Commutation factors on generalized Lie algebras

R. Kleeman
Department of Mathematical Physics, University of Adelaide, GPO Box 498, Adelaide, South Australia 5001, Australia

(Received 21 February 1985; accepted for publication 31 May 1985)

Generalized Lie algebras or color algebras, as we shall call them, are described by an Abelian grading group Γ and a commutation factor ϵ defined on Γ. In this paper Γ is assumed to be finite. It is shown that color algebras with the pair $\langle \Gamma,\epsilon\rangle$ can also be considered as color algebras with the different pair $\langle \Gamma',\epsilon'\rangle$ and that as a result a canonical pair $\langle \Gamma_c,\epsilon_c\rangle$ is possible. It is further shown that, in fact, a unique "minimal" $\langle \Gamma_c,\epsilon_c\rangle$ can be used for all algebras with the pair $\langle \Gamma,\epsilon\rangle$.

I. INTRODUCTION

Color algebras are a generalization of superalgebras and were introduced, from a physical point of view, by Rittenberg and Wyler[1] in 1978. Part of the motivation was to throw some light on generalized methods of field quantization.[2] In 1979 Scheunert[3] studied the algebras from a rigorous mathematical viewpoint and succeeded in proving generalizations of the Poincaré–Birkhoff–Witt and Ado theorems. Perhaps the most important result he obtained was that there is a unique "canonical" superalgebra for every color algebra. Moreover this correspondence carries over to the representation theory and a simple generalization of the Klein transformation[4] allows one to obtain all representations of color algebras from the representations of the canonical superalgebra (see Kleeman[5] for details).

A color algebra **A** is firstly a *graded* algebra. That means one assigns elements of the *Abelian* grading group Γ to each element of **A** and then supposes that the algebra satisfies

$$a_\alpha \circ b_\beta = c_{\alpha+\beta}, \quad \alpha,\beta \in \Gamma. \tag{1.1}$$

One further supposes that there is a commutation factor $\epsilon:\Gamma\times\Gamma\to C$, which satisfies

$$\begin{aligned} &\epsilon(\alpha,\beta)\epsilon(\beta,\alpha) = 1,\\ &\epsilon(\alpha+\beta,\gamma) = \epsilon(\alpha,\gamma)\epsilon(\beta,\gamma),\\ &\epsilon(\alpha,\beta+\gamma) = \epsilon(\alpha,\beta)\epsilon(\alpha,\gamma). \end{aligned} \tag{1.2}$$

The definition of the color algebra is completed by supposing that its elements further satisfy

$$a_\alpha \circ b_\beta = -\epsilon(\alpha,\beta) b_\beta \circ a_\alpha, \tag{1.3}$$

$$\sum_{\mathrm{cycl}(\alpha,\beta,\gamma)} \epsilon(\gamma,\alpha) a_\alpha \circ (b_\beta \circ c_\gamma) = 0. \tag{1.4}$$

The last equation being a generalization of the Jacobi identity of Lie algebras.

In order to introduce the notion of a *representation* of a color algebra one considers graded vector spaces

$$V = \bigoplus_{\alpha\in\Gamma} V_\alpha \tag{1.5}$$

and graded linear maps upon such spaces, i.e.,

$$g_\alpha(V_\beta) \subseteq V_{\alpha+\beta}. \tag{1.6}$$

The color algebra $\mathrm{gl}(V,\epsilon)$ is then the set of such maps with a color algebra product defined by

$$g_a \circ g_b \equiv g_\alpha\, g_\beta - \epsilon(\alpha,\beta) g_\beta\, g_\alpha. \tag{1.7}$$

Usage of (1.2) confirms that this product satisfies (1.3) and (1.4). A *representation* is then a homomorphic map of a color algebra **A** into the color algebra $\mathrm{gl}(V,\epsilon)$.

It is fairly clear from the above discussion that a color algebra **A** has two essential elements: namely its grading group Γ and its commutation factor ϵ. The basic philosophy of this paper is contained in the observation that for a given color algebra it may also be possible to consider it as a color algebra with a different grading group Γ' and commutation factor ϵ'. If we add the condition that the representations of **A** with $\langle \Gamma,\epsilon\rangle$ are also representations with $\langle \Gamma',\epsilon'\rangle$ then there seems no reason to prefer $\langle \Gamma,\epsilon\rangle$ over $\langle \Gamma',\epsilon'\rangle$. The essential result of this paper is that when Γ is finite there is a *canonical* set of $\langle \Gamma_c,\epsilon_c\rangle$, which will serve as grading groups and commutation factors for *all* color algebras. The ϵ_c are *almost* determined by the Γ_c—hence the expression canonical. We shall further show that for all color algebras with $\langle \Gamma,\epsilon\rangle$ there is a *unique* $\langle \Gamma_c,\epsilon_c\rangle$ able to replace $\langle \Gamma,\epsilon\rangle$. The uniqueness will be in the sense that any other $\langle \Gamma_c',\epsilon_c'\rangle$ able to replace $\langle \Gamma,\epsilon\rangle$ will satisfy $\Gamma_c' \subseteq \Gamma_c$.

It should be noticed that the scenario described above has already been carried through by Scheunert in the case that Γ is a vector space as well. The vector space property enabled him to use certain standard results on canonical forms on vector spaces. As we shall see below the extension to the more general case requires a somewhat more basic approach.

To clarify the above abstract discussion it may be helpful to discuss a concrete example.

To this end we consider a color algebra which is graded by the group $Z_4\oplus Z_4\oplus Z_2\oplus Z_2$. Let us assume further that this color algebra has the commutation factor

$$\epsilon(\alpha,\beta) = (-1)^{\phi(\alpha,\beta)}, \tag{1.8}$$

where $\phi(\alpha,\beta)$ is a bilinear form given by

$$\phi(\alpha,\beta) = \boldsymbol{\alpha}^t M \boldsymbol{\beta}. \tag{1.9}$$

The "vectors" $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the decomposition of α and β with respect to the direct sum decomposition of the grading group. In other words $\boldsymbol{\alpha}$ has the form

$$\boldsymbol{\alpha} = (\alpha_1,\alpha_2,\alpha_3,\alpha_4),$$

with $\alpha_1 \in Z_4$, $\alpha_2 \in Z_4$, $\alpha_3 \in Z_2$, and $\alpha_4 \in Z_2$.

The matrix M we shall assume to take the following form:

 0022-2488/85/102405-08$02.50

$$M=\begin{pmatrix}0&1&0&1\\1&0&0&1\\0&0&0&1\\1&1&1&0\end{pmatrix}. \tag{1.10}$$

Now as a consequence of the results in Secs. III and IV it shall be possible to grade our color algebras by the new grading group $Z_2\oplus Z_2\oplus Z_2\oplus Z_2$ and to have a new commutation factor defined on the algebra. This factor will have the same basic form as (1.8) and (1.9) except that M will be altered to the simpler M' given by

$$M'=\begin{pmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}. \tag{1.11}$$

II. COMMUTATION FACTORS

We can immediately deduce from (1.2) a number of basic results:

$$\begin{aligned}&\epsilon(\alpha,\alpha)=\pm 1,\\&\epsilon(\alpha,0)=\epsilon(0,\alpha)=1,\\&\epsilon(\alpha,n\beta)=\epsilon(n\alpha,\beta)=\epsilon^n(\alpha,\beta).\end{aligned} \tag{2.1}$$

These will be used repeatedly below.

Consider now a finite Abelian grading group Γ. The very well known and old result[6] concerning such groups is that they have a unique, up to isomorphism, decomposition into p-groups:

$$\Gamma=\Gamma_{p_1}\oplus\Gamma_{p_2}\oplus\cdots\oplus\Gamma_{p_n}, \tag{2.2}$$

with the p_i being distinct primes. These p-groups have a further unique decomposition into cyclic groups of prime power order,

$$\Gamma_{p_i}=Z_{(p_i)^{r_1}}\oplus Z_{(p_i)^{r_2}}\oplus\cdots\oplus Z_{(p_i)^{r_m}}, \tag{2.3}$$

and we assume for convenience that $r_1\geqslant r_2\geqslant\cdots\geqslant r_m$. Now let us denote the generator of $Z_{(p_i)^{r_j}}$ by q_i^j and define

$$E_{ij}^{st}\equiv\epsilon(q_i^s,q_j^t). \tag{2.4}$$

It follows from (2.2) and (2.3) that an arbitrary $\alpha\in\Gamma$ can be written as

$$\alpha=\sum_{i,s}k_i^s q_i^s, \tag{2.5}$$

where the k_i^s are integers, and so, by using (2.1),

$$\epsilon(\alpha,\beta)=\prod_{i,s}(E_{ii}^{ss})^{k_i^s l_i^s}\prod_{i>j}\prod_{s>t}(E_{ij}^{st})^{k_i^s l_j^t-k_j^t l_i^s}, \tag{2.6}$$

with

$$\beta=\sum_{i,s}l_i^s q_i^s;$$

and the first equation of (2.1) shows that $E_{ii}^{ss}=\pm 1$.

One can obtain further information on the E_{ij}^{st} by using (2.1) to deduce that

$$(E_{ij}^{st})^{(p_i)^{r_s}}=(E_{ij}^{st})^{(p_j)^{r_t}}=1, \tag{2.7}$$

and so if $i\neq j$ then $E_{ij}^{st}=1$, whereas if $i=j$ then E_{ii}^{st} is a νth root of unity with

$$\nu=\min((p_i)^{r_s},(p_i)^{r_t}). \tag{2.8}$$

This allows us to decompose ϵ into a product of commutation factors on the individual p-groups of (2.2):

$$\epsilon(\alpha,\beta)=\epsilon_1(\alpha_1,\beta_1)\epsilon_2(\alpha_2,\beta_2)\cdots\epsilon_n(\alpha_n,\beta_n), \tag{2.9}$$

where the α_i and β_i are the projections of α and β onto Γ_{p_i}. Finally we may write the ϵ_i, with the aid of (2.6), as

$$\epsilon_i(\alpha_i,\beta_i)=\eta_{i1}^{\psi(\alpha_i,\beta_i)} \tag{2.10}$$

with

$$\psi(\alpha_i,\beta_i)=\mathbf{k}_i^t\mathbf{M}_i\mathbf{l}_i,$$

where $\mathbf{k}_i$ and $\mathbf{l}_i$ are vectors with elements $\{k_i^s\}$ and $\{l_i^s\}$, respectively; η_{i1} is the primitive $(p_i)^{r_1}$-th root of unity and $\mathbf{M}_i$ is an antisymmetric matrix of integers modulo $(p_i)^{r_1}$.

It is to be noted that if we had considered finitely generated Abelian grading groups instead, then one would have added copies of Z (the integers) to (2.2). In this case, however, one does not get the decomposition of ϵ given by (2.9) and, in fact, there may be "cross terms" between the Z groups and the p-groups. It is for this reason that we defer discussion of this more general situation to a later paper.

III. COVERING

We discuss in this section when a grading group Γ and commutation factor ϵ may be replaced. Consider the class $C_{\Gamma\epsilon}$ of color algebras with commutation factor ϵ and grading group Γ. We say that $\langle\Gamma',\epsilon'\rangle$ *covers* $\langle\Gamma,\epsilon\rangle$ if every member of $C_{\Gamma\epsilon}$ is *also* a member of $C_{\Gamma'\epsilon'}$ *and* any representation of an algebra in $C_{\Gamma\epsilon}$ is also a representation of the algebra when considered as a member of $C_{\Gamma'\epsilon'}$. Notice that this relation is not necessarily symmetric and in fact is a partial ordering—inherited from the class containment relation.

In order to use this relation we need a more technical definition of a color algebra than that given in Sec. I. Suppose we have an algebra A, whose elements we denote by a^i (i belonging to some set Ω), then we may define this algebra through its structure constants C_k^{ij}. Or, in other words, the product on A is defined through the equation

$$a^i\circ a^j=C_k^{ij}a^k, \tag{3.1}$$

where summation over Ω is implied by the repeated index. We now say that A is a *color* algebra with Abelian grading group Γ and commutation factor ϵ, or more briefly A is *colored* by $\langle\Gamma,\epsilon\rangle$, if the following hold.

(a) There exists a map $\phi:\Omega\to\Gamma$ such that whenever $C_k^{ij}\neq 0$ then $\phi(i)+\phi(j)=\phi(k)$.

(b) ϵ is a commutation factor in the sense of (1.2).

(c) $C_k^{ij}=-\epsilon(\phi(i),\phi(j))C_k^{ji}$, $\forall i,j,k\in\Gamma$.

(d) $$\sum_{\text{cycl}(i,j,k)}\epsilon(\phi(k),\phi(i))C_m^{il}C_l^{jk}=0,$$
$$\forall i,j,k,m\in\Omega.$$

It is clear from this definition that A will also be a color algebra with grading group Γ' and commutation factor ϵ' if there exists a map $\phi':\Omega\to\Gamma'$ satisfying condition (a), and if

$$\epsilon(\phi(i),\phi(j)) = \epsilon'(\phi'(i),\phi'(j)), \quad \forall i,j \in \Omega. \tag{3.2}$$

In addition (3.2) also implies [upon consideration of (1.7)] that any representation of A with $\langle \Gamma,\epsilon\rangle$ will also be a representation with $\langle \Gamma',\epsilon'\rangle$.

A very general situation where $\langle \Gamma',\epsilon'\rangle$ covers $\langle \Gamma,\epsilon\rangle$ is when there exists what we shall term a *covering homomorphism* between Γ and Γ'. We define this as follows: $h:\Gamma\to\Gamma'$ is a covering homomorphism if (a) it is a homomorphism; and (b) ϵ and ϵ' satisfy the relation

$$\epsilon(\alpha,\beta) = \epsilon'(h(\alpha),h(\beta)), \quad \forall \alpha,\beta \in \Gamma. \tag{3.3}$$

To show that $\langle \Gamma',\epsilon'\rangle$ covers $\langle \Gamma,\epsilon\rangle$ we observe that given any algebra A with a coloring $\langle \Gamma,\epsilon\rangle$ we can obtain a coloring by $\langle \Gamma',\epsilon'\rangle$ with the new grading map ϕ' given by

$$\phi' = h\circ\phi,$$

which satisfies condition (a) of the color algebra definition because h is a homomorphism. Equation (3.2) now follows directly from (3.3).

If we suppose that h is onto and satisfies the relation

$$\alpha \in \ker(h) \Rightarrow \epsilon(\alpha,\beta) = 1, \quad \forall \beta \in \Gamma, \tag{3.4}$$

then ϵ' will be *induced* from ϵ via (3.3). This is so because if h is onto then ϵ' will be defined from (3.3). This definition will make sense since if there is a $\gamma \neq \alpha$ such that $h(\alpha) = h(\gamma)$ then $\alpha - \gamma \in \ker(h)$ and so $\epsilon(\alpha-\gamma,\beta) = 1$ or $\epsilon(\alpha,\beta) = \epsilon(\gamma,\beta)$. A similar argument holds for the second argument of ϵ. Finally it is easy to establish that ϵ' will be a commutation factor on Γ'. The first equation of (1.2) follows from (3.3) and the fact that ϵ is a commutation factor. The other two follow also this way with the additional use of the homomorphic property of h.

Notice that if h were an isomorphism, i.e., 1:1 *as well* as onto, condition (3.4) is fulfilled trivially because $\ker(h) = 0$ and so (3.4) follows from (2.1).

It should be observed at this point that Scheunert[3] has considered what he terms *equivalence* of commutation factors. Thus two commutation factors ϵ and ϵ', defined on the *same* Γ are termed equivalent if there exists an *automorphism* $g:\Gamma\to\Gamma$ such that

$$\epsilon'(\alpha,\beta) = \epsilon(g(\alpha),g(\beta)). \tag{3.5}$$

It is clear that in this case we can conclude that g is a covering homomorphism, as is g^{-1}. Thus in our terminology $\langle \Gamma,\epsilon\rangle$ covers $\langle \Gamma,\epsilon'\rangle$ and vice versa.

It is an interesting question as to whether a covering homomorphism between Γ and Γ' is necessarily implied when $\langle \Gamma',\epsilon'\rangle$ covers $\langle \Gamma,\epsilon\rangle$. We shall provide a partial answer to this question in Theorem 3.2 below. Before this result is proved we need a preliminary definition and lemma.

We say that $\langle \Gamma,\epsilon\rangle$ is *reduced* if

$$\epsilon(\alpha,\beta) = 1, \quad \forall \beta \in \Gamma,$$

implies $\alpha = 0$.

Lemma 3.1: There always exists an *onto* covering homomorphism between Γ and a Γ^r, where $\langle \Gamma^r,\epsilon^r\rangle$ is reduced.

Proof: We define Γ_0, the ϵ-trivial subgroup of Γ, as follows:

$$\Gamma_0 \equiv \{\alpha \in \Gamma : \epsilon(\alpha,\gamma) = 1, \ \forall \gamma \in \Gamma\}. \tag{3.6}$$

To see that it is a subgroup of Γ suppose $\alpha,\beta \in \Gamma_0$, then

$$\begin{aligned}\epsilon(\alpha-\beta,\gamma) &= \epsilon(\alpha,\gamma)\epsilon(-\beta,\gamma)\\ &= \epsilon(\alpha,\gamma)\epsilon^{-1}(\beta,\gamma) = 1, \quad \forall \gamma \in \Gamma.\end{aligned}$$

We identify the Γ^r with Γ/Γ_0 and choose the homomorphism $g \in \Gamma\to\Gamma/\Gamma_0$ to be the *natural homomorphism* (Fuchs[7]), which is onto and has kernel Γ_0, and thus by (3.4) is a covering homomorphism. As we have seen this means there is an induced ϵ^r. Finally $\langle \Gamma^r,\epsilon^r\rangle$ is reduced since suppose $g(\alpha) \in \Gamma^r$ is an arbitrary element of Γ^r, then

$$\begin{aligned}&\epsilon^r(g(\alpha),\gamma^r) = 1, \quad \forall \gamma^r \in \Gamma^r\\ &\Rightarrow \epsilon^r(g(\alpha),g(\gamma)) = 1, \quad \forall \gamma \in \Gamma\\ &\Rightarrow \epsilon(\alpha,\gamma) = 1, \quad \forall \gamma \in \Gamma\\ &\Rightarrow \alpha \in \Gamma_0\\ &\Rightarrow g(\alpha) = 0.\end{aligned}$$

□

Theorem 3.2: Suppose $\langle \Gamma',\epsilon'\rangle$ covers $\langle \Gamma,\epsilon\rangle$ then there exists a subgroup $\Gamma^h \subseteq \Gamma'$ such that there is a covering homomorphism between Γ and the reduced Γ^h/Γ_0^h (Γ_0^h being the ϵ'-trivial subgroup of Γ^h).

