u = &-\frac{4}{3}\left(\frac{3}{2}Ry\right) - \frac{4}{3}\left(-\frac{3}{8}\frac{2y}{R}\right)\begin{Bmatrix} -\xi \\ \xi \end{Bmatrix} \\
&+ \frac{4}{3}\begin{Bmatrix} \tfrac{3}{2}y\xi^{1/2} \\ 0 \end{Bmatrix} - 2D_0x_2(-y/2R) \\
&+ \ln|R_u(z)| = 0, \quad \begin{cases} x < -x_0D_0 \\ x > -x_0D_0 \end{cases}.
\end{aligned}
$$

So

$$
y = \begin{cases} \dfrac{-1}{-2R + D_0x_2/R + \xi/R}\ln|R_u(z)|, \\ \qquad\qquad x > -x_0D_0, \\ \dfrac{-1}{-2R + 2\xi^{1/2} + D_0x_2/R - \xi/R}\ln|R_u(z)|, \\ \qquad\qquad x < -x_0D_0, \end{cases} \qquad \text{(A29)}
$$

(iii) $a_1 \neq a_2$,

$|z| \gg x_0D_0$.

Here $z = \Delta z - x_0D_0$ with $|\Delta z| > x_0D_0$

$$
z^{3/2} \simeq (\Delta z)^{3/2} - (3x_0D_0/2)(\Delta z)^{1/2} + \tfrac{3}{8}(x_0D_0)^2/(\Delta z)^{1/2},
$$
$$
z^{1/2} \simeq (\Delta z)^{1/2} - x_0D_0/2(\Delta z)^{1/2}.
$$

From (A16), the real part of the root satisfies

$$
\begin{aligned}
iv = &-\tfrac{4}{3}(\Delta z)^{3/2} + 2R^2(\Delta z)^{1/2} - R^2/2(\Delta z)^{1/2} + \tfrac{4}{3}(\Delta z)^{3/2} \\
&- 2D_0x_2(\Delta z)^{1/2} + R^2D_0x_0/(\Delta z)^{1/2} + i2\pi(m - \tfrac{3}{8}) \\
&= 0,
\end{aligned}
$$

or

$$
\Delta z + \frac{i2\pi((m-\tfrac{3}{8})(\Delta z)^{1/2}}{2(R^2 - D_0x_2)} - \frac{R^2(\tfrac{1}{2}R^2 - D_0x_0)}{2(R^2 - D_0x_2)} = 0,
$$

and, after some algebra,

$$
x = \frac{-\pi^2(m-\tfrac{3}{8})^2}{(R^2 - D_0x_2)^2} - \frac{x_0D_0}{2}. \qquad \text{(A30)}
$$

From (A16), for $u = 0$ we have

$$
\begin{aligned}
&(\Delta z)^{3/2} \simeq i\xi^{3/2} + \tfrac{3}{2}y\xi^{1/2}, \quad \xi > 0, \\
&z^{3/2} \simeq i\xi^{3/2} + \tfrac{3}{2}y\xi^{1/2} + iR\xi^{1/2} + R^2y/2\xi^{1/2}, \\
&z^{1/2} \simeq -i\xi^{1/2} - y/2\xi^{1/2} - iR/2\xi^{1/2}, \\
&u = -\tfrac{4}{3}(\tfrac{3}{2}y\xi^{1/2} + R^2y/2\xi^{1/2}) + \tfrac{4}{3}(\tfrac{3}{2}y\xi^{1/2}) \\
&\qquad + D_0x_2y/\xi^{1/2} + \ln|R_u(z)| = 0,
\end{aligned}
$$

or

$$
y = \frac{\xi^{1/2}}{\tfrac{2}{3}(R^2 - \tfrac{3}{2}D_0x_2)}\ln|R_u(z)|. \qquad \text{(A31)}
$$

(iv) $a_1 \neq a_2$,

$|z| \ll x_0D_0$.