Proof: We begin by stating the following.

Proposition 3.3: There exists a well-defined mapping $h:\Gamma\to\Gamma'$ satisfying

$$\epsilon(\alpha,\beta) = \epsilon'(h(\alpha),h(\beta)).$$

The proof of this proposition depends on the covering condition that representations of algebras in $C_{\Gamma\epsilon}$ are also representations of the same algebra when considered as a member of $C_{\Gamma'\epsilon'}$. We do not present the proof here as it would require some development of color algebra representation theory. It may be found in Kleeman.[5]

We define Γ^h to be the subgroup of Γ generated by $h(\Gamma)$. The covering homomorphism we require is just the composition of the map h and the onto covering homomorphism $k:\Gamma^h \to \Gamma^h/\Gamma_0^h$ given by Lemma 3.1. To see this, first we observe that

$$\epsilon(\alpha,\beta) = \epsilon'(h(\alpha),h(\beta)) = \epsilon^r(kh(\alpha),kh(\beta)),$$

using Lemma 3.1 and Proposition 3.3, and where ϵ^r is the induced commutation factor on Γ^h/Γ_0^h. It remains to be shown that kh is, in fact, a homomorphism. Define

$$l(\alpha,\beta) \equiv kh(\alpha+\beta) - kh(\alpha) - kh(\beta),$$

and so

$$\epsilon^r(l(\alpha,\beta),kh(\gamma)) = \epsilon(0,\gamma) = 1, \quad \gamma \in \Gamma.$$

Now Γ^h/Γ_0^h consists of elements of the form

$$k\left(\sum_i n^i h(\alpha_i)\right) = \sum_i n^i kh(\alpha_i), \quad \alpha_i \in \Gamma,$$

but

$$\epsilon^r\left(l(\alpha,\beta),\sum_i n^i kh(\alpha_i)\right) = \prod_i \left[\epsilon^r(l(\alpha,\beta),kh(\alpha_i))\right]^{n_i} = 1,$$

but since Γ^h/Γ_0^h is reduced this means that $l(\alpha,\beta) = 0$, which in turn means that kh must be homomorphic. □

IV. CANONICAL FORMS

The obvious question now arises: Does there exist a canonical set of $\langle \Gamma_c,\epsilon_c\rangle$'s that cover all possible $\langle \Gamma,\epsilon\rangle$'s? We

provide a complete answer to this question in the case that Γ is finite.

Theorem 4.1: Every color algebra A, which can be colored by $\langle \Gamma, \epsilon \rangle$, where Γ is a finite group, can also be colored by a $\langle \Gamma_c, \epsilon_c \rangle$. The Γ_c are of the form

$$\Gamma_c = \Gamma_{p_1} \oplus \Gamma_{p_1} \oplus \cdots \oplus \Gamma_{p_n}, \tag{4.1}$$

where the p_i are distinct primes and each p_i-group Γ_{p_i}, with $p_i \neq 2$, is of the form

$$\Gamma_{p_i} = Z^1_{(p_i)^{r_1}} \oplus Z^2_{(p_i)^{r_1}} \oplus \cdots \oplus Z^{2j}_{(p_i)^{r_1}} \oplus Z^1_{(p_i)^{r_2}} \oplus \cdots \oplus Z^{2k}_{(p_i)^{r_2}} \oplus \cdots \oplus Z^{2l}_{(p_i)^{r_m}}, \tag{4.2}$$

where $Z^u_{(p_i)^{r_k}}$ means the uth copy of the cyclic group of order $(p_i)^{r_k}$. For $p_i = 2$ the group has the same form *except* that an odd number of copies of Z_2 are allowed. The ϵ_c defined on the Γ_c have the decomposition given by (2.9), i.e.,

$$\epsilon_c = \epsilon_1 \epsilon_2 \cdots \epsilon_n,$$

with ϵ_i defined on the Γ_{p_i}. The ϵ_i will be *unique* in the case $p_i \neq 2$, and are given by

$$\epsilon_i(\alpha, \beta) = \eta_{i1}^{\psi_1(\alpha_1, \beta_1)} \eta_{i2}^{\psi_2(\alpha_2, \beta_2)} \cdots \eta_{im}^{\psi_m(\alpha_m, \beta_m)}, \tag{4.3}$$

with η_{ik} being the primitive $(p_i)^{r_k}$ root of unity: α_k and β_k the projections of α and β onto the copies of $Z_{(p_i)^{r_k}}$ in Γ_{p_i}; and, for example, $\psi_1(\alpha_1, \beta_1)$ is the following *antisymmetric* bilinear form defined on copies of $Z_{(p_i)^{r_1}}$:

$$\psi_1(\alpha_1, \beta_1) = \sum_{i=1}^{j} \left[k_1^{2i-1} l_1^{2i} - k_1^{2i} l_1^{2i-1} \right]. \tag{4.4}$$

The k and l are as in (2.5). In the case of $p_i = 2$, ϵ_i has the form

$$\epsilon_i(\alpha, \beta) = \eta_{i1}^{\psi_1(\alpha_1, \beta_1)} \cdots \eta_{im}^{\psi_m(\alpha_m, \beta_m)} (-1)^{\phi(\alpha', \beta')}, \tag{4.5}$$

where the η_{ik}, ψ_k, α_k, and β_k are the same as before with the restriction that $r_k \neq 1$. The ϕ is defined on the copies of Z_2 (α' and β' being the projections of α and β onto these copies) and has *two* possible forms (already discussed by Scheunert[3]). The first is

$$\phi(\alpha', \beta') = \sum_{i=1}^{q} k_i' l_i', \tag{4.6a}$$

with q being the number of copies of Z_2 in Γ_2. The second is the antisymmetric form

$$\phi(\alpha', \beta') = \sum_{i=1}^{q/2} \left[k'_{2i-1} l'_{2i} - k'_{2i} l'_{2i-1} \right]. \tag{4.6b}$$

Notice that in this case q must be even.

Proof: We shall show that a $\langle \Gamma_c, \epsilon_c \rangle$ covers an arbitrary $\langle \Gamma, \epsilon \rangle$ by considering a sequence of covering homomorphisms between Γ and Γ_c. It shall be sufficient to restrict the covering homomorphisms to a particular p_i-group with commutation factor obtained from the decomposition (2.9). It is obvious that these restricted covering homomorphisms extend to the whole group—just set the action on the other p_i-groups to the identity.

Let us write the p_i-group of Γ as follows:

$$Z^1_{(p_i)^{r_1}} \oplus Z^2_{(p_i)^{r_1}} \oplus \cdots \oplus Z^{n_1}_{(p_i)^{r_2}} \oplus Z^1_{(p_i)^{r_2}} \oplus \cdots \oplus Z^{n_2}_{(p_i)^{r_2}} \oplus \cdots \oplus Z^{n_m}_{(p_i)^{r_m}}, \tag{4.7}$$

where $r_1 > r_2 > \cdots > r_m$, with the notation as in (4.2). With respect to this basis of the p-group[8] the matrix $\mathbf{M}$ of (2.10), which determines the commutation factor on the p-group, can be written as

$$M = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1m} \\ M_{21} & M_{22} & \cdots & M_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mm} \end{pmatrix}, \tag{4.8}$$

where the M_{jk} are the submatrices of dimension $n_j \times n_k$. Using (2.8) we deduce that these submatrices have the form

$$M_{jk} = p^{r_i - \min(r_j, r_k)} R_{jk}, \tag{4.9}$$

where R_{jk} is an arbitrary $n_j \times n_k$ matrix of integers modulo p^{r_1}. In the same manner as with $\mathbf{M}$ we can break up the $\mathbf{k}$ and $\mathbf{l}$ of (2.10) into subvectors $\mathbf{k}_j$ and $\mathbf{l}_j$, where j runs from 1 to m, as in the notation of (4.7).

According to Shoda[9] the automorphisms on the p-group have the expression

$$\mathbf{k}_i' = \sum_{j=1}^{m} P_{ij} \mathbf{k}_j \tag{4.10}$$

through the $\mathbf{k}$, where the P_{ij} have the following form: For $i > j$ the entries of the matrix P_{ij} are integers modulo p^{r_i}; while for $i \leqslant j$ $P_{ij} = p^{r_i - r_j} Q_{ij}$ with Q_{ij} having the form of P_{ij} when $i > j$; finally one requires that $\det(P_{ij})$ not be a multiple of p. This final condition is to ensure that the homomorphism is 1:1 and onto.

Now given a commutation factor ϵ on our p-group the automorphism of (4.10) induces a new commutation factor ϵ' via (3.5), which is given by

$$\epsilon'(\alpha, \beta) = \epsilon(\alpha', \beta') = \eta_{i1}^{\psi(\alpha', \beta')}$$

and

$$\psi(\alpha', \beta') = (\mathbf{Pk})^t \mathbf{M} (\mathbf{Pl}) = \mathbf{k}^t (\mathbf{P}^t \mathbf{M} \mathbf{P}) \mathbf{l}, \tag{4.11}$$

or in other words the $\mathbf{M}$ is transformed to $\mathbf{M}'$ given by

$$\mathbf{M}' = \mathbf{P}^t \mathbf{M} \mathbf{P}. \tag{4.12}$$

As is usual in reduction problems of this kind we shall be interested in particular types of $\mathbf{P}$, which correspond to column and row operations on $\mathbf{M}$. From the form of (4.12) it is clear that a given row operation must always be followed by the *corresponding* column operation. The conditions on $\mathbf{P}$ outlined above mean that there are restrictions on the row (*and* corresponding column) operations allowed. These are easily seen to be the following.

(1) For the addition of a multiple of a row to another row, which we denote by $sr_1 + r_2 = r_2'$: If r_1 belongs to the ith block row $(M_i)_j = M_{ij}$ and r_2 to the kth block row $(M_k)_j = M_{kj}$ then if (a) $i \geqslant k$, then s may be an arbitrary number modulo p^{r_1}; and (b) $i < k$, then s must be a multiple of $p^{r_i - r_k}$.

(2) The multiplication of a row by a constant s has the restriction that s may not be divisible by p. This follows from $\det(P_{ii})$ not being a multiple of p.

(3) The interchange of two rows is only possible when they belong to the same block row.

Apart from the above automorphisms we shall also be interested in the following nonautomorphic covering homo-

morphism: Suppose the mth row in the first block row $(M_1)_j = M_{1j}$ is a multiple of p [this occurs when the mth row of M_{11} is a multiple of p—due to (4.10)], then there is a covering homomorphism which maps the $Z^m_{(p)^{r_1}}$ summand of (4.7) into a $Z_{(p)^{r_1-1}}$ summand and leaves all other summands untouched. This map ϕ is defined as follows: We can write any element of $Z^m_{(p)^{r_1}}$ uniquely as

$$kp^{r_1-1}+l, \tag{4.13}$$

with $k<p$ and l not divisible by p^{r_1-1}. Clearly l corresponds to an element of $Z_{(p)^{r_1-1}}$ and then ϕ is simply given by

$$\phi(kp^{r_1-1}+l)=l. \tag{4.14}$$

To show that ϕ is a covering homomorphism, we first observe that it is obviously onto by its definition. Second, the kernel of ϕ just consists of elements of the form kp^{r_1-1} (which is just a Z_p subgroup of $Z^m_{(p)^{r_1}}$) from $Z^m_{(p)^{r_1}}$ and zeros from all the other summands in (4.7). Given now that the mth row of the first block is a multiple of p it is clear that if $\alpha \in \ker\phi$ then, *a la* (2.10), $\mathbf{k}^t\mathbf{M}$ is a vector, which is a multiple of p^{r_1} and so $\epsilon(\alpha,\beta)=1, \forall\beta\in\Gamma$, using the first equation of (2.10).

Using the above covering homomorphisms we can proceed to reduce M_{11}. For $p\neq 2$ it is clear from (2.6) that

$$M_{11}=\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n_1} \\ -a_{12} & 0 & \cdots & a_{2n_1} \\ -a_{13} & -a_{23} & \cdots & a_{3n_1} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n_1} & -a_{2n_1} & \cdots & 0 \end{pmatrix}. \tag{4.15}$$

Consider now the first column; two possibilities arise—either it is a multiple of p or else there exists an a_{1j} not a multiple of p. In the first case we apply the covering homomorphism ϕ of (4.14) to the summand $Z^1_{(p)^{r_1}}$ converting it to a $Z_{(p)^{r_1-1}}$ summand and then relegate the column (and row) to being in the second block column (or row). Note that it may or may not be a *new* block column (or row) depending on whether $r_2=r_1-1$. We then restart the analysis with a smaller $(n_1-1)\times(n_1-1)$ M_{11}. In the second case we multiply the column by the inverse of a_{1j} (which exists and is not a multiple of p because a_{1j} is not a multiple of p) and then interchange the second and jth row. Now M_{11} has been reduced to the following form:

$$M'_{11}=\begin{pmatrix} 0 & 1 & a'_{13} & \cdots & a'_{1n_1} \\ -1 & 0 & a'_{23} & \cdots & a'_{2n_1} \\ -a'_{13} & -a'_{23} & 0 & \cdots & a'_{3n_1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a'_{1n_1} & -a'_{2n_1} & -a'_{3n_1} & \cdots & 0 \end{pmatrix}. \tag{4.16}$$

We now eliminate all other elements in the first column (and row) by multiplying the second row by a'_{1j} and subtracting it from the jth row. Now we can multiply the first row by a'_{2j} and *add* it to the jth row—thereby eliminating all but the 1 from the second column. Now M_{11} becomes

$$M'_{11}=\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & b_{34} & \cdots & b_{3n_1} \\ 0 & 0 & -b_{34} & 0 & \cdots & b_{4n_1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -b_{3n_1} & -b_{4n_1} & \cdots & 0 \end{pmatrix}.$$

It is obvious that the above analysis can now be applied to the third column and so on and so we conclude that M_{11} may be reduced to the form

$$M^r_{11}=\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

which is of dimension $n'_1\times n'_1$ with $n'_1\leqslant n_1$.

Now we can use a row operation of type (1) (b) to eliminate all elements in M^r_{j1} $(j>1)$. This is because elements in M_{j1} (and thus M^r_{j1}) have the form $kp^{r_1-r_j}$ by (4.9) and by (1) (b) we are allowed to multiply a row from M^r_{11} by a number of this form and add it to a row in M^r_{j1}. *It is to be noted that such an elimination would not have been possible in general, if we had not allowed the nonautomorphic covering homomorphism* (consider the extreme example of when M_{11} consists entirely of zeros).

We have now reduced $\mathbf{M}$ to $\mathbf{M}^r$ with

$$M^r=\begin{pmatrix} M^r_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \bar{\mathbf{M}} & \\ 0 & & & \end{pmatrix}. \tag{4.17}$$

Clearly now we can regard $\bar{\mathbf{M}}$ as determining a commutation factor on a group with cyclic summands of order strictly less than p^{r_1}. We may now repeat the analysis of above on this $\bar{\mathbf{M}}$ *without affecting* the decomposition in (4.17). The only complication with continuing the analysis iteratively is at the end where there may be rows (and columns) of zeros left. It is fairly obvious that the cyclic summands corresponding to these rows (and columns) may be mapped into the trivial group with a covering homomorphism. Equations (4.3) and (4.4) are now clear.

In the case that $p=2$, (2.6) shows that we may get diagonal elements in $\mathbf{M}$ and this will interfere with the reduction process outlined above. The approach we shall follow will be governed by the nature of these diagonal elements. First, if there are no such elements then clearly we may pursue the previous reduction. Second, if there exists an $\alpha\in\Gamma$ such that

$$\epsilon(\alpha,\alpha)=-1$$

and

$$\epsilon(\alpha,\gamma)=\pm 1, \quad \forall\gamma\in\Gamma, \tag{4.18}$$

then we shall show that there exists an onto covering homomorphism between Γ and a Γ_c. Third, if there exist elements satisfying the first equation of (4.18), but none of them satisfy

the second, then we shall demonstrate a *nononto* covering homomorphism between Γ and Γ_c. We proceed now to prove the second case.

First, we observe that if there is a $\beta \in \Gamma$ satisfying (4.18) then β cannot have the form 2γ since in this case

$$-1 = \epsilon(\beta,\beta) = \epsilon(2\gamma,2\gamma) = \epsilon^4(\gamma,\gamma),$$

which means that $\epsilon(\gamma,\gamma) \neq \pm 1$ which contradicts (2.1). We conclude therefore that β must have the form

$$\beta = e_1 + 2^{r_1}e_2 + \cdots + 2^{r_n-1}e_n,$$

where the e_i are the generators of the cyclic summands of Γ. Consider now the covering homomorphism g, given by Lemma 3.1, onto a reduced Γ^r. Consider $g(\beta)$, now

$$\begin{aligned}&\epsilon^r(g(\beta),\gamma^r) \text{ for arbitrary } \gamma^r \in \Gamma^r\\ &= \epsilon^r(g(\beta), g(\gamma)) \text{ for some } \gamma \in \Gamma\\ &= \epsilon(\beta,\gamma) = \pm 1.\end{aligned}$$

Now since Γ^r is reduced and

$$\epsilon^r(2g(\beta),\gamma^r) = 1, \quad \forall \gamma^r \in \Gamma^r,$$

it follows that $g(\beta)$ has order 2 and since

$$\epsilon^r(g(\beta), g(\beta)) = \epsilon(\beta,\beta) = -1$$

we have

$$g(\beta) = e_1^r + 2^{s_1}e_2^r + \cdots + 2^{s_n-1}e_n^r,$$

with

$$o(e_1^r) = o(2^{s_1}e_2^r) = \cdots = o(2^{s_n-1}e_n^r) = 2.$$

From the form of the isomorphisms given by (4.1) it follows that there exists an isomorphism f mapping $g(\beta)$ into e_1^r. Now consider any e_i^r with $o(e_i^r) > 2$ and $\epsilon^r(e_i^r,e_i^r) = -1$, then apply the following isomorphism k to the e_i^r:

$$k(e_i^r) = e_i^r + e_1^r.$$

However,

$$\epsilon^r(k(e_i^r),k(e_i^r)) = \epsilon^r(e_i^r + e_1^r, e_i^r + e_1^r) = -1.1. - 1 = 1,$$

which means that there are no diagonal elements in the reduced $\mathbf{M}^r$ except those in the final Z_2 block. We can now apply the iterative process used in the case of $p \neq 2$ (which consists only of onto covering homomorphisms) until we are left only with a subblock corresponding to the Z_2 summands. From there the proof is completed by use of a theorem of Scheunert.[3]

In the third case mentioned above consider an e_i (notation as in the second case) with least order such that $\epsilon(e_i,e_i) = -1$. Now let the other e_k satisfying $\epsilon(e_k,e_k) = -1$, have the following isomorphism applied to them:

$$e_k' = e_k + e_i.$$

It is fairly clear that after such transformations only e_i will contribute a diagonal element to $\mathbf{M}$. As we shall see later this diagonal element cannot be removed to the Z_2 subblock by means of an *onto* covering homomorphism and instead we put it there by the following nononto covering homomorphism: Let e_i generate a $Z_{2^{r_i}}^{j}$ cyclic summand; then we map $Z_{2^{r_i}}^{j}$ into $Z_2 \oplus Z_{2^{r_i}}$ as follows. It is clear that any element of $Z_{2^{r_i}}^{j}$ can be written as $2k + l$ with $l = 0,1$. The covering homomorphism h is then given by

$$h(2k + l) = (l, 2k + l) \tag{4.19}$$

and the new commutation factor on the expanded Γ' will have a new $\mathbf{M}'$, which will be the same as $\mathbf{M}$ except that all diagonal entries, not belonging to the Z_2 subblock, will be zero. In addition there will be a new Z_2 summand in Γ' whose effect on $\mathbf{M}$ will be to introduce a single diagonal element 2^{r_i-1} and all new off-diagonal elements will be zeros.

It is obvious that (4.19) describes a homomorphic map and a little thought then shows that the new commutation factor we have defined will satisfy the condition (3.3). We can now repeat the comments that applied for the final reduction in the second case and obtain the stated result.

V. UNIQUENESS RESULTS

Another important question to be considered is the uniqueness or otherwise of the canonical $\langle \Gamma_c,\epsilon_c \rangle$ for a particular color algebra. A little thought will show that if $\langle \Gamma_b,\epsilon_b \rangle$ covers $\langle \Gamma,\epsilon \rangle$ and if $\Gamma_b \subset \Gamma_a$ then it may be possible in general to extend ϵ_b to a commutation factor Γ_a and then obviously $\langle \Gamma_a,\epsilon_a \rangle$ covers $\langle \Gamma,\epsilon \rangle$. Clearly then what we may hope for is that there is a unique *smallest* $\langle \Gamma_c,\epsilon_c \rangle$ covering $\langle \Gamma,\epsilon \rangle$. More precisely what we shall prove is that there exists a $\langle \Gamma_c,\epsilon_c \rangle$ covering $\langle \Gamma,\epsilon \rangle$ such that any other $\langle \Gamma_c',\epsilon_c' \rangle$ covering $\langle \Gamma,\epsilon \rangle$ satisfies $\Gamma_c \subseteq \Gamma_c'$. Further we shall give a criterion to determine what this minimal $\langle \Gamma_c,\epsilon_c \rangle$ is.

In order to prove the above result we need to introduce a little machinery from elementary Abelian group theory. This shall differ somewhat from the standard treatment (see Fuchs[7]) and so we shall be forced to state a number of elementary results in this field. The proofs may be found in Kleeman.[5]

First we define two notions of *linear independence*. We say that the set $\{\alpha_i\}$ of elements of Γ is *p-linearly independent* if $o(\alpha_i)$ is a power of p and if

$$\sum_i n^i\alpha_i = 0 \tag{5.1}$$

means that $\forall i$, $n^i \equiv 0 \bmod p$. Second, we say the set $\{\alpha_i\}$ is *p-linearly independent with respect to* ϵ if again $o(\alpha_i)$ is a power of p and if

$$\epsilon\left(\sum_i n^i\alpha_i,\gamma\right) = 1, \quad \forall \gamma \in \Gamma, \tag{5.2}$$

means that $\forall i$, $n^i \equiv 0 \bmod p$.

From these two definitions we are further able to define two notions of rank. We say the *p^k-rank* of Γ is the maximal number of p-linearly independent elements in $p^k\Gamma$. A similar definition holds for *p^k-rank with respect to* ϵ. In the interests of brevity we use rank when we want to talk of p^0-rank and l.i. when we wish to talk of p-linear independence.

The following two results we quote without comment.

Lemma 5.1: The rank of Z_{p^r} is 1.

Proposition 5.2:

$$\text{rank}(A \oplus B) = \text{rank}(A) + \text{rank}(B).$$

From these we derive the following important corollary.

Corollary: A finite Abelian group is determined uniquely by its p^k-ranks.

Proof: First, we use Lemma 5.1 and Proposition 5.2 to conclude that the rank of a p-group Γ_p is equal to the number of cyclic summands in its unique decomposition (2.3). Now obviously the rank of $p\Gamma_p$ is the same as that for Γ_p *less* the number of Z_p summands in Γ_p. This argument extends in an obvious way to the rank of $p^k\Gamma_p$, which is equal to the p^{k-1}-rank *less* the number of Z_{pk} cyclic summands in Γ_p. We can thus conclude that Γ_p is uniquely specified by its p^k-ranks. To extend this result to an arbitrary finite Γ, it is sufficient to observe that elements of nonprime power order do not contribute to the p^k-ranks of Γ by the definition of p-linear independence. Hence the p^k-ranks of Γ determine uniquely the unique p-groups making up the total group. □

We examine now the connection between rank and rank with respect to (w.r.t.) ϵ.

Proposition 5.3: The p^k-rank of Γ is at least as large as its p^k-rank w.r.t. ϵ. Equality holds when Γ is reduced.

Proof: Let $\alpha_i \in p^k\Gamma$ and suppose $\Sigma_i m^i\alpha_i = 0$. Further suppose that $\{\alpha_i\}$ is l.i. w.r.t. ϵ

$$\Rightarrow \epsilon\left(\sum_i m^i\alpha_i,\gamma\right) = 1, \quad \forall\gamma\in\Gamma$$

$$\Rightarrow m^i \equiv 0 \mod p, \quad \forall i,$$

which means that $\{\alpha_i\}$ are l.i. For the second part of the proposition suppose further that Γ is reduced and now that $\{\alpha_i\}$ are l.i. Now if

$$\epsilon\left(\sum_i m^i\alpha_i,\gamma\right) = 1, \quad \forall\gamma\in\Gamma,$$

then, because Γ is reduced,

$$\Rightarrow \sum_i m^i\alpha_i = 0$$

$$\Rightarrow m^i \equiv 0 \mod p, \quad \forall i,$$

which means that $\{\alpha_i\}$ are l.i. w.r.t. ϵ. □

The reason for the usefulness, from our point of view, of rank w.r.t. ϵ is contained in the following.

Proposition 5.4: The p^k-rank w.r.t. ϵ is preserved by an *onto* covering homomorphism.

Proof: Suppose $h:\Gamma\to\Gamma'$ is the onto covering homomorphism and suppose that $h(\alpha_i)\in\Gamma'$ are l.i. w.r.t. ϵ'. Now if

$$\epsilon\left(\sum_i m^i\alpha_i,\gamma\right) = 1, \quad \forall\gamma\in\Gamma$$

$$\Rightarrow \epsilon'\left(h\left(\sum_i m^i\alpha_i\right),h(\gamma)\right) = 1, \quad \forall\gamma\in\Gamma,$$

which implies, since h is onto, that

$$\epsilon'\left(\sum_i m^i h(\alpha_i),\gamma'\right) = 1, \quad \forall\gamma'\in\Gamma'$$

$$\Rightarrow m^i \equiv 0 \mod p, \quad \forall i,$$

and this therefore means that $\{\alpha_i\}$ are l.i. w.r.t. ϵ.

Conversely suppose that $\{\alpha_i\}$ are l.i. w.r.t ϵ. Now if

$$\epsilon'\left(\sum_i m^i h(\alpha_i),\gamma'\right) = 1, \quad \forall\gamma'\in\Gamma'$$

$$\Rightarrow \epsilon'\left(h\left(\sum_i m^i\alpha_i\right),h(\gamma)\right) = 1, \quad \forall\gamma\in\Gamma$$

$$\Rightarrow \epsilon\left(\sum_i m^i\alpha_i,\gamma\right) = 1, \quad \forall\gamma\in\Gamma$$

$$\Rightarrow m^i \equiv 0 \mod p, \quad \forall i,$$

and thus we conclude that $\{h(\alpha_i)\}$ is l.i. w.r.t. ϵ'.

Having dispensed with the algebraic preliminaries we are now able to prove the second major result of this paper.

Theorem 5.5: There exists a unique canonical $\langle\Gamma_c,\epsilon_c\rangle$ covering $\langle\Gamma,\epsilon\rangle$ such that if another canonical $\langle\Gamma'_c,\epsilon'\rangle$ also covers $\langle\Gamma,\epsilon\rangle$ then $\Gamma_c\subset\Gamma'_c$. Furthermore the p^k-rank of Γ_c is equal to the p^k-rank w.r.t. ϵ of Γ *unless* there exist $\beta\in\Gamma$ such that $\epsilon(\beta,\beta) = -1$ and none of these β satisfy

$$\epsilon(\beta,\gamma) = \pm 1, \quad \forall\gamma\in\Gamma. \tag{5.3}$$

In this latter case the 2^0-rank of Γ_c is one greater than the 2^0-rank w.r.t. ϵ of Γ but all other p^k-ranks are identical. Finally the unique $\langle\Gamma_c,\epsilon_c\rangle$ for every $\langle\Gamma,\epsilon\rangle$ is the one achieved in the proof of Theorem 4.1.

Proof: We begin with the following essential lemma.

Lemma 5.6: If $\Gamma_a\subseteq\Gamma_b$ then the p^k-rank w.r.t. ϵ_a of Γ_a is no greater than the corresponding rank on Γ_b (providing, of course, that ϵ_a and ϵ_b agree on Γ_a).

Proof: Suppose $\alpha_i\in p^k\Gamma_a$. Since $p^k\Gamma_a\subseteq p^k\Gamma_b$ this means that the α_i are also in $p^k\Gamma_b$. Further suppose that $\{\alpha_i\}$ is l.i. w.r.t. ϵ_a in Γ_a. Now if

$$\epsilon_b\left(\sum_i m^i\alpha_i,\gamma\right) = 1, \quad \forall\gamma\in\Gamma_b$$

$$\Rightarrow \epsilon_a\left(\sum_i m^i\alpha_i,\gamma'\right) = 1, \quad \forall\gamma'\in\Gamma_a$$

$$\Rightarrow m^i \equiv 0 \mod p, \quad \forall i,$$

which shows that $\{\alpha_i\}\subset p^k\Gamma_b$ is l.i. w.r.t. ϵ_b. □

Now if $\langle\Gamma',\epsilon'\rangle$ covers $\langle\Gamma,\epsilon\rangle$ then Theorem 3.2 tells us that Γ'/Γ'_0 contains the image of a covering homomorphism from Γ and so Proposition 5.4 and Lemma 5.6 show that this quotient group must have p^k-ranks w.r.t. ϵ' at least as large as the corresponding ranks of Γ w.r.t. ϵ. By Lemma 3.1 Γ'/Γ'_0 is the image of a covering homomorphism from Γ' and so this group has the same p^k-ranks w.r.t. ϵ'. Finally, since $\Gamma'\subseteq\Gamma'$ we deduce from Lemma 5.6 that the p^k-ranks w.r.t. ϵ' of Γ' are at least as great as the corresponding ranks w.r.t. ϵ of Γ. We also have the following lemma.

Lemma 5.7: Any canonical $\langle\Gamma_c,\epsilon_c\rangle$ is reduced.

Proof: Let $\alpha\in\Gamma_c$ satisfy

$$\epsilon_c(\alpha,\gamma) = 1, \quad \forall\gamma\in\Gamma_c,$$

and for a $p_i\neq 2$ let s^t_j be the generator of $Z^t_{(p_i)^{r_j}}$, then we have, using (4.3) and (4.4), that

$$1 = \epsilon(\alpha,s^t_j) = \epsilon_i(\alpha,s^t_j) = \eta_{ij}^{\pm k^t_j \mp 1},$$

where $k^t_j\, s^t_j$ is the projection of α onto $Z^t_{(p_i)^{r_j}}$ and the $\pm$ depends on whether t is odd or even (t even gives the $+$). It follows immediately that $k^t_j \equiv 0 \mod p_i^{r_j}$ and so $k^t_j\, s^t_j = 0$.

Since this holds for all j and t we conclude that α can have no components in p_i-subgroups of Γ_c. For $p_i = 2$ the argument is identical, except in the case that copies of Z_2 have the form (4.6a) defined on them, in which case we get the simpler equation $(-1)^{k^t} = 1$, where $k^t s^t$ is the projection onto the t th copy of Z_2 in Γ_c. Again this implies that $k^t s^t = 0$ and so we conclude that $\alpha = 0$. □

This lemma allows us to conclude from Proposition 5.3 that the p^k-rank of any covering Γ_c must be at least as large as the p^k-ranks w.r.t. ϵ of Γ. Now by the corollary to Proposition 5.2 the p^k-ranks of a group uniquely determine it and so there must be a unique minimal canonical Γ_c^m with p^k-ranks *equal* to the p^k-ranks w.r.t. ϵ of Γ. Clearly then, any $\langle \Gamma_c, \epsilon_c \rangle$ covering $\langle \Gamma, \epsilon \rangle$ must satisfy $\Gamma_c^m \subseteq \Gamma_c$.

Except in the pathological case outlined in the statement of the theorem, the proof of Theorem 4.1 has shown that there exists an *onto* covering homomorphism between Γ and Γ_c. Proposition 5.4 shows therefore that this must be in fact Γ_c^m.

In the pathological case we have seen in the proof of Theorem 4.1 that there exist onto covering homomorphisms at all but one place in the reduction—where we are forced to append an extra Z_2 summand. It follows, again from Proposition 5.4, that a reduction to $\Gamma_c^m \oplus Z_2$ is possible. Finally we complete the proof by showing that, in the pathological case, if Γ_c has 2^k ranks equal to those of Γ_c^m then it cannot cover Γ. A little thought will show that this implies that any Γ_c covering Γ must satisfy $\Gamma_c^m \oplus Z_2 \subseteq \Gamma_c$.

Let us assume that a Γ_c with 2^k ranks equal to those of Γ_c^m does cover Γ. The covering of Γ implies there exists a covering homomorphism $h: \Gamma \to \Gamma'/\Gamma_0'$ with $\Gamma' \subseteq \Gamma_c$; also there is an onto covering homomorphism $g: \Gamma' \to \Gamma'/\Gamma_0'$. Consider now the 2-subgroups of $\Gamma, \Gamma'/\Gamma_0', \Gamma'$ and Γ_c; denote them by Γ_2, Q_2, Γ_{2c}', and Γ_{2c}, respectively. Now, by Ref. 7, the covering homomorphisms h and g restrict to covering homomorphisms $\Gamma_2 \to Q_2$ and $\Gamma_{2c}' \to Q_2$, respectively; moreover, it is easily seen that the latter must be onto. Denote by $r(\Gamma)$ and $r(\Gamma, \epsilon)$ the 2^k-rank and 2^k-rank w.r.t ϵ of Γ, respectively. We have, by use of the technical lemmas and propositions above, the following inequalities:

$$r(\Gamma_2, \epsilon) = r(h(\Gamma_2), \epsilon_r) \leqslant r(Q_2, \epsilon_r) = r(\Gamma_{2c}', \epsilon_c)$$
$$\leqslant r(\Gamma_{2c}, \epsilon_c) = r(\Gamma_{2c}) = r(\Gamma_2, \epsilon), \tag{5.4}$$

which shows that equality must hold amongst all of them. Now we have seen that Γ'/Γ_0' is reduced and we now show that this implies that $Q_2 \subseteq \Gamma'/\Gamma_0'$ is reduced. Suppose $\alpha \in Q_2$ satisfies

$$\epsilon_r(\alpha, \gamma) = 1, \quad \forall \gamma \in Q_2,$$

then it is quite clear from (2.9) that in fact

$$\epsilon_r(\alpha, \gamma') = 1, \quad \forall \gamma' \in \Gamma'/\Gamma_0',$$

and so $\alpha = 0$, which shows Q_2 is reduced. We have thus that $r(\Gamma_{2c}) = r(Q_2, \epsilon_r) = r(Q_2)$ and hence

$$\Gamma_{2c} \cong Q_2. \tag{5.5}$$

A further set of inequalities are

$$r(h(\Gamma_2), \epsilon_r) < r(h(\Gamma_2)) < r(Q_2), \tag{5.6}$$

which, when (5.4) and (5.5) are considered, become equalities and lead to the conclusion that

$$h(\Gamma_2) = Q_2.$$

Combining h with the isomorphism (5.5) leads one to conclude that there is an *onto* covering homomorphism k between Γ_2 and Γ_{2c}.