In this case

$$
x_1(z) = \frac{D_2}{D_0}\frac{w_1'(R_2z + x_D)}{w_1(R_2z + x_D)}, \qquad \text{(A32)}
$$

and

$$
R_u(z) \simeq \frac{z^{1/2} - (z + x_D/R_2)^{1/2}}{z^{1/2} + (z + x_D/R_2)^{1/2}}. \qquad \text{(A33)}
$$

Now, for this case we have

$$
\arg[R_u(z)] = \begin{cases} 0, & x_D < |z| \ll x_0D_0, \\ 2\tan^{-1}(z^{1/2}x_D), & |z| < x_D, \end{cases} \qquad \text{(A34)}
$$

which we approximate by

$$
\arg\{R_u(z)\} = \begin{cases} 0, & x_D < |z| < x_0D_0, \\ 2z^{1/2}/x_D, & |z| < x_D. \end{cases} \qquad \text{(A35)}
$$

Then, from

$$\exp(u+iv)=\frac{w_1(z)}{w_2(z)}R_u(z)=1, \tag{A36}$$

we have

$$iv=-\tfrac{4}{3}z^{3/2}+i\pi/2+\begin{Bmatrix}2z^{1/2}/x_D\\ 0\end{Bmatrix}+i2\pi(m-1)=0,$$

$$m=1,2,\cdots\begin{cases}|z|<x_D\\ x_D<|z|<x_0D_0\end{cases} \tag{A37}$$

and substituting in (A37)

$$z^{3/2}\simeq i\xi^{3/2}+\tfrac{3}{2}y\xi^{1/2},\quad \xi>0,$$

$$z^{1/2}\simeq -i\xi^{1/2},\quad \xi>0,$$

gives

$$-x=\xi\approx\begin{cases}[\tfrac{3}{2}\pi(m-\tfrac{3}{4})]^{2/3}, & x_D<|z|<x_0D_0,\\ [\tfrac{3}{2}\pi(m-\tfrac{3}{4})]^{2/3}+1/x_D, & |z|<x_D.\end{cases} \tag{A38}$$

From (A36), $u=0$, gives

$$u=-\tfrac{4}{3}(\tfrac{3}{2}y\xi^{1/2})+\ln\left|\frac{\xi^{1/2}+(\xi-x_D/R_2)^{1/2}}{\xi^{1/2}-(\xi-x_D/R_2)^{1/2}}\right|=0, \tag{A39}$$

or

$$y=(1/2\xi^{1/2})\ln|R_u(\xi)|,$$

where

$$|R_u(\xi)|=\left|\frac{\xi^{1/2}+(\xi-x_D/R_2)^{1/2}}{\xi^{1/2}-(\xi-x_D/R_2)^{1/2}}\right|. \tag{A40}$$

(v) $a_1=a_2$,

$|z|\simeq x_0D_0$.

We again assume the asymptotics can be pushed into this case as in case (ii). The modal equation is given in (A36) and we find

$$iv=-\tfrac{4}{3}z^{3/2}+(i\pi/2)+\tfrac{4}{3}(z+x_0D_0)^{3/2}-(i\pi/2)+i\pi+i2\pi(m-1)=0, \tag{A41}$$

or

$$-\tfrac{4}{3}z^{3/2}+\tfrac{4}{3}(z+x_0D_0)^{3/2}+i2\pi(m-\tfrac{1}{2})=0.$$

Let $z^{3/2}\simeq iR^3-i\tfrac{3}{2}R\Delta z$, then

$$-\tfrac{4}{3}(iR^3-i\tfrac{3}{2}R\Delta z)+\tfrac{4}{3}(\Delta z)^{3/2}+i2\pi(m-\tfrac{1}{2})=0,$$

$$\tfrac{4}{3}(\Delta z)^{3/2}+2iR\Delta z-\tfrac{4}{3}iR^3+i2\pi(m-\tfrac{1}{2})=0,$$

$$(\Delta z)^{3/2}+\tfrac{3}{2}iR\Delta z-iR^3+i\tfrac{3}{2}\pi(m-\tfrac{1}{2})=0.$$

Let $\omega=(\Delta z)^{1/2}$, then

$$\omega^3+i\tfrac{3}{2}R\omega^2-iR^3+i\tfrac{3}{2}\pi(m-\tfrac{1}{2})=0.$$

Let $\omega=s-iR/2$, then

$$s^3-\tfrac{3}{4}R^2s-i\tfrac{5}{4}R^3+i\tfrac{3}{2}\pi(m-\tfrac{1}{2})=0.$$

Let

$$q=-i[\tfrac{5}{4}R^3-\tfrac{3}{2}\pi(m-\tfrac{1}{2})],$$

$$p=(3/4)R^2,$$

$$A=(i/2)[\tfrac{5}{4}R^3-\tfrac{3}{2}\pi(m-\tfrac{1}{2})]+i[-\tfrac{3}{8}R^6+\tfrac{15}{16}\pi R^3(m-\tfrac{1}{2})-\tfrac{9}{4}\pi^2(m-\tfrac{1}{2})^2]^{1/2},$$