We define a diagonal element $\beta \in \Gamma$ to be one satisfying $\epsilon(\beta, \beta) = -1$; such an element must exist in the pathological case. Now $\Gamma = \Gamma_2 \oplus \Gamma^t$ with Γ^t being a direct sum of $p \neq 2$ groups. The diagonal $\beta \in \Gamma$ must be able, therefore, to be written as $\beta = \alpha + \gamma$ with $\alpha \in \Gamma_2$ and $\gamma \in \Gamma^t$. Now

$$-1 = \epsilon(\beta, \beta) = \epsilon(\alpha + \gamma, \alpha + \gamma)$$
$$= \epsilon(\alpha, \alpha)\epsilon(\gamma, \gamma) = \epsilon(\alpha, \alpha), \tag{5.7}$$

the last step following from the results of Sec. II. Clearly, then, $\alpha \in \Gamma_2$ is diagonal, which implies, from the definition of the covering homomorphism, that $k(\Gamma_2) = \Gamma_{2c}$ also has a diagonal element. Examination of (4.6a) then shows that Γ_{2c} must possess a diagonal element satisfying (5.3), but since Γ_{2c} is the image of a covering homomorphism from Γ_2 it follows that this latter group must also contain such an element. This, however, contradicts the assumption that Γ is pathological and so we are done. □

VI. CONCLUSIONS

We have seen that for a finite grading group Γ any color algebra A with commutation factor ϵ can also be considered to have grading group Γ_c and a commutation factor almost determined by Γ_c. Moreover we have seen that there exists a unique "minimal" Γ_c^m, which will do the job.

Clearly the next step would be to consider finitely generated grading groups. Despite the complication mentioned in Sec. II this is not likely to be all that difficult given the mathematical machinery that has been developed here. The concept of the covering homomorphism is likely to prove central to any such extension. Any further extension to more infinite groups will require more development of the basic work of Scheunert[3]—particularly as to the question of multipliers on such groups.

[1] V. Rittenberg and D. Wyler, Nucl. Phys. B **139**, 189 (1978).
[2] R. Kleeman, J. Math. Phys. **24**, 166 (1983).
[3] M. Scheunert, J. Math. Phys. **20**, 712 (1979).
[4] O. Klein, J. Phys. Radium **9**, 1 (1938).
[5] R. Kleeman, Ph.D. Thesis, University of Adelaide, 1985.
[6] L. Fuchs, *Abelian Groups* (Pergamon, New York, 1960), Chap. II.
[7] Reference 6, Chap. I.
[8] We drop the subscript i here for notational ease.
[9] K. Shoda, Math. Ann. **100**, 674 (1928).

On a labeling for point group harmonics. I. Cubic group

Robert Conte
Service de Physique du Solide et de Résonance Magnétique, CEN-SACLAY, 91191 Gif-sur-Yvette Cedex, France

Jacques Raynal
Service de Physique Théorique, CEN-SACLAY, 91191 Gif-sur-Yvette Cedex, France

(Received 17 July 1984; accepted for publication 4 April 1985)

Expressions for the Γ_4, Γ_6, and Γ_8 representations are derived in terms of the Γ_1 representations for neighboring angular momenta. Expressions for Γ_5, Γ_7, and again for Γ_8 are derived in terms of the Γ_2 representations. A third set of expressions for Γ_8 is derived in terms of the Γ_3 representations. With an arbitrary choice of orthonormal sets of the Γ_1, Γ_2, and Γ_3 representations, orthonormal sets of all other kinds of representations are thus well defined and can be entirely labeled using the parent representations. All Clebsch–Gordan coefficients are expressed in terms of those between parent representations (and a few other ones). The Γ_4 and Γ_5 representations defined here are not those conventionally used, but they provide a simpler expression for the fictitious spin coefficients. Tables of the parent representations Γ_1, Γ_2, and Γ_3 are given for quaternary and ternary axes of quantization. Using the usual Clebsch–Gordan coefficients of SU(2), these tables allow us to obtain any representation for integer or half-integer angular momenta up to $j = 25$.

I. INTRODUCTION

In a previous article,[1] we gave expressions for cubic harmonics quantized along an axis of order 4 in terms of reduced rotation matrix elements with argument $\pi/2$. These expressions are uniquely defined (to within a sign) only if the multiplicity of the involved representation is unity.

The number $n_i(j)$ of representations Γ_i obtained for a given value of j can be derived from the characters of SU(2) and of the double cubic group. They are given in Table I, with their generating functions,[2,3] which we define by

$$g_i(x) = \sum_j x^{2j} n_i(j). \tag{1}$$

Among them, $g_1(x)$ has been known for a long time.[4] Explicit values for the multiplicities for the 12 first integer and half-integer angular momenta are given in Table II.

The coefficients obtained using the standard basis $|jm\rangle$ of the Γ_1 representations, for which the multiplicity is unity, are very simple, even for the case of $j = 23$. Their squares are the products of prime numbers smaller than $2j$, divided by a large power of 2 and also sometimes by 3. The only exceptions are for $j = 13$, where 5 appears to the power 2, and $j = 14$ and 15, when 3 appears to the power 1 instead of the usual 0 or -1. The coefficients of the Γ_2 representations are similar in the 12 cases for which the multiplicity is unity: besides some anomalous powers of 3, the only exceptions are powers of 5 and 7 for $j = 17$. For all the other representations where the multiplicity is unity, the coefficients have the above described form, with only some anomalous powers of 3.

One of the components of the representations Γ_4 and Γ_5 spans a complete subspace of D_4. When the multiplicity is larger than unity, this component can be chosen to be diagonal in this subspace: the coefficients of the other components are square roots of a product of prime numbers smaller than $2j$ multiplied by some integer which can be very large. A natural labeling of these representations is obtained by attributing the label of the vector of the subspace of D_4, which is one of their components. We shall show later that this labeling is not always the best.

For the other kinds of representations, when the multiplicity is larger than unity, an arbitrary rotation in the subspace of these representations can be performed in the hope of obtaining the simplest possible coefficients. Searches for Γ_1, Γ_2, and Γ_3 lead to the results given later in Table V: a large prime number cannot always be avoided in the denominator. Such a search for the "simplest" expressions was the criterium adopted by Butler[5] when computing the Clebsch–Gordan coefficients between all kinds of representations.

In the next section, we show that the coefficients of the Γ_6, Γ_4, and Γ_8 representations for a given value of j can be obtained by multiplying the coefficients of neighboring Γ_1 representations by a Clebsch–Gordan coefficient taken from SU(2). The set of representations obtained in this way is orthonormal. So, there is a choice of Γ_6, Γ_4, and Γ_8 representations, which can be labeled by the angular momentum l and the extra label α of the "parent" Γ_1 representations from which they are built. Due to the relative simplicity of the Clebsch–Gordan coefficients of SU(2) when a low angular momentum is involved, we believe that the coefficients of the Γ_6, Γ_4, and Γ_8 representations are the simplest possible if the parent Γ_1 representations are either nondegenerate or well chosen. A glimpse at the decomposition into irreducible representations of the direct product of two representations, as given in Table III, shows that the Γ_7, Γ_5, and Γ_8 representations are obtained in a similar way if the Γ_2 representations are used instead of Γ_1. Only the Γ_3 representations are left; the Γ_8 representations can also be obtained from them. So, when the Γ_1, Γ_2, and Γ_3 representations are known (and labeled), all the other kinds of representations are known (and labeled). However, there are three labelings of Γ_8 representations, with respect to the Γ_1, Γ_2, and Γ_3 representa-

 0022-2488/85/102413-28$02.50

TABLE I. Multiplicities and their generating functions. The second column gives the multiplicity for a given value of j, using the a's quoted at the bottom of the table. The generating functions of the last column use x^2 instead of x in order to simplify relations between multiplicities for different values of j, integer or half-integer.

	$n(j)$	$g(x)$
Γ_1	$(a_1+9a_2+8a_3+6a_4)/24$	$\dfrac{1+x^{18}}{(1-x^8)(1-x^{12})}$
Γ_2	$(a_1-3a_2+8a_3-6a_4)/24$	$\dfrac{x^6}{(1-x^6)(1-x^8)}$
Γ_3	$(a_1+3a_2-4a_3)/12$	$\dfrac{x^4}{(1-x^4)(1-x^6)}$
Γ_4	$(a_1-3a_2+2a_4)/8$	$\dfrac{x^2(1+x^6)}{(1-x^4)(1-x^8)}$
Γ_5	$(a_1+a_2-2a_4)/8$	$\dfrac{x^4}{(1-x^2)(1-x^8)}$
Γ_6	$(a_1+4a_3+3\sqrt{2}a_4)/12$	$\dfrac{x+x^{11}}{(1-x^6)(1-x^8)}$
Γ_7	$(a_1+4a_3-3\sqrt{2}a_4)/12$	$\dfrac{x^5(1+x^2)}{(1-x^6)(1-x^8)}$
Γ_8	$(a_1-2a_3)/6$	$\dfrac{x^3}{(1-x^2)(1-x^6)}$
	$a_1=2j+1$ $a_3=2\sin(a_1\pi/3)/\sqrt{3}$	$a_2=\sin(a_1\pi/2)$ $a_4=\sqrt{2}\sin(a_1\pi/4)$

tions, respectively; to distinguish between them, we shall quote them as Γ_8, Γ_8', and Γ_8'', respectively. The representations published by Butler[5] are the same, to within a sign, as those constructed by this method, because the consequences of this construction coincide with the criteria of simplicity in Ref. 5. However, in his tables established up to $j=8$, Butler gives Γ_8 representations for $j=\frac{11}{2}$ and $\frac{15}{2}$, Γ_8' representations for $j=\frac{13}{2}$, and Γ_8'' representations for $j=\frac{9}{2}$.

In the third section, we study some consequences of these labelings in all problems which involve Clebsch–Gordan coefficients, such as the decomposition of a product of two representations as a sum of irreducible representations or the computation of the matrix element of a given operator. Notations in this field vary strongly among different authors.[5–11] With the components of three representations

$$|\Gamma_i\, j\,\alpha\hat{m}\rangle=\sum_{\mu}x_{\mu}^{j\alpha}|j\mu\rangle, \tag{2}$$

we define the quantity

$$\begin{pmatrix}\Gamma_i\, j\alpha & \Gamma_{i'}\, j'\alpha' & \Gamma_{i''}\, j''\alpha''\\ \hat{m} & \hat{m}' & \hat{m}''\end{pmatrix}=\sum_{\mu,\mu',\mu''}x_{\mu}^{j\alpha}x_{\mu'}^{j'\alpha'}x_{\mu''}^{j''\alpha''}\begin{pmatrix}j & j' & j''\\ \mu & \mu' & \mu''\end{pmatrix}, \tag{3}$$

which we call the "cubic $3jm$ symbol" because we consider it simply as a sum of $3jm$ symbols of SU(2) weighted by the coefficients of cubic harmonics. In fact, since all our coefficients $x_{\mu}^{j\alpha}$ are real, this cubic $3jm$ symbol coincides with the $\bar{f}$ symbol introduced by Kibler.[6] The dependence of the cubic $3jm$ symbol on magnetic quantum numbers $\hat{m}$ is eliminated by dividing it by a similar expression obtained with the "basic" representations

$$|\gamma_i\, q_i\, \hat{m}\rangle=\sum_{\mu}x_{\mu}^{q_i}|q_i\,\mu\rangle, \tag{4}$$

which are the representations of the same kind as Γ_i but with the lowest possible angular momentum. The values of q_i are 0, 3, 2, 1, 2, $\frac{1}{2}$, $\frac{5}{2}$, and $\frac{3}{2}$ for Γ_1 to Γ_8, respectively. Except for two Γ_8 and a Γ_4 or Γ_5 representations, writing

$$\begin{pmatrix}\Gamma_i\, j\alpha & \Gamma_{i'}\, j'\alpha' & \Gamma_{i''}\, j''\alpha''\\ \hat{m} & \hat{m}' & \hat{m}''\end{pmatrix}=(\Gamma_i\, j\alpha\ \Gamma_{i'}\, j'\alpha'\ \Gamma_{i''}\, j''\alpha'')\begin{pmatrix}\gamma_i q_i & \gamma_{i'} q_{i'} & \gamma_{i''} q_{i''}\\ \hat{m} & \hat{m}' & \hat{m}''\end{pmatrix}, \tag{5}$$

we define what we call a "3Γ symbol" $(\Gamma_i\, j\alpha\ \Gamma_{i'}\, j'\alpha'\ \Gamma_{i''}\, j''\alpha'')$, which is a reduced matrix element. With two Γ_8 and one Γ_4 or Γ_5 representations, a sum of two terms has to be written on the right-hand side of (5), the second of them involving a "second basic" Γ_4, Γ_5, or Γ_8 representation with the next lowest angular momentum which is 3, 3, and $\frac{5}{2}$, respectively. In this definition, the cubic $3jm$ symbols of the "basic" representations are not the V coefficients[7] of the cubic group: for example, with two Γ_2 and a Γ_1 representations our cubic $3jm$ symbol is $7^{-1/2}$ while the V coefficient is 1. All the 3Γ symbols are unity between basic representations. They are the quotient of two reduced coefficients $\bar{f}(( ))$ of Kibler or of the $3jm$ factor tabulated by Butler by the $3jm$ for "basic representations." The reason why we adopt such definitions is that it allows us to use the Racah algebra of SU(2) without introducing any numerical factor coming from the cubic group itself, and also that the results can be translated by anybody into his own notation. These cubic $3jm$ symbols and the first 3Γ symbols are invariant with respect to a circular permutation of their arguments; for an odd permutation, there is a phase $(-)^{j+j'+j''}$ for the cubic $3jm$ symbols and $(-)^{j+j'+j''-q_i-q_{i'}-q_{i''}}$ for the 3Γ symbols.

As seen in Table III, there are 36 different kinds of 3Γ symbols, including the four cases in which a second term is

TABLE II. Multiplicities of the irreducible representations for the first 12 integer and half-integer angular momenta. The last lines give the period of the multiplicity and the increment: when the angular momentum is increased by the period, the multiplicity is increased by the increment.

l	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	j
0	1	0	0	0	0				
						1	0	0	$\frac{1}{2}$
1	0	0	0	1	0				
						0	0	1	$\frac{3}{2}$
2	0	0	1	0	1				
						0	1	1	$\frac{5}{2}$
3	0	1	0	1	1				
						1	1	1	$\frac{7}{2}$
4	1	0	1	1	1				
						1	0	2	$\frac{9}{2}$
5	0	0	1	2	1				
						1	1	2	$\frac{11}{2}$
6	1	1	1	1	2				
						1	2	2	$\frac{13}{2}$
7	0	1	1	2	2				
						1	1	3	$\frac{15}{2}$
8	1	0	2	2	2				
						2	1	3	$\frac{17}{2}$
9	1	1	1	3	2				
						2	2	3	$\frac{19}{2}$
10	1	1	2	2	3				
						1	2	4	$\frac{21}{2}$
11	0	1	2	3	3				
						2	2	4	$\frac{23}{2}$
Per.	12	12	6	4	4	12	12	3	
Inc.	1	1	1	1	1	2	2	1	

needed in (5). In fact, the existence of alternate Γ_8' and Γ_8'' representations introduces 57 other kinds of 3Γ symbols. With the definition of the Γ_i representation for $i = 4$ to 8, all these 3Γ symbols can be expressed in terms of those of the parent Γ_1, Γ_2, or Γ_3 representations.

In the fourth section, we consider the results obtained when one of the three cubic harmonics is a vector, i.e., the Γ_4 representation with $l = 1$, and the two others are of the same kind. Except with two Γ_8 representations, the result is diagonal. One of the consequences of this result is that the fictitious spin[12] is diagonal for the Γ_4 and Γ_5 representations labeled as proposed here: labeling with respect to a D_4 subspace as is usual[12–14] does not provide this property. Butler's[5] Γ_4 and Γ_5 representations are the same but the numerical value of their fictitious spin is not given by their label.

In the next section, we give the expressions of the components of Γ_4 to Γ_8 representations when a ternary axis is used for quantization. The last section gives tables of Γ_1 and Γ_2 up to $l = 26$, Γ_3 up to $l = 25$. With the formulas derived in the second and fifth sections, components of any representation up to $j = 25$ can be obtained from these values. Some examples are given.

II. EXPRESSION OF OTHER REPRESENTATIONS IN TERMS OF THE Γ_1, Γ_2, AND Γ_3 REPRESENTATIONS

For a given value of the angular momentum l, the expression of the Γ_1 representation is

$$|\Gamma_1\, l\, \alpha\, \hat{0}\rangle = \sum_{m\equiv 0(4)} a^{\alpha}_{lm} |l\, m\rangle, \tag{6}$$

where $a^{\alpha}_{lm} = (-)^l a^{\alpha}_{l-m}$. This notation is not the same as that used in a previous work.[1] The sum over m is extended to negative values and, for an odd value of l, $m = 0$ is taken into account with a vanishing coefficient in order to obtain formulas that are easier to read. The index α is introduced to label the $n_1(l)$ different orthonormal Γ_1 representations which can be obtained for this value of l.

We shall show below that

$$|\Gamma_1\, j\, l\, \alpha\, \hat{\sigma}\rangle = (-)^{l-s+\mu}\sqrt{(2s+1)(2l+1)} \times \sum_{m\equiv 0(4)} a^{\alpha}_{lm} \begin{pmatrix} l & s & j \\ m & \hat{\sigma} & -\mu \end{pmatrix} |j\mu\rangle \tag{7}$$

is an orthonormal set of the Γ_6, Γ_4, and Γ_8 representations for $s = \frac{1}{2}$, 1, and $\frac{3}{2}$, respectively. In (7) there is only the Clebsch–Gordan coefficient coupling $|lm\rangle|s\hat{\sigma}\rangle$ to $|j\mu\rangle$ and a normalization constant; the angular momenta dependence is rather that of the Clebsch–Gordan coefficient coupling $|j\mu\rangle|s-\hat{\sigma}\rangle$ to $|lm\rangle$, plus a phase. This result is trivial when written between basic representations (4) because $|s\sigma\rangle$ for $s = \frac{1}{2}$, 1, and $\frac{3}{2}$ is the basic Γ_6, Γ_4, and Γ_8 representation of the cubic group, respectively, as can be seen in Table II.

For a given value of l, the Γ_2 representations are

$$|\Gamma_2\, l\, \beta\, \hat{2}\rangle = \sum_{m\equiv 2(4)} b^{\beta}_{lm} |l\, m\rangle, \tag{8}$$

where $b^{\beta}_{lm} = (-)^l b^{\beta}_{l-m}$. Here also, the sum over m is extended to negative values and β labels the $n_2(l)$ different orthonormal Γ_2 representations which can be obtained for this value of l. As above,

TABLE III. Decomposition of the product of two irreducible representations.

	$\Gamma_1\Gamma_2\Gamma_3$	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8
Γ_1	$\Gamma_1\Gamma_2\Gamma_3$	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8
Γ_2	$\Gamma_2\Gamma_1\Gamma_3$	Γ_5	Γ_4	Γ_7	Γ_6	Γ_8
Γ_3	$\Gamma_3\Gamma_3\Gamma_1+\Gamma_2+\Gamma_3$	$\Gamma_4+\Gamma_5$	$\Gamma_4+\Gamma_5$	Γ_8	Γ_8	$\Gamma_6+\Gamma_7+\Gamma_8$
Γ_4	$\Gamma_4\Gamma_5\Gamma_4+\Gamma_5$	$\Gamma_1+\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_2+\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_6+\Gamma_8$	$\Gamma_7+\Gamma_8$	$\Gamma_6+\Gamma_7+2\Gamma_8$
Γ_5	$\Gamma_5\Gamma_4\Gamma_4+\Gamma_5$	$\Gamma_2+\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_1+\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_7+\Gamma_8$	$\Gamma_6+\Gamma_8$	$\Gamma_6+\Gamma_7+2\Gamma_8$
Γ_6	$\Gamma_6\Gamma_7\Gamma_8$	$\Gamma_6+\Gamma_8$	$\Gamma_7+\Gamma_8$	$\Gamma_1+\Gamma_4$	$\Gamma_2+\Gamma_5$	$\Gamma_3+\Gamma_4+\Gamma_5$
Γ_7	$\Gamma_7\Gamma_6\Gamma_8$	$\Gamma_7+\Gamma_8$	$\Gamma_6+\Gamma_8$	$\Gamma_2+\Gamma_5$	$\Gamma_1+\Gamma_4$	$\Gamma_3+\Gamma_4+\Gamma_5$
Γ_8	$\Gamma_8\Gamma_8\Gamma_6+\Gamma_7+\Gamma_8$	$\Gamma_6+\Gamma_7+2\Gamma_8$	$\Gamma_6+\Gamma_7+2\Gamma_8$	$\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_3+\Gamma_4+\Gamma_5$	$\Gamma_1+\Gamma_2+\Gamma_3+2\Gamma_4+2\Gamma_5$

$$|\Gamma_i\, j\, l\, \beta\, \hat{\sigma}\rangle = \epsilon_{i\hat{\sigma}}(-)^{l-s+\mu}\sqrt{(2s+1)(2l+1)} \times \sum_{m\equiv\hat{2}(4)} b^{\beta}_{lm}\begin{pmatrix} l & s & j \\ m & \sigma & -\mu \end{pmatrix}|j\mu\rangle \tag{9}$$

is an orthonormal set of Γ_7, Γ_5, and Γ_8 representations for $s=\frac{1}{2}$, 1, and $\frac{3}{2}$, respectively. The value of σ is related to $\hat{\sigma}$ by $\sigma+2\equiv\hat{\sigma}(4)$. The phase $\epsilon_{i\hat{\sigma}}$ can be obtained for each kind of representation, component after component, using standard representation in (9): the phase turns out to be $(-)^{j+s}$ for Γ_7 and Γ_5 representations, 1 for Γ_8 components with $|\hat{\sigma}|=\frac{1}{2}$, and -1 for Γ_8 components with $|\hat{\sigma}|=\frac{3}{2}$. Obviously, we could choose the phase of Γ_7 and Γ_5 representations so as to eliminate this phase, but this is not possible for Γ_8 representations, already defined by (7).

So, using (7) or (9), the Γ_4 to Γ_8 representations can be labeled by the angular momentum j and the labels l and α or β of the parent Γ_1 or Γ_2 representation. The explicit formulas are

$$|\Gamma_4\, j\, l\, \alpha\hat{\sigma}\rangle = (-)^{l-1+\mu}\sqrt{3(2l+1)} \times \sum_{m\equiv 0(4)} a^{\alpha}_{lm}\begin{pmatrix} l & 1 & j \\ m & \hat{\sigma} & -\mu \end{pmatrix}|j\mu\rangle, \tag{10}$$

$$|\Gamma_5\, j\, l\, \beta\hat{\sigma}\rangle = (-)^{l+j+\mu}\sqrt{3(2l+1)} \times \sum_{m\equiv\hat{2}(4)} b^{\beta}_{lm}\begin{pmatrix} l & 1 & j \\ m & \sigma & -\mu \end{pmatrix}|j\mu\rangle, \tag{11}$$

$$|\Gamma_6\, j\, l\, \alpha\, \hat{\sigma}\rangle = (-)^{l-1/2+\mu}\sqrt{2(2l+1)} \times \sum_{m\equiv 0(4)} a^{\alpha}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \hat{\sigma} & -\mu \end{pmatrix}|j\mu\rangle, \tag{12}$$

$$|\Gamma_7\, j\, l\, \beta\, \hat{\sigma}\rangle = (-)^{l+j+\mu}\sqrt{2(2l+1)} \times \sum_{m\equiv\hat{2}(4)} b^{\beta}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \sigma & -\mu \end{pmatrix}|j\mu\rangle, \tag{13}$$

$$|\Gamma_8\, j\, l\, \alpha\, \hat{\sigma}\rangle = (-)^{l-3/2+\mu}2\sqrt{2l+1} \times \sum_{m\equiv 0(4)} a^{\alpha}_{lm}\begin{pmatrix} l & \frac{3}{2} & j \\ m & \hat{\sigma} & -\mu \end{pmatrix}|j\mu\rangle, \tag{14}$$

and the alternate choice

$$|\Gamma'_8\, j\, l\, \beta\, \pm\hat{\sigma}\rangle = \mp(-)^{l}2\sqrt{2l+1} \times \sum_{m\equiv\hat{2}(4)} b^{\beta}_{lm}\begin{pmatrix} l & \frac{3}{2} & j \\ m & \pm\sigma & -\mu \end{pmatrix}|j\mu\rangle, \tag{15}$$

where $\hat{\sigma}$ is positive. In (11), (13), and (15), $\sigma+2\equiv\hat{\sigma}(4)$.

The Γ_3 representations are not given by these formulas. Their two components are

$$|\Gamma_3\, l\, \gamma\, \hat{0}\rangle = \sum_{m\equiv 0(4)} c^{\gamma}_{lm}|l\, m\rangle,$$
$$|\Gamma_3\, l\, \gamma\, \hat{2}\rangle = \sum_{m\equiv\hat{2}(4)} d^{\gamma}_{lm}|l\, m\rangle, \tag{16}$$

with $c^{\gamma}_{l-m}=(-)^{l}c^{\gamma}_{lm}$ and $d^{\gamma}_{l-m}=(-)^{l}d^{\gamma}_{lm}$ and γ labels the $n_3(l)$ orthonormal Γ_3 representations, which can be obtained for this value of l. With them can be defined a third equivalent set of orthonormal Γ_8 representations, which is

$$|\Gamma''_8\, j\, l\, \gamma\, \pm\hat{\tfrac{3}{2}}\rangle = (-)^{l+j-1/2}\sqrt{2(2l+1)} \times \sum_{m\equiv\hat{2}(4)} d^{\gamma}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \mp\frac{1}{2} & -\mu \end{pmatrix}|j\mu\rangle,$$
$$|\Gamma''_8\, j\, l\, \gamma\pm\hat{\tfrac{1}{2}}\rangle = (-)^{l+j-1/2}\sqrt{2(2l+1)} \times \sum_{m\equiv 0(4)} c^{\gamma}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \pm\frac{1}{2} & -\mu \end{pmatrix}|j\mu\rangle. \tag{17}$$

With the first Γ_2 and Γ_3 representations and the first two Γ_1 representations given in tables of Refs. 12 and 13, the first two representations of each kind agree except for the Γ_6 representation with $j=\frac{7}{2}$; the Γ_4 representation with $j=3$ agrees with Ref. 13 but has opposite signs in Ref. 12.

In order to prove that Eqs. (10)–(15) and (17) define orthonormal sets of representations, let us consider the scalar product of two components:

$$\begin{aligned}\langle\Gamma_{i_1} j l_1 \nu_1 \hat{\sigma}_1|\Gamma_{i_2} j\, l_2\, \nu_2\, \hat{\sigma}_2\rangle &= \epsilon_1\epsilon_2(-)^{l_1-s_1+\hat{\sigma}_1+l_2-s_2+\hat{\sigma}_2} \times\sqrt{(2s_1+1)(2s_2+1)(2l_1+1)(2l_2+1)} \\ &\times\sum_{m_1,m_2} x^{\nu_1}_{l_1m_1}x^{\nu_2}_{l_2m_2} \times\begin{pmatrix} l_1 & s_1 & j \\ m_1 & \sigma_1 & -\mu \end{pmatrix}\begin{pmatrix} l_2 & s_2 & j \\ m_2 & \sigma_2 & -\mu \end{pmatrix},\end{aligned} \tag{18}$$

where x stands for a, b, c, or d; ν for α, β, or γ; and ϵ is an extra sign which appears in (9). It is convenient to define

$$N(L,M) = (-)^{l_2}\sqrt{2l_2+1} \times\sum_{m_1m_2} x^{\nu_1}_{l_1m_1}x^{\nu_2}_{l_2m_2}\begin{pmatrix} l_1 & l_2 & L \\ m_1 & -m_2 & M \end{pmatrix}, \tag{19}$$

from which the product $x^{\nu_1}_{l_1m_1}x^{\nu_2}_{l_1m_2}$ can be obtained using the orthogonality properties of the $3jm$ symbols and inserted in (18). The result is

$$\begin{aligned}\langle\Gamma_{i_1} j\, l_1\, \nu_1\, \hat{\sigma}_1|\Gamma_{i_2} j l_2\, \nu_2\, \hat{\sigma}_2\rangle &= \epsilon_1\epsilon_2(-)^{l_1+j+\sigma_1}\sqrt{(2s_1+1)(2s_2+1)(2l_1+1)} \times\sum_{L,M}(2L+1)\begin{Bmatrix} s_1 & s_2 & L \\ l_2 & l_1 & j \end{Bmatrix} \\ &\times\begin{pmatrix} s_2 & s_1 & L \\ \sigma_2 & -\sigma_1 & M \end{pmatrix}N(L,M),\end{aligned} \tag{20}$$

which implies $L\leqslant s_1+s_2$.

If the parent representations are both Γ_1 or Γ_2 representations, the vector

$$\sum_M N(L,M)|LM\rangle \tag{21}$$

is a Γ_1 representation and the allowed values of L are 0,4,6, etc. Therefore, 0 is the only value allowed in the sum of (20). If the parent states are both Γ_3 representations, the vector (21) can be a mixture of Γ_1, Γ_2, and Γ_3 representations and the allowed values for L are 0, 2, 3, etc.; but $s_1+s_2\leqslant 1$ and the sum in (20) is again restricted to $L=0$. So, when both parent representations are of the same kind,

$$\langle \Gamma_{i_1} j l_1 \nu_1 \hat{\sigma}_1 | \Gamma_{i_2} j l_2 \nu_2 \hat{\sigma}_2 \rangle = \delta_{s_1 s_2} \delta_{\sigma_1 \sigma_2} \delta_{l_1 l_2} N(0,0), \tag{22}$$

$$N(0,0) = \sum_m x_{l_1 m}^{\nu_1} x_{l_1 m}^{\nu_2} = \delta_{\nu_1 \nu_2}.$$

If the parents are a Γ_1 and a Γ_2 representation, respectively, the vector (21) is a Γ_2 representation and the minimum value of L is 3. There are no terms in the sum of (20) except when used to compute the overlap between a Γ_8 and Γ_8' representation. With the notations defined by (3) and (5) and taking into account the coefficients $b_{3,2} = -b_{3,-2} = 2^{-1/2}$ of the basic Γ_2 representation

$$N(3,\pm 2) = \pm\left(\frac{2l_2+1}{2}\right)^{1/2} \begin{pmatrix} \Gamma_1 l_1 \alpha & \Gamma_2 l_2 \beta & \Gamma_2 3 \\ \hat{0} & \hat{2} & \hat{2} \end{pmatrix}$$

$$= \pm\left(\frac{(2l_2+1)}{14}\right)^{1/2} (\Gamma_1 l_1 \alpha \quad \Gamma_2 l_2 \beta \quad \Gamma_2 3) \tag{23}$$

and

$$\langle \Gamma_8 j l_1 \alpha \hat{\sigma} | \Gamma_8' j l_2 \beta \hat{\sigma} \rangle = (-)^{l_1+j+1/2} 2\sqrt{(2l_1+1)(2l_2+1)} \times \begin{Bmatrix} \frac{3}{2} & \frac{3}{2} & 3 \\ l_1 & l_2 & j \end{Bmatrix} (\Gamma_1 l_1 \alpha \quad \Gamma_2 l_2 \beta \quad \Gamma_2 3), \tag{24}$$

where

$$\begin{Bmatrix} \frac{3}{2} & \frac{3}{2} & 3 \\ l_1 & l_2 & j \end{Bmatrix} = \frac{(-)^{l_1+l_2+1}}{2\sqrt{5.7}} \left(\frac{(l_1+l_2+4)!(l_1-l_2+3)!(l_2-l_1+3)!(l_1+j-\frac{3}{2})!(l_2+j-\frac{3}{2})!}{(l_1+l_2-3)!(l_1+j+\frac{5}{2})!(l_1-j+\frac{3}{2})!(j-l_1+\frac{3}{2})!(l_2+j+\frac{5}{2})!(l_2-j+\frac{3}{2})!(j-l_2+\frac{3}{2})!} \right)^{1/2}. \tag{25}$$

If one of the parent representations is a Γ_3 representation and the other a Γ_1 or Γ_2 representation, the vector (21) is a Γ_3 representation and the value of L is 2. As in the previous case, $N(2,\pm 2)$ can be expressed in terms of the 3Γ symbol involving the parent representations and the basic Γ_3 representation for which the angular momentum is 2. We get

$$\langle \Gamma_8 j l_1 \alpha \hat{\sigma} | \Gamma_8'' j l_2 \gamma \hat{\sigma} \rangle = (-)^{l_1} 2\sqrt{(2l_1+1)(2l_2+1)} \begin{Bmatrix} \frac{3}{2} & \frac{1}{2} & 2 \\ l_1 & l_2 & j \end{Bmatrix} (\Gamma_1 l_1 \alpha \quad \Gamma_3 l_2 \gamma \quad \Gamma_3 2) \tag{26}$$

or

$$\langle \Gamma_8' j l_1 \beta \hat{\sigma} | \Gamma_8'' j l_2 \gamma \hat{\sigma} \rangle = (-)^{l_1} \sqrt{10(2l_1+1)(2l_2+1)} \begin{Bmatrix} \frac{3}{2} & \frac{1}{2} & 2 \\ l_2 & l_1 & j \end{Bmatrix} (\Gamma_2 l_1 \beta \quad \Gamma_3 l_1 \gamma \quad \Gamma_3 2), \tag{27}$$

where

$$\begin{Bmatrix} \frac{3}{2} & \frac{1}{2} & 2 \\ l_1 & l_2 & j \end{Bmatrix} = \frac{(-)^{l_1+l_2}}{2\sqrt{5}} \left(\frac{(l_1+l_2+3)!(l_1-l_2+2)!(l_2-l_1+2)!(l_1+j-\frac{1}{2})!(l_2+j-\frac{3}{2})!}{(l_1+l_2-2)!(l_1-j+\frac{1}{2})!(l_1+j+\frac{3}{2})!(j-l_1+\frac{1}{2})!(l_2+j-\frac{3}{2})!} \right)^{1/2}$$

$$\times [(j-l_2+\tfrac{3}{2})!(l_2+j+\tfrac{5}{2})!]^{-1/2}. \tag{28}$$

The Γ_8, Γ_8', Γ_8'' representations are independently defined. They coincide for the two lowest angular momenta. They must not be confused even if there is only one representation. For example, the parent of the Γ_8 representation for $j=2.5$ and $j=3.5$ is the Γ_1 representation for $l=4$, the parent of the related Γ_8' representations is the Γ_2 representation for $l=3$: the coefficients of the Γ_8 and the Γ_8' representations for $j=2.5$ are identical and those for $j=3.5$ are opposite.