$$=(i/2)[\tfrac{5}{4}R^3-\tfrac{3}{2}\pi(m-\tfrac{1}{2})]-i[-\tfrac{3}{8}R^6+\tfrac{15}{16}\pi R^3(m-\tfrac{1}{2})-\tfrac{9}{4}\pi^2(m-\tfrac{1}{2})^2]^{1/2},$$

and the solution to (A41) is

$$x=\mathrm{Re}(s-iR/2)^2, \tag{A42}$$

where

$$s=\begin{cases}A^{1/3}+B^{1/3}\exp(i2\pi/3), & \Delta<0,\\ A^{1/3}+B^{1/3}\exp(-i4\pi/3), & \Delta>0,\end{cases} \tag{A43}$$

and

$$\Delta=-\tfrac{3}{8}R^6+\tfrac{15}{16}\pi R^3(m-\tfrac{1}{2})-\tfrac{9}{4}\pi^2(m-\tfrac{1}{2})^2. \tag{A44}$$

We take as transition from case (iv) to (v) when $\xi=x_0D_0$ or

$$m_1\triangleq\tfrac{3}{8}+(2/3\pi)R^3, \tag{A45}$$

and the transition from case (v) to case (vi) when $z=-3x_0D_0/2$, giving

$$m_2\triangleq(\sqrt{2}R^3/6\pi)(3\sqrt{3}-1)+\tfrac{1}{2}. \tag{A46}$$

From (A36), with $u=0$, we have

$$u=\frac{4}{3}\left(\frac{3}{2}Ry\right)-\frac{4}{3}\left(-\frac{3}{8}\frac{2y}{R}\right)\begin{Bmatrix}-\xi\\ \xi\end{Bmatrix}+\frac{4}{3}\begin{Bmatrix}\tfrac{3}{2}y\xi^{1/2}\\ 0\end{Bmatrix}+\ln|R_u(\xi)|=0,\quad \begin{cases}x<-x_0D_0,\\ x>-x_0D_0,\end{cases}$$

so

$$y=\begin{cases}\dfrac{1}{2R-\xi/R}\ln|R_u(\xi)|, & x>-x_0D_0,\\ \dfrac{1}{2R-2\xi^{1/2}+\xi/R}\ln|R_u(\xi)|, & x<-x_0D_0.\end{cases} \tag{A47}$$

(vi) $a_1=a_2$,

$|z|\gg x_0D_0$.

Here $z=\Delta z-x_0D_0$, $|\Delta z|>x_0D_0$. From (A36), we have

$$iv=-\tfrac{4}{3}z^{3/2}+i\pi/2+\tfrac{4}{3}(z+x_0D_0)^{3/2}-i\pi/2+i2\pi(m-\tfrac{1}{2})=0.$$

Now

$$z^{3/2}\simeq(\Delta z)^{3/2}-\tfrac{3}{2}R^2(\Delta z)^{1/2}+\tfrac{3}{8}R^4/(\Delta z)^{1/2}$$

and

$$iv=-\tfrac{4}{3}(\Delta z)^{3/2}+2R^2(\Delta z)^{1/2}-\tfrac{1}{2}R^4/(\Delta z)^{1/2}+\tfrac{4}{3}(\Delta z)^{3/2}+i2\pi(m-\tfrac{1}{2})=0$$

and

$$\Delta z+\frac{i\pi(m-\tfrac{1}{2})(\Delta z)^{1/2}}{R^2}-\frac{R^2}{4}=0,$$

giving

$$\Delta z=-\pi^2(m-\tfrac{1}{2})^2/R^4+R^2/2.$$

and $\Delta z \simeq x + x_0 D_0$, so

$$x = -\pi^2(m - \tfrac{1}{2})^2/R^4 - R^2/2. \tag{A48}$$

From $u = 0$, we have

$$u = -\tfrac{4}{3}(\tfrac{3}{2} y\xi^{1/2} + R^2 y/2\xi^{1/2}) + \tfrac{4}{3}(\tfrac{3}{2} y\xi^{1/2}) + \ln|R_u(\xi)| = 0,$$