Formulas (7) and (9) cannot be extended to a larger value of s because the sum in (20) does not reduce to a single term. When multiplied by $|s\sigma\rangle$ with s larger than $\frac{3}{2}$, a Γ_1 representation generates at least two different kinds of representations with different normalizations. For example, a Γ_7 and a Γ_8 representation are obtained with $s=\frac{5}{2}$. In this case, the representations obtained from $l=j-\frac{5}{2}$ are orthogonal to those obtained from $l=j+\frac{5}{2}$ and are a complete set for the Γ_7 representations because $n_7(j)$ is the sum of $n_1(j-\frac{5}{2})$ and $n_1(j+\frac{5}{2})$. Unfortunately, the Γ_7 representations obtained from a set of orthogonal Γ_1 representations with the same $l=j\pm\frac{5}{2}$ are not mutually orthogonal. In fact, their norm and their mutual overlap depend on the cubic $3jm$ symbol of the two Γ_1 parent representations and the Γ_1 $l=4$. Therefore we could get a set of orthonormal Γ_7 representations from a set of Γ_1 representations which diagonalizes the invariant polynomial of degree 4; such representations would, however, have the inconvenience that the square of their coefficients is not a rational number.

Relations (10)–(15) and (17) are the only ones which can be used to generate an orthonormal set of representations without condition on the parent representations. They are related to the following relations between multiplicities:

$$\begin{aligned} n_4(j) &= \sum_{l=|j-1|}^{j+1} n_1(l), \\ n_5(j) &= \sum_{l=j-1}^{j+1} n_2(l), \\ n_6(j) &= \sum_{l=|j-1/2|}^{j+1/2} n_1(l), \\ n_7(j) &= \sum_{l=j-1/2}^{j+1/2} n_2(l), \\ n_8(j) &= \sum_{l=|j-3/2|}^{j+3/2} n_1(l) = \sum_{l=j-3/2}^{j+3/2} n_2(l) = \sum_{l=j-1/2}^{j+1/2} n_3(l), \end{aligned} \tag{29}$$

which can also be checked on the generating functions of Table I as

$$g_4(x) = (1/x^2 + 1 + x^2)g_1(x) - 1/x^2 - 1,$$
$$g_5(x) = (1/x^2 + 1 + x^2)g_2(x), \tag{30}$$
$$\vdots .$$

III. THE 3*jm* SYMBOLS OF THE CUBIC HARMONICS

The cubic $3jm$ symbols defined by (3) between any three kinds of representations can be expressed in terms of those involving the parent Γ_1, Γ_2, or Γ_3 representations. This expression is

$$\begin{pmatrix} \Gamma_{i_1} j_1 l_1 \nu_1 & \Gamma_{i_2} j_2 l_2 \nu_2 & \Gamma_{i_3} j_3 l_3 \nu_3 \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix}$$
$$= \sum_{k,L,M} A(s_1 s_2 s_3; \sigma_1 \sigma_2 \sigma_3; kLM)$$
$$\times \sum_{L'} B(j_1 j_2 j_3 s_1 s_2 l_1 l_2; L'kL)$$
$$\times C(\Gamma_{p_1} l_1 \nu_1 \hat{m}_1, \Gamma_{p_2} l_2 \nu_2 \hat{m}_2, L', \Gamma_{p_3} l_3 \nu_3 \hat{m}_3; LM), \tag{31}$$

where

$$A(s_1 s_2 s_3; \sigma_1 \sigma_2 \sigma_3; kLM)$$
$$= (-)^{s_3 + \sigma_3} \epsilon_1 \epsilon_2 \epsilon_3 \sqrt{2k+1}$$
$$\times \begin{pmatrix} s_1 & s_2 & k \\ \sigma_1 & \sigma_2 & \kappa \end{pmatrix} \begin{pmatrix} s_3 & k & L \\ -\sigma_3 & \kappa & M \end{pmatrix} \tag{32}$$

is the dependence on magnetic quantum numbers,

$$B(j_1 j_2 j_3 s_1 s_2 l_1 l_2; L'kL)$$
$$= (-)^{k+L'+j_3} (2L+1) \sqrt{(2k+1)(2L'+1)}$$
$$\times \sqrt{(2s_1+1)(2s_2+1)(2s_3+1)(2l_1+1)(2l_2+1)(2l_3+1)}$$
$$\times \begin{Bmatrix} j_1 & j_2 & j_3 \\ s_1 & s_2 & k \\ l_1 & l_2 & L' \end{Bmatrix} \begin{Bmatrix} j_3 & k & L' \\ L & l_3 & s_3 \end{Bmatrix} \tag{33}$$

describes the dependence on angular momenta, and

$$C(\Gamma_{p_1} l_1 \nu_1 \hat{m}_1, \Gamma_{p_2} l_2 \nu_2 \hat{m}_2, L', \Gamma_{p_3} l_3 \nu_3 \hat{m}_3, LM)$$
$$= (-)^{L'} \sqrt{(2L'+1)} \sum_{m_1 m_2 m_3 M'} x^{\nu_1}_{l_1 m_1} x^{\nu_2}_{l_2 m_2} x^{\nu_3}_{l_3 m_3}$$
$$\times \begin{pmatrix} l_1 & l_2 & L' \\ m_1 & m_2 & M' \end{pmatrix} \begin{pmatrix} L' & l_3 & L \\ -M' & m_3 & M \end{pmatrix}, \tag{34}$$

in which Γ_p are the parent representations of the Γ_i and x^{ν}_{lm} their coefficients, is the only part which does not involve only geometrical coefficients of SU(2). In (32) and (33), the triangular relations give $|s_1 - s_2| \leqslant k \leqslant s_1 + s_2$ and $|k - s_3| \leqslant L \leqslant k + s_3$, so that the largest value of L is 4 and occurs only for two Γ_8 and a Γ_4 or a Γ_5 representation. The values of L are also limited by (34), which are the coefficients of a representation of the cubic group for the angular momentum L: as the product of the parent representations can be only Γ_1, Γ_2, or Γ_3 representations, the possible values of L are 0, 2, 3, and 4. Combining those two limitations, there appears only one set of values (k,L) in the sum (31) except for a cubic $3jm$ symbol dealing with two Γ_8, Γ_8', or Γ_8'' and one Γ_4 or Γ_5 representation, in which cases there are two sets (k,L). When the product of the three parent states includes a Γ_1 representation, i.e., when there is a cubic $3jm$ symbol between the parent representations, the first term is $(k = s_3, L = 0)$ and the second term is $(k = s_1 + s_2, L = 4)$ if no parent representation is Γ_3, and $(k = s_1 + s_2, L = 2)$ otherwise. However, for $\Gamma_3\,\Gamma_8\,\Gamma_8''$ and $\Gamma_3\,\Gamma_8'\Gamma_8''$, only $L = 2$ exists because $L = 0$ is not allowed with the values 0, $\frac{1}{2}$, and $\frac{3}{2}$ for s. With the five other combinations of parent representations for which there are no cubic $3jm$ symbols, $L = 3$ if there is no Γ_3 representation among them and $L = 2$ otherwise; the values of k are $k = L - s_3$ for the case $L = 3$ and $k = s_1 + s_2$ for the case $L = 2$ which occurs only if $s_1 + s_2 + s_3 = L + 1$.

For $L = 0$ the relation (31) reduces to

$$\begin{pmatrix} \Gamma_{i_1} j_1 l_1 \nu_1 & \Gamma_{i_2} j_2 l_2 \nu_2 & \Gamma_{i_3} j_3 l_3 \nu_3 \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} = \epsilon_1 \epsilon_2 \epsilon_3 \sqrt{(2s_1+1)(2s_2+1)(2s_3+1)(2l_1+1)(2l_2+1)(2l_3+1)}$$
$$\times \begin{Bmatrix} j_1 & j_2 & j_3 \\ s_1 & s_2 & s_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} \begin{pmatrix} s_1 & s_2 & s_3 \\ \sigma_1 & \sigma_2 & \sigma_3 \end{pmatrix} \begin{pmatrix} \Gamma_{p_1} l_1 \nu_1 & \Gamma_{p_2} l_2 \nu_2 & \Gamma_{p_3} l_3 \nu_3 \\ \hat{m}_1 & \hat{m}_2 & \hat{m}_3 \end{pmatrix}. \tag{35}$$

When all the Γ_p are Γ_1 representations and when there is only one 3Γ symbol, this symbol is

$$(\Gamma_{i_1} j_1 l_1 \nu_1 \quad \Gamma_{i_2} j_2 l_2 \nu_2 \quad \Gamma_{i_3} j_3 l_3 \nu_3) = \sqrt{(2s_1+1)(2s_2+1)(2s_3+1)(2l_1+1)(2l_2+1)(2l_3+1)}$$
$$\times \begin{Bmatrix} j_1 & j_2 & j_3 \\ s_1 & s_2 & s_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} (\Gamma_{p_1} l_1 \nu_1 \quad \Gamma_{p_2} l_2 \nu_2 \quad \Gamma_{p_3} l_3 \nu_3). \tag{36}$$

In all other cases, the dependence on s_1,s_2,s_3 is more complicated.

For $L \neq 0$, the expression (34) can be rewritten as

$$C(\Gamma_{p_1} l_1 \nu_1 \hat{m}_1, \Gamma_{p_2} l_2 \nu_2 \hat{m}_2, \Gamma_{p_3} l_3 \nu_3 \hat{m}_3; LM) = \sqrt{2L'+1} \sum_{i',\nu'} \begin{pmatrix} \Gamma_{p_1} l_1 \nu_1 & \Gamma_{p_2} l_2 \nu_2 & \Gamma_{i'} L' \nu' \\ \hat{m}_1 & \hat{m}_2 & \hat{m}' \end{pmatrix} \times \begin{pmatrix} \Gamma_{i'} L' \nu' & \Gamma_{p_3} l_3 \nu_3 & \Gamma_i L \\ \hat{m}' & \hat{m}_3 & \hat{M} \end{pmatrix} x_{L,M}, \tag{37}$$

where Γ_i stands for Γ_3, Γ_2, and Γ_1 when $L = 2,3,4$, respectively, and $x_{L,M}$ are its components. The sum on i' and ν' extends over all kinds of representations which are obtained in the product of Γ_{p_1} and Γ_{p_2} with the angular momentum L'. The sum on i' reduces to only one kind of representation except with Γ_{p_1} and Γ_{p_2} being both Γ_3 representations: this case can be avoided using permutation symmetries except for three Γ_3 parent representations, in which case there is no problem because $L = 0$ is the only term of the sum. However, a careful look at all possible cases shows that the sum over $\Gamma_{i'}$ reduces to only one term, even if Γ_{p_1} and Γ_{p_2} are both Γ_3: in that case, for the $L = 0$ term, $\Gamma_{i'}$ is Γ_{p_3}; for the $L = 2$ term, Γ_{p_3} can be only Γ_1 or Γ_2, and $\Gamma_{i'}$ is Γ_3.

Shifting the components $x_{L,M}$ from C to A, i.e., using $C' = C/x_{L,M}$ and $A' = Ax_{L,M}$, all the dependence on magnetic quantum numbers is gathered in the coefficient A'. This coefficient is proportional to the cubic $3jm$ symbol between basic states or a linear combination of these cubic $3jm$ symbols and those for which the third basic representation has been replaced by the next basic one. So

$$\begin{pmatrix} \Gamma_{i_1} j_1 l_1 \nu_1 & \Gamma_{i_2} j_2 l_2 \nu_2 & \Gamma_{i_3} j_3 l_3 \nu_3 \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} = (\Gamma_{i_1} j_1 l_1 \nu_1 \quad \Gamma_{i_2} j_2 l_2 \nu_2 \quad \Gamma_{i_3} j_3 l_3 \nu_3) \times \begin{pmatrix} \gamma_{i_1} q_{i_1} & \gamma_{i_2} q_{i_2} & \gamma_{i_3} q_{i_3} \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} + (\Gamma_{i_1} j_1 l_1 \nu_1 \quad \Gamma_{i_2} j_2 l_2 \nu_2 \quad \Gamma_{i_3} j_3 l_3 \nu_3)' \times \begin{pmatrix} \gamma_{i_1} q_{i_1} & \gamma_{i_2} q_{i_2} & \gamma_{i_3} q'_{i_3} \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix}. \tag{38}$$

When there is only one 3Γ symbol

$$(\Gamma_{i_1} j_1 l_1 \nu_1 \quad \Gamma_{i_2} j_2 l_2 \nu_2 \quad \Gamma_{i_3} j_3 l_3 \nu_3) = D(i_1, i_2, i_3; kL) \sum_{L'} \sqrt{(2L'+1)} \times B(j_1 j_2 j_3 s_1 s_2 l_1 l_2; L' kL) \times \sum_{\nu'} (\Gamma_{p_1} l_1 \nu_1 \Gamma_{p_2} l_2 \nu_2 \Gamma_{i'} L' \nu') \times (\Gamma_{i'} L' \nu' \Gamma_{p_3} l_3 \nu_3 \Gamma_i L), \tag{39}$$

where

$$D(i_1,i_2,i_3;kL) = A(s_1 s_2 s_3; \sigma_1 \sigma_2 \sigma_3; kLM) x_{L,M} \times \begin{pmatrix} \gamma_{p_1} q_{p_1} & \gamma_{p_2} q_{p_2} & \gamma_{i'} q_{i'} \\ \hat{m}_1 & \hat{m}_2 & \hat{m}_3 \end{pmatrix} \times \begin{pmatrix} \gamma_{i'} q_{i'} & \gamma_{p_3} q_{p_3} & \gamma_i q_i \\ \hat{m}' & \hat{m}_3 & \hat{M} \end{pmatrix} \times \begin{pmatrix} \gamma_{i_1} q_{i_1} & \gamma_{i_2} q_{i_2} & \gamma_{i_3} q_{i_3} \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix}^{-1} \tag{40}$$

does not depend on magnetic quantum numbers but only on the kinds of representation involved. In fact, there is a small dependence on quantum numbers via a phase $(-)^{j+s}$ for Γ_5 and Γ_7 representations. These coefficients are given in Table IV.

When there are two 3Γ symbols, a permutation of the third representation with one of the first two ones must take into account the relation

$$\begin{pmatrix} \gamma_4 1 & \gamma_{8_2^3} & \gamma_{8_2^5} \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \sigma_3 \end{pmatrix} - \frac{\sqrt{14}}{3} \begin{pmatrix} \gamma_{8_2^3} & \gamma_{8_2^3} & \gamma_4 3 \\ \hat{\sigma}_2 & \hat{\sigma}_3 & \hat{\sigma}_1 \end{pmatrix} = 0 \tag{41}$$

or

$$\begin{pmatrix} \gamma_5 2 & \gamma_{8_2^3} & \gamma_{8_2^3} \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} - \frac{3\sqrt{7}}{2\sqrt{2}} \begin{pmatrix} \gamma_5 2 & \gamma_{8_2^3} & \gamma_{8_2^5} \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} + \frac{\sqrt{35}}{2} \begin{pmatrix} \gamma_{8_2^3} & \gamma_{8_2^3} & \gamma_5 3 \\ \hat{\sigma}_2 & \hat{\sigma}_3 & \hat{\sigma}_1 \end{pmatrix} = 0. \tag{42}$$

In the general case, the 3Γ symbols are a sum over the two values of (k,L) with the coefficients listed in Table IV. In many cases one of the two terms vanishes: indeed, using (31) for the basic representations, only the first term (lower L or lower k) remains and, using (31) for two basic representations and the next basic in the third place, the first term vanishes and only the second one remains. Only nonvanishing coefficients are given in Table IV.

In conclusion, the cubic $3jm$ symbols and the 3Γ symbols can be expressed in terms of those of the parent representations when they exist and when there is only one 3Γ symbol. In the other cases, they can be expressed in terms of "$5jm$ symbols" or "5Γ symbols" involving, besides the parent representations, a neighboring $\Gamma_{i'}$, the Γ_1 representation with $L = 4$, or the basic Γ_2 or Γ_3 representations.

IV. ON SOME PROPERTIES OF THE LABELING

When one of the parent representations is the Γ_1 representation for $L = 0$ and $L = 0$ is the only term in (31), the 3Γ symbol of the parent representations reduces to the orthonormalization condition of the two other parent representations

$$\begin{pmatrix} \Gamma_{i_1} j_1 l_1 \nu_1 & \Gamma_{i_2} s_2 0 & \Gamma_{i_3} j_3 l_3 \nu_3 \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} = (-)^{j_1+s_1+l_1+s_2} \epsilon_{i_1 \hat{\sigma}_1} \epsilon_{i_2 \hat{\sigma}_2} \sqrt{(2s_1+1)(2s_3+1)} \times \begin{pmatrix} s_1 & s_2 & s_3 \\ \sigma_1 & \sigma_2 & \sigma_3 \end{pmatrix} \begin{Bmatrix} s_1 & s_2 & s_3 \\ j_3 & l_1 & j_1 \end{Bmatrix} \delta_{l_1 l_3} \delta_{\nu_1 \nu_3}. \tag{43}$$

TABLE IV. Coefficients $D(i_1,i_2,i_3,k,L)$ as defined by formula (39). They should be read columnwise. Example: $D(\Gamma_4,\Gamma_4,\Gamma_5,2,3) = -7\sqrt{3}/\sqrt{5}$.

$i_1i_2i_3$	k	L	i	i'	Coefficient
$\Gamma_1\Gamma_1\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_2\Gamma_2$	0	0	Γ_1	Γ_2	1
$\Gamma_1\Gamma_3\Gamma_3$	0	0	Γ_1	Γ_3	5
$\Gamma_1\Gamma_4\Gamma_4$	1	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_5\Gamma_5$	1	0	Γ_1	Γ_2	$3.7^2/5$
$\Gamma_1\Gamma_6\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_7\Gamma_7$	$\frac{1}{2}$	0	Γ_1	Γ_2	$7^2/3$
$\Gamma_1\Gamma_8\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_8\Gamma_8'$	$\frac{3}{2}$	3	Γ_2	Γ_1	-7^2
$\Gamma_1\Gamma_8\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_1	$5^2/2$
$\Gamma_1\Gamma_8'\Gamma_8$	$\frac{3}{2}$	3	Γ_2	Γ_2	7^3
$\Gamma_1\Gamma_8'\Gamma_8'$	$\frac{3}{2}$	0	Γ_1	Γ_2	7^2
$\Gamma_1\Gamma_8'\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_2	5.7^2
$\Gamma_1\Gamma_8''\Gamma_8$	$\frac{1}{2}$	2	Γ_3	Γ_3	$-5^3/2$
$\Gamma_1\Gamma_8''\Gamma_8'$	$\frac{1}{2}$	2	Γ_3	Γ_3	$5^2.7$
$\Gamma_1\Gamma_8''\Gamma_8''$	$\frac{1}{2}$	0	Γ_1	Γ_3	$5^2/2$
$\Gamma_2\Gamma_2\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_3\Gamma_3$	0	0	Γ_1	Γ_3	5
$\Gamma_2\Gamma_4\Gamma_5$	1	0	Γ_1	Γ_2	-7
$\Gamma_2\Gamma_5\Gamma_4$	1	0	Γ_1	Γ_1	-1
$\Gamma_2\Gamma_6\Gamma_7$	$\frac{1}{2}$	0	Γ_1	Γ_2	7
$\Gamma_2\Gamma_7\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	-1
$\Gamma_2\Gamma_8\Gamma_8$	$\frac{3}{2}$	3	Γ_2	Γ_2	-7^2
$\Gamma_2\Gamma_8\Gamma_8'$	$\frac{3}{2}$	0	Γ_1	Γ_2	-7
$\Gamma_2\Gamma_8\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_2	-5.7
$\Gamma_2\Gamma_8'\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_8'\Gamma_8'$	$\frac{3}{2}$	3	Γ_2	Γ_1	-7^2
$\Gamma_2\Gamma_8'\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_1	$5^2/2$
$\Gamma_2\Gamma_8''\Gamma_8$	$\frac{1}{2}$	2	Γ_3	Γ_3	-5^2
$\Gamma_2\Gamma_8''\Gamma_8'$	$\frac{1}{2}$	2	Γ_3	Γ_3	$2.5.7$
$\Gamma_2\Gamma_8''\Gamma_8''$	$\frac{1}{2}$	0	Γ_1	Γ_3	5
$\Gamma_3\Gamma_3\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_3\Gamma_3\Gamma_2$	0	0	Γ_1	Γ_2	7
$\Gamma_3\Gamma_3\Gamma_3$	0	0	Γ_1	Γ_3	5
$\Gamma_3\Gamma_4\Gamma_4$	1	2	Γ_3	Γ_3	5^2
$\Gamma_3\Gamma_4\Gamma_5$	1	2	Γ_3	Γ_3	$2.5.7$
$\Gamma_3\Gamma_5\Gamma_4$	1	2	Γ_3	Γ_3	$-2.5.7$
$\Gamma_3\Gamma_5\Gamma_5$	1	2	Γ_3	Γ_3	$2^2.3.7$
$\Gamma_3\Gamma_6\Gamma_8$	$\frac{1}{2}$	2	Γ_3	Γ_3	-5^2
$\Gamma_3\Gamma_6\Gamma_8'$	$\frac{1}{2}$	2	Γ_3	Γ_3	$2.5.7$
$\Gamma_3\Gamma_6\Gamma_8''$	$\frac{1}{2}$	0	Γ_1	Γ_3	5
$\Gamma_3\Gamma_7\Gamma_8$	$\frac{1}{2}$	2	Γ_3	Γ_3	-2.5^2
$\Gamma_3\Gamma_7\Gamma_8'$	$\frac{1}{2}$	2	Γ_3	Γ_3	$2^2.5.7$
$\Gamma_3\Gamma_7\Gamma_8''$	$\frac{1}{2}$	0	Γ_1	Γ_3	2.5
$\Gamma_3\Gamma_8\Gamma_6$	$\frac{3}{2}$	2	Γ_3	Γ_3	-5^2
$\Gamma_3\Gamma_8\Gamma_7$	$\frac{3}{2}$	2	Γ_3	Γ_3	-2.5^2
$\Gamma_3\Gamma_8\Gamma_8$	$\frac{3}{2}$	2	Γ_3	Γ_3	5^2
$\Gamma_3\Gamma_8\Gamma_8'$	$\frac{3}{2}$	2	Γ_3	Γ_3	$2.5.7$
$\Gamma_3\Gamma_8\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_3	$5^2.7/2^2$
$\Gamma_3\Gamma_8'\Gamma_6$	$\frac{3}{2}$	2	Γ_3	Γ_3	$2.5.7$
$\Gamma_3\Gamma_8'\Gamma_7$	$\frac{3}{2}$	2	Γ_3	Γ_3	$2^2.5.7$
$\Gamma_3\Gamma_8'\Gamma_8$	$\frac{3}{2}$	2	Γ_3	Γ_3	$2.5.7$
$\Gamma_3\Gamma_8'\Gamma_8'$	$\frac{3}{2}$	2	Γ_3	Γ_3	$2^2.7^2$
$\Gamma_3\Gamma_8'\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_3	$5.7^2/2$
$\Gamma_3\Gamma_8''\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	-1
$\Gamma_3\Gamma_8''\Gamma_7$	$\frac{1}{2}$	0	Γ_1	Γ_2	2.7
$\Gamma_3\Gamma_8''\Gamma_8$	$\frac{1}{2}$	2	Γ_3	Γ_3	$-5^2.7/2^2$
$\Gamma_3\Gamma_8''\Gamma_8'$	$\frac{1}{2}$	2	Γ_3	Γ_3	$5.7^2/2$
$\Gamma_3\Gamma_8''\Gamma_8''$	$\frac{1}{2}$	0	Γ_1	Γ_3	$5.7/2^2$
$\Gamma_4\Gamma_4\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_4\Gamma_4\Gamma_3$	2	2	Γ_3	Γ_1	5

TABLE IV. (*Continued.*)

$i_1i_2i_3$	k	L	i	i'	Coefficient
$\Gamma_4\Gamma_4\Gamma_4$	1	0	Γ_1	Γ_1	1
$\Gamma_4\Gamma_4\Gamma_5$	2	3	Γ_2	Γ_1	$-3.7^2/5$
$\Gamma_4\Gamma_5\Gamma_2$	0	0	Γ_1	Γ_2	-7
$\Gamma_4\Gamma_5\Gamma_3$	2	2	Γ_3	Γ_2	-2.7^2
$\Gamma_4\Gamma_5\Gamma_4$	2	3	Γ_2	Γ_2	$-3.7^3/5$
$\Gamma_4\Gamma_5\Gamma_5$	1	0	Γ_1	Γ_2	$-7^2/5$
$\Gamma_4\Gamma_6\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_4\Gamma_6\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_4\Gamma_6\Gamma_8'$	$\frac{3}{2}$	3	Γ_2	Γ_1	-7^2
$\Gamma_4\Gamma_6\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_1	$5^2/2$
$\Gamma_4\Gamma_7\Gamma_7$	$\frac{1}{2}$	0	Γ_1	Γ_2	$-5.7/3^2$
$\Gamma_4\Gamma_7\Gamma_8$	$\frac{3}{2}$	3	Γ_2	Γ_2	$-7^3/3$
$\Gamma_4\Gamma_7\Gamma_8'$	$\frac{3}{2}$	0	Γ_1	Γ_2	$-7^2/3$
$\Gamma_4\Gamma_7\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_2	$-5.7^2/3$
$\Gamma_4\Gamma_8\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_4\Gamma_8\Gamma_7$	$\frac{5}{2}$	3	Γ_2	Γ_1	$7^2/3$
$\Gamma_4\Gamma_8\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_4\Gamma_8\Gamma_8(2)$	$\frac{5}{2}$	4	Γ_1	Γ_1	-2.3
$\Gamma_4\Gamma_8\Gamma_8'$	$\frac{3}{2}$	3	Γ_2	Γ_1	-7^2
$\Gamma_4\Gamma_8\Gamma_8'(2)$	$\frac{5}{2}$	3	Γ_2	Γ_1	$2.7^2/3$
$\Gamma_4\Gamma_8\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_1	$5^2/2$
$\Gamma_4\Gamma_8\Gamma_8''(2)$	$\frac{5}{2}$	2	Γ_3	Γ_1	$5^2/3$
$\Gamma_4\Gamma_8'\Gamma_6$	$\frac{5}{2}$	3	Γ_2	Γ_2	7^3
$\Gamma_4\Gamma_8'\Gamma_7$	$\frac{1}{2}$	0	Γ_1	Γ_2	$-7^2/3$
$\Gamma_4\Gamma_8'\Gamma_8$	$\frac{3}{2}$	3	Γ_2	Γ_2	$-3^2.7^3/5^2$
	$\frac{5}{2}$	3	Γ_2	Γ_2	$2^4.7^3/5^2$
$\Gamma_4\Gamma_8'\Gamma_8(2)$	$\frac{3}{2}$	3	Γ_2	Γ_2	$-2^5.7^3/(3.5^2)$
	$\frac{5}{2}$	3	Γ_2	Γ_2	$-2.3.7^3/5^2$
$\Gamma_4\Gamma_8'\Gamma_8'$	$\frac{3}{2}$	0	Γ_1	Γ_2	$-3^2.7^2/5^2$
	$\frac{5}{2}$	4	Γ_1	Γ_2	$2^4.3^2.7^2/5^2$
$\Gamma_4\Gamma_8'\Gamma_8'(2)$	$\frac{3}{2}$	0	Γ_1	Γ_2	$-2^5.7^2/(3.5^2)$
	$\frac{5}{2}$	4	Γ_1	Γ_2	$-2.3^3.7^2/5^2$
$\Gamma_4\Gamma_8'\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_2	$-3^2.7^2/5$
	$\frac{5}{2}$	2	Γ_3	Γ_2	$-2^4.7^2/5$
$\Gamma_4\Gamma_8'\Gamma_8''(2)$	$\frac{3}{2}$	2	Γ_3	Γ_2	$-2^5.7^2/(3.5)$
	$\frac{5}{2}$	2	Γ_3	Γ_2	$2.3.7^2/5$
$\Gamma_4\Gamma_8''\Gamma_6$	$\frac{5}{2}$	2	Γ_3	Γ_3	$-5^3/2$
$\Gamma_4\Gamma_8''\Gamma_7$	$\frac{3}{2}$	2	Γ_3	Γ_3	$-5^2.7/3$
$\Gamma_4\Gamma_8''\Gamma_8$	$\frac{1}{2}$	2	Γ_3	Γ_3	$5^2/2$
	$\frac{3}{2}$	2	Γ_3	Γ_3	-2.5^2
$\Gamma_4\Gamma_8''\Gamma_8(2)$	$\frac{1}{2}$	2	Γ_3	Γ_3	$-2^2.5^2/3$
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-5^2/3$
$\Gamma_4\Gamma_8''\Gamma_8'$	$\frac{1}{2}$	2	Γ_3	Γ_3	-5.7
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-2^2.5.7$
$\Gamma_4\Gamma_8''\Gamma_8'(2)$	$\frac{1}{2}$	2	Γ_3	Γ_3	$2^3.5.7/3$
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-2.5.7/3$
$\Gamma_4\Gamma_8''\Gamma_8''$	$\frac{1}{2}$	0	Γ_1	Γ_3	$-5/2$
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-5^2.7/2$
$\Gamma_4\Gamma_8''\Gamma_8''(2)$	$\frac{1}{2}$	0	Γ_1	Γ_3	$2^2.5/3$
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-5^2.7/(2^2.3)$
$\Gamma_5\Gamma_5\Gamma_1$	0	0	Γ_1	Γ_1	$3.7/5$
$\Gamma_5\Gamma_5\Gamma_3$	2	2	Γ_3	Γ_1	-3.5
$\Gamma_5\Gamma_5\Gamma_4$	1	0	Γ_1	Γ_1	$-7/5$
$\Gamma_5\Gamma_5\Gamma_5$	2	3	Γ_2	Γ_1	$-3^2.7^2/5$
$\Gamma_5\Gamma_6\Gamma_7$	$\frac{1}{2}$	0	Γ_1	Γ_2	$7^2/3$
$\Gamma_5\Gamma_6\Gamma_8$	$\frac{3}{2}$	3	Γ_2	Γ_2	$3.7^3/5$
$\Gamma_5\Gamma_6\Gamma_8'$	$\frac{3}{2}$	0	Γ_1	Γ_2	$3.7^2/5$
$\Gamma_5\Gamma_6\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_2	3.7^2
$\Gamma_5\Gamma_7\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	$-7/3$
$\Gamma_5\Gamma_7\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	$1/3$

TABLE IV. (*Continued.*)

$i_1 i_2 i_3$	k	L	i	i'	Coefficient
$\Gamma_5\Gamma_7\Gamma_8'$	$\frac{3}{2}$	3	Γ_2	Γ_1	$-7^2/3$
$\Gamma_5\Gamma_7\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_1	$7^2/(2.3)$
$\Gamma_5\Gamma_8\Gamma_6$	$\frac{5}{2}$	3	Γ_2	Γ_2	$3.7^3/5$
$\Gamma_5\Gamma_8\Gamma_7$	$\frac{1}{2}$	0	Γ_1	Γ_2	$-7/3$
$\Gamma_5\Gamma_8\Gamma_8$	$\frac{3}{2}$	3	Γ_2	Γ_2	$-3.7^3/5^2$
	$\frac{5}{2}$	3	Γ_2	Γ_2	$-2^2.3.7^3/5^2$
$\Gamma_5\Gamma_8\Gamma_8(2)$	$\frac{3}{2}$	3	Γ_2	Γ_2	$2^7.7^2/(3.5^2)$
	$\frac{5}{2}$	3	Γ_2	Γ_2	$-2.3^3.7^2/5^2$
$\Gamma_5\Gamma_8\Gamma_8'$	$\frac{3}{2}$	0	Γ_1	Γ_2	$-3.7^2/5^2$
	$\frac{5}{2}$	4	Γ_1	Γ_2	$-2^2.3^3.7^2/5^2$
$\Gamma_5\Gamma_8\Gamma_8'(2)$	$\frac{3}{2}$	0	Γ_1	Γ_2	$2^7.7/(3.5^2)$
	$\frac{5}{2}$	4	Γ_1	Γ_2	$-2.3^5.7/5^2$
$\Gamma_5\Gamma_8\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_2	$-3.7^2/5$
	$\frac{5}{2}$	2	Γ_3	Γ_2	$2^2.3.7^2/5$
$\Gamma_5\Gamma_8\Gamma_8''(2)$	$\frac{3}{2}$	2	Γ_3	Γ_2	$2^7.7/(3.5)$
	$\frac{5}{2}$	2	Γ_3	Γ_2	$2.3^3.7/5$
$\Gamma_5\Gamma_8'\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	$-3.7/5$
$\Gamma_5\Gamma_8'\Gamma_7$	$\frac{5}{2}$	3	Γ_2	Γ_1	$-7^2/3$
$\Gamma_5\Gamma_8'\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	$3.7/5^2$
	$\frac{5}{2}$	4	Γ_1	Γ_1	$2^2.3^3.7/5^2$
$\Gamma_5\Gamma_8'\Gamma_8(2)$	$\frac{3}{2}$	0	Γ_1	Γ_1	$2^5.3/5^2$
	$\frac{5}{2}$	4	Γ_1	Γ_1	$-2.3/5^2$
$\Gamma_5\Gamma_8'\Gamma_8'$	$\frac{3}{2}$	3	Γ_2	Γ_1	$-3.7^3/5^2$
	$\frac{5}{2}$	3	Γ_2	Γ_1	$-2^2.3.7^3/5^2$
$\Gamma_5\Gamma_8'\Gamma_8'(2)$	$\frac{3}{2}$	3	Γ_2	Γ_1	$-2^5.3.7^2/5^2$
	$\frac{5}{2}$	3	Γ_2	Γ_1	$2.7^2/(3.5^2)$
$\Gamma_5\Gamma_8'\Gamma_8''$	$\frac{3}{2}$	2	Γ_3	Γ_1	$3.7/2$
	$\frac{5}{2}$	2	Γ_3	Γ_1	$-2.3.7$
$\Gamma_5\Gamma_8'\Gamma_8''(2)$	$\frac{3}{2}$	2	Γ_3	Γ_1	$2^4.3$
	$\frac{5}{2}$	2	Γ_3	Γ_1	$1/3$
$\Gamma_5\Gamma_8''\Gamma_6$	$\frac{3}{2}$	2	Γ_3	Γ_3	$3.5.7$
$\Gamma_5\Gamma_8''\Gamma_7$	$\frac{3}{2}$	2	Γ_3	Γ_3	$2.5.7/3$
$\Gamma_5\Gamma_8''\Gamma_8$	$\frac{1}{2}$	2	Γ_3	Γ_3	$3.5.7$
$\Gamma_5\Gamma_8''\Gamma_8(2)$	$\frac{1}{2}$	2	Γ_3	Γ_3	$2^3.5/3$
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-2.5^3/3$
$\Gamma_5\Gamma_8''\Gamma_8'$	$\frac{1}{2}$	2	Γ_3	Γ_3	$-2.3.7^2$
$\Gamma_5\Gamma_8''\Gamma_8'(2)$	$\frac{1}{2}$	2	Γ_3	Γ_3	$-2^4.7/3$
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-2^2.5^2.7/3$
$\Gamma_5\Gamma_8''\Gamma_8''$	$\frac{1}{2}$	0	Γ_1	Γ_3	-3.7
$\Gamma_5\Gamma_8''\Gamma_8''(2)$	$\frac{1}{2}$	0	Γ_1	Γ_3	$-2^3/3$
	$\frac{3}{2}$	2	Γ_3	Γ_3	$-5^3.7/(2.3)$
$\Gamma_6\Gamma_6\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_6\Gamma_6\Gamma_4$	1	0	Γ_1	Γ_1	1
$\Gamma_6\Gamma_7\Gamma_2$	0	0	Γ_1	Γ_2	7
$\Gamma_6\Gamma_7\Gamma_5$	1	0	Γ_1	Γ_2	$7^2/3$
$\Gamma_6\Gamma_8\Gamma_3$	2	2	Γ_3	Γ_1	5
$\Gamma_6\Gamma_8\Gamma_4$	1	0	Γ_1	Γ_1	1
$\Gamma_6\Gamma_8\Gamma_5$	2	3	Γ_2	Γ_1	$-3.7^2/5$
$\Gamma_6\Gamma_8'\Gamma_3$	2	2	Γ_3	Γ_2	2.7^2
$\Gamma_6\Gamma_8'\Gamma_4$	2	3	Γ_2	Γ_2	7^3
$\Gamma_6\Gamma_8'\Gamma_5$	1	0	Γ_1	Γ_2	$3.7^2/5$
$\Gamma_6\Gamma_8''\Gamma_3$	0	0	Γ_1	Γ_3	5
$\Gamma_6\Gamma_8''\Gamma_4$	1	2	Γ_3	Γ_3	$-5^3/2$
$\Gamma_6\Gamma_8''\Gamma_5$	1	2	Γ_3	Γ_3	$3.5.7$
$\Gamma_7\Gamma_7\Gamma_1$	0	0	Γ_1	Γ_1	$7/3$
$\Gamma_7\Gamma_7\Gamma_4$	1	0	Γ_1	Γ_1	$-5/3^2$
$\Gamma_7\Gamma_8\Gamma_3$	2	2	Γ_3	Γ_2	$-2.5.7$
$\Gamma_7\Gamma_8\Gamma_4$	2	3	Γ_2	Γ_2	$7^3/3$
$\Gamma_7\Gamma_8\Gamma_5$	1	0	Γ_1	Γ_2	$7/3$
$\Gamma_7\Gamma_8'\Gamma_3$	2	2	Γ_3	Γ_1	5^2
$\Gamma_7\Gamma_8'\Gamma_4$	1	0	Γ_1	Γ_1	$-7/3$