or

$$y = (3\xi^{1/2}/2R^2)\ln|R_u(\xi)|. \tag{A49}$$

[1] S.H. Cho and J.R. Wait, EM Rep. No. 1, Cooperative Institute for Research in Environmental Sciences, Boulder, Colorado, June, 1977. A summarized version is given in S.H. Cho and J.R. Wait, *Analysis of Microwave Ducting in an Inhomogeneous Troposphere, Pure and Applied Geophysics* (1978), Vol. 116, pp. 1118–42.
[2] C.L. Goodhart and R.A. Pappert, Tech. Rep. No. 153, NOSC, Sept. 1977.
[3] K.G. Budden, Philos. Mag. **42**, 833 (1952).
[4] J.R. Wait, *Electromagnetic Waves in Stratified Media*, 2nd ed. (Pergamon, New York, 1970).
[5] J.R. Wait and K. Spies, Radio Sci. **4** (4), 319 (1969).
[6] Lord Rayleigh, Philos. Mag. **20**, 1001 (1910).

ERRATA

Erratum: Four Euclidean conformal group in atomic calculations: Exact analytical expressions for the bound–bound two photon transition matrix elements in the H atom [J. Math. Phys. 19, 1041 (1978)]

J. P. Gazeau
Laboratoire de Chimie Physique de l'Université Pierre et Marie Curie, 75 231 Paris Cedex 05, France

(Received 13 December 1979; accepted for publication 4 January 1980)

P. 1042: Eq. (1.9)′ should read:

$$(\phi_1, \phi_2)_{\mathscr{H}(p_0)} = \left(\psi_1, \left(\frac{p_0^2 + p^2}{2p_0^2}\right)\psi_2\right)_{L^2_{\mathbb{C}}(\mathbb{R}^3)}$$

P. 1044: In Eq. (3.3), in place of "$\cdots K_{1\pm} = (p_0 \pm p_{n'}, \mathbf{k}_1)\cdots$," read: "$\cdots K_{1\pm} = (p_0 \pm p_n, \mathbf{k}_1)$.[11]

P. 1048: In Eq. (D2), in place of:

"$\cdots T^m_{nl,n'l'}\cdots,\ U^m_{nl,n'l'},\ U^{-m}_{nl,n'l'}\cdots,$"

read: "$\cdots T^m_{n'l',nl}\cdots,\ U^m_{n'l',nl}\cdots,\ U^{-m}_{n'l',nl}\cdots.$"

In the Eq. (D3) in place of "$T^m_{nl,n'l'} = \cdots,$" read: "$T^m_{n'l',nl} = \cdots.$"

Finally, the Eq. (D4) should read:

$$\begin{aligned} U^m_{n'l',nl} = {} & \delta_{n',n+1}(\delta_{l',l-1}\, a(n+1, -l)c(l,m) \\ & + \delta_{l',l+1}\, a(n+1, l+1)c(l+1, 1-m)) \\ & - \delta_{n',n-1}(\delta_{l',l-1}\, a(n,l)c(l,m) \\ & + \delta_{l',l+1}\, a(n, -l-1)c(l+1, 1-m)). \end{aligned}$$

Erratum: Conserved densities for nonlinear evolution equations. I. Even order case [J. Math. Phys. 20, 1239 (1979)]

L. Abellanas, A. Galindo
Department of Theoretical Physics, Universidad Complutense de Madrid, Madrid-3, Spain

(Received 31 December 1979; accepted for publication 4 January 1980)

A pitiful error has been detected by the authors in Sec. 3. Equation (11) is wrong and thus the criteria (2) and (3) are incorrect. As to example $u_t = (u_4 + u_5)u_6$, which was exhibited as an illustration of criterion (2), a direct calculation shows that it has no conserved density. The remaining criteria (1), (4), (5), and (6) hold.

Finally, several minor corrections are in order: In Eq. (16b) $\exists\, \rho(\cdots u_{M/2-2})$ should read $\exists \rho \in C_{M/2-2}(P)$. In criterion (6), $\nexists \rho(\cdots u_{M/2-2})$ and $\nexists \rho$ should read, respectively, $\nexists \rho \in C_{M/2-2}(P)$ and $\nexists \rho \in C(P)$. In Eq. (18), the first symbol ρ should be replaced by f. In Eq. (24), $T_{u_{M-2}}$ must read $K(\cdots u_{M-2})$. In Proposition 3(ii), $C_{M/2-1}(P)$ should read $C_{M/2-2}(P)$. In the first formula following Proposition 4, replace $\delta\rho/\delta u$ by $(-1)^n(\delta\rho/\delta u)$. In Eq. (38), a term $b_1(u)u_1$ should be added to its right-hand side, with $b_1(u)$ an arbitrary polynomial.

 0022-2488/80/051267-01$0.00