TABLE IV. (*Continued.*)

$i_1 i_2 i_3$	k	L	i	i'	Coefficient
$\Gamma_7\Gamma_8'\Gamma_5$	2	3	Γ_2	Γ_1	$7^2/3$
$\Gamma_7\Gamma_8''\Gamma_3$	0	0	Γ_1	Γ_3	2.5
$\Gamma_7\Gamma_8''\Gamma_4$	1	2	Γ_3	Γ_3	$5^2.7/3$
$\Gamma_7\Gamma_8''\Gamma_5$	1	2	Γ_3	Γ_3	$-2.5.7/3$
$\Gamma_8\Gamma_8\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_8\Gamma_8\Gamma_2$	3	3	Γ_2	Γ_1	7
$\Gamma_8\Gamma_8\Gamma_3$	2	2	Γ_3	Γ_1	5
$\Gamma_8\Gamma_8\Gamma_4$	1	0	Γ_1	Γ_1	1
$\Gamma_8\Gamma_8\Gamma_4(2)$	3	4	Γ_1	Γ_1	$3^3/7$
$\Gamma_8\Gamma_8\Gamma_5$	2	3	Γ_2	Γ_1	$-3.7^2/5$
$\Gamma_8\Gamma_8\Gamma_5(2)$	3	3	Γ_2	Γ_1	-3.7
$\Gamma_8\Gamma_8'\Gamma_1$	3	3	Γ_2	Γ_2	7^3
$\Gamma_8\Gamma_8'\Gamma_2$	0	0	Γ_1	Γ_2	-7
$\Gamma_8\Gamma_8'\Gamma_3$	2	2	Γ_3	Γ_2	-2.7^2
$\Gamma_8\Gamma_8'\Gamma_4$	2	3	Γ_2	Γ_2	$-7^3/5$
	3	3	Γ_2	Γ_2	$2^2.7^3/5$
$\Gamma_8\Gamma_8'\Gamma_4(2)$	2	3	Γ_2	Γ_2	$2^2.3.7^2/5$
	3	3	Γ_2	Γ_2	$3.7^2/5$
$\Gamma_8\Gamma_8'\Gamma_5$	1	0	Γ_1	Γ_2	$-3.7^2/5^2$
	3	4	Γ_1	Γ_2	$2^2.3^3.7^2/5^2$
$\Gamma_8\Gamma_8'\Gamma_5(2)$	1	0	Γ_1	Γ_2	$2^2.3.7/5$
	3	4	Γ_1	Γ_2	$3^3.7/5$
$\Gamma_8\Gamma_8''\Gamma_1$	2	2	Γ_3	Γ_3	$-5^3/2$
$\Gamma_8\Gamma_8''\Gamma_2$	2	2	Γ_3	Γ_3	-5^2
$\Gamma_8\Gamma_8''\Gamma_3$	2	2	Γ_3	Γ_3	$-5^2.7/2^2$
$\Gamma_8\Gamma_8''\Gamma_4$	1	2	Γ_3	Γ_3	$5^2/2^2$
	2	2	Γ_3	Γ_3	$-3^2.5^2/2^2$
$\Gamma_8\Gamma_8''\Gamma_4(2)$	1	2	Γ_3	Γ_3	$3^3.5^2/(2^2.7)$
	2	2	Γ_3	Γ_3	$3.5^2/(2^2.7)$
$\Gamma_8\Gamma_8''\Gamma_5$	1	2	Γ_3	Γ_3	$-3.5.7/2$
	2	2	Γ_3	Γ_3	$-3.5.7/2$
$\Gamma_8\Gamma_8''\Gamma_5(2)$	1	2	Γ_3	Γ_3	$-3.5^2/2$
	2	2	Γ_3	Γ_3	$3.5^2/2$
$\Gamma_8'\Gamma_8'\Gamma_1$	0	0	Γ_1	Γ_1	7
$\Gamma_8'\Gamma_8'\Gamma_2$	3	3	Γ_2	Γ_1	7^2
$\Gamma_8'\Gamma_8'\Gamma_3$	2	2	Γ_3	Γ_1	-5.7
$\Gamma_8'\Gamma_8'\Gamma_4$	1	0	Γ_1	Γ_1	$-3^2.7/5^2$
	3	4	Γ_1	Γ_1	$-2^4.3^2.7/5^2$
$\Gamma_8'\Gamma_8'\Gamma_4(2)$	1	0	Γ_1	Γ_1	$-2^4.3/5^2$
	3	4	Γ_1	Γ_1	$3^5/5^2$
$\Gamma_8'\Gamma_8'\Gamma_5$	2	3	Γ_2	Γ_1	$-3.7^3/5$
$\Gamma_8'\Gamma_8'\Gamma_5(2)$	3	3	Γ_2	Γ_1	3.7^2
$\Gamma_8'\Gamma_8''\Gamma_1$	2	2	Γ_3	Γ_3	$-5^2.7$
$\Gamma_8'\Gamma_8''\Gamma_2$	2	2	Γ_3	Γ_3	$-2.5.7$
$\Gamma_8'\Gamma_8''\Gamma_3$	2	2	Γ_3	Γ_3	$5.7^2/2$
$\Gamma_8'\Gamma_8''\Gamma_4$	1	2	Γ_3	Γ_3	$-3^2.5.7/2$
	2	2	Γ_3	Γ_3	$5.7/2$
$\Gamma_8'\Gamma_8''\Gamma_4(2)$	1	2	Γ_3	Γ_3	$3.5/2$
	2	2	Γ_3	Γ_3	$3^3.5/2$
$\Gamma_8'\Gamma_8''\Gamma_5$	1	2	Γ_3	Γ_3	-3.7^2
	2	2	Γ_3	Γ_3	-3.7^2
$\Gamma_8'\Gamma_8''\Gamma_5(2)$	1	2	Γ_3	Γ_3	$3.5.7$
	2	2	Γ_3	Γ_3	$-3.5.7$
$\Gamma_8''\Gamma_8''\Gamma_1$	0	0	Γ_1	Γ_1	$5/2$
$\Gamma_8''\Gamma_8''\Gamma_2$	0	0	Γ_1	Γ_2	7
$\Gamma_8''\Gamma_8''\Gamma_3$	0	0	Γ_1	Γ_3	$5.7/2^2$
$\Gamma_8''\Gamma_8''\Gamma_4$	1	0	Γ_1	Γ_1	$-1/2$
	1	2	Γ_3	Γ_3	$5^2.7/2$
$\Gamma_8''\Gamma_8''\Gamma_4(2)$	1	0	Γ_1	Γ_1	$2.3/7$
	1	2	Γ_3	Γ_3	$3.5^2/2^3$
$\Gamma_8''\Gamma_8''\Gamma_5$	1	0	Γ_1	Γ_2	$-3.7^2/5$
$\Gamma_8''\Gamma_8''\Gamma_5(2)$	1	2	Γ_3	Γ_3	$3.5^2.7/2^2$

Parent representations of Γ_{i_1} and Γ_{i_3} must be of the same kind. In this formula the dependence on magnetic quantum numbers is gathered in the $3jm$ symbol: component $|\Gamma_5\,\hat{2}\rangle$ behaves like $|\Gamma_4\,\hat{0}\rangle$, $|\Gamma_5 \pm \hat{1}\rangle$ like $|\Gamma_4 \mp \hat{1}\rangle$, and $|\Gamma_7 \pm \hat{\tfrac{3}{2}}\rangle$ like $|\Gamma_6 \mp \hat{\tfrac{1}{2}}\rangle$. Such a formula can be used with $s_2 = \frac{1}{2}$, 1, or $\frac{3}{2}$, i.e., the basic Γ_6, Γ_4, or Γ_8 representations. The main interest is when Γ_{i_2} is the basic Γ_4 representation with $j_2 = s_2 = 1$, which corresponds to a vectorial term in the Hamiltonian.

When this vectorial term is the vector **J** itself, j_1 equals j_3 and the $3jm$ symbol associated to the matrix element is the product of (43) by the reduced matrix element $\langle j_1|\,|\mathbf{J}|\,|j_3\rangle = \delta_{j_1 j_3}\sqrt{j_1(j_1+1)(2j_1+1)}$. If Γ_{i_1} and Γ_{i_3} are both Γ_4 or Γ_6 representations, dividing this matrix element by the one obtained with basic representations $l_1 = l_3 = 0$, we get

$$\alpha = (-1)^{j_1+l_1+s_1+1}\sqrt{(2s_1+1)/(s_1(s_1+1))} \times\sqrt{j_1(j_1+1)(2j_1+1)}\begin{Bmatrix} s_1 & 1 & s_1 \\ j_1 & l_1 & j_1 \end{Bmatrix} \times\delta_{l_1 l_3}\delta_{j_1 j_3}\delta_{\nu_1\nu_3}, \tag{44}$$

which is usually called the "fictitious spin."[12] If Γ_{i_1} and Γ_{i_3} are both Γ_5 or Γ_7 representations, we can divide by the matrix element obtained with the corresponding Γ_4 or Γ_6 basic representations and we get the same result. The $6j$ symbol in (44) has a simple expression and α evaluates to

$$\alpha = \tfrac{1}{2}[\,j_1(j_1+1) + s_1(s_1+1) - l_1(l_1+1)]/(s_1(s_1+1)), \tag{45}$$

which happens to be $(\mathbf{J}.\mathbf{S})/S^2$. Actually, the operator **J** has matrix elements between different kinds of representations. Some of them, namely those with the same kind of parents, such as (Γ_1,Γ_4), (Γ_6,Γ_8), (Γ_2,Γ_5), (Γ_7,Γ_8'), obey the conditions of formula (43); a nondiagonal fictitious spin can be similarly defined, provided $\langle j_1|\,|\mathbf{J}|\,|j_3\rangle$ between basic representations is replaced by a nonvanishing expression. The result looks like formula (44), but a simple formula such as (45) does not hold.

For a Γ_8 representation, the role of the fictitious spin is played by two coefficients, which are usually[12,15,16] chosen as the mean values P and Q of J_z for components $|\Gamma_8\hat{\tfrac{3}{2}}\rangle$ and $|\Gamma_8\hat{\tfrac{1}{2}}\rangle$. The Wigner–Eckart theorem yields

$$\langle\Gamma_8 j_1\, l_1\,\nu_1\,\hat{m}_1|\,J_q\,|\Gamma_8 j_3\, l_3\,\nu_3\,\hat{m}_3\rangle = (-1)^{3/2-\hat{m}_1}\langle j_1|\,|J|\,|j_3\rangle \times\begin{pmatrix}\Gamma_8 j_1 l_1\nu_1 & 1 & \Gamma_8 j_3 l_3\nu_3 \\ -\hat{m}_1 & q & \hat{m}_3\end{pmatrix}. \tag{46}$$

Using results of the previous section, we get

$$\frac{\langle\Gamma_8 j\, l_1\,\nu_1|\mathbf{J}|\Gamma_8 j\, l_3\,\nu_3\rangle}{\langle\frac{3}{2}|\,|\mathbf{J}|\,|\frac{3}{2}\rangle} = \alpha\langle\Gamma_8\tfrac{3}{2}|\Gamma_4\,1|\Gamma_8\tfrac{3}{2}\rangle + \alpha'\langle\Gamma_8\tfrac{3}{2}|\Gamma_4\,3|\Gamma_8\tfrac{3}{2}\rangle = \mathrm{id} + (\sqrt{14}/3)\alpha'\langle\Gamma_8\tfrac{3}{2}|\Gamma_4\,1|\Gamma_8\tfrac{5}{2}\rangle \tag{47}$$

and the correspondence between (α,α') and (P,Q) is

$$\alpha = \tfrac{1}{5}(3P+Q), \quad \alpha' = (\sqrt{7}/(5\sqrt{3}))(3Q-P). \tag{48}$$

The expression for α comes from $L=0$ in (31) and is therefore given by (45). The expression for α' is obtained from (31) as

$$\alpha' = \frac{2^2.3^3.7}{\sqrt{5}}\sqrt{j(j+1)(2j+1)(2l_1+1)(2l_3+1)}\begin{Bmatrix} j & j & 1 \\ \frac{3}{2} & \frac{3}{2} & 3 \\ l_1 & l_3 & 4\end{Bmatrix}(\Gamma_1 l_3\nu_3\Gamma_1 l_1\nu_1\Gamma_1 4)$$
$$= (-1)^{l_1+j+3/2}\sqrt{(2l_1+1)(2l_3+1)} \times\left(\frac{3^3.7}{2^5.5^2}\frac{(l_1+l_3+5)!(l_1+j-\frac{3}{2})!(l_3+j-\frac{3}{2})!(l_1-l_3+4)!(l_3-l_1+4)!}{(l_1+l_3-4)!(l_1+j+\frac{5}{2})!(l_3+j+\frac{5}{2})!(l_1-j+\frac{3}{2})!(l_3-j+\frac{3}{2})!(j-l_1+\frac{3}{2})!(j-l_3\,\frac{3}{2})!}\right)^{1/2} \times(\Gamma_1 l_3\nu_3 \quad \Gamma_1 l_1\nu_1 \quad \Gamma_1 4). \tag{49}$$

α' depends on the cubic 3Γ symbol between the parent representations and the Γ_1 representation for $l=4$; in contrast to α it has no diagonality property. For Γ_8' and Γ_8'' representations, the first coefficient α is not given by an $L=0$ term, and therefore neither α nor α' are simple. However, the term $L=0$ is a combination of P and Q, which is diagonal; these diagonal combinations are $(P+3Q)$ for Γ_8' and $(P-Q)$ for Γ_8''.

Let us recall that there exists a natural labeling for Γ_4 and Γ_5 representations quantized along a quaternary axis, since $|\Gamma_4\,\hat{0}\rangle$ and $|\Gamma_5\,\hat{2}\rangle$ span the whole subspaces $|j\,0\,-\rangle$ and $|j\,2\,-\rangle$ of D_4: $|\Gamma_4\,\hat{0}\rangle$ and $|\Gamma_5\,\hat{2}\rangle$ can be chosen as the base vectors of those subspaces and their labels will be those of these vectors with respect to D_4. In case of degeneracy, this orthonormal set differs from that built in this paper, which has the advantage of diagonalizing the fictitious spin.

For example, the multiplet $j=5$ contains 2 Γ_4, and the reduced matrix elements of **J** within the natural set are

$$\langle\Gamma_4\,5\,0\,-|\,|J|\,|\Gamma_4\,5\,0\,-\rangle = \tfrac{15}{8},$$
$$\langle\Gamma_4\,5\,4\,-|\,|J|\,|\Gamma_4\,5\,0\,-\rangle = 3\sqrt{35}/8, \tag{50}$$
$$\langle\Gamma_4\,5\,4\,-|\,|J|\,|\Gamma_4\,5\,4\,-\rangle = -\tfrac{11}{8}.$$

Within the orthonormal set built from the Γ_1 representations for $l=4$ and $l=6$, the reduced matrix elements are

$$\langle\Gamma_4\,5\,4|\,|J|\,|\Gamma_4\,5\,4\rangle = 3,$$
$$\langle\Gamma_4\,5\,6|\,|J|\,|\Gamma_4\,5\,6\rangle = -\tfrac{5}{2}, \tag{51}$$

i.e., the eigenvalues of the preceding matrix. We can take another example within 2 Γ_5 of the multiplet $j=6$: the reduced matrix elements of **J** between the natural set are

$$\langle \Gamma_5\, 6\, 2- \,||\, J \,||\, \Gamma_5\, 6\, 2- \rangle = -\tfrac{37}{16},$$
$$\langle \Gamma_5\, 6\, 6\, -\,||\, J \,||\, \Gamma_5\, 6\, 2- \rangle = -3\sqrt{55}/16, \qquad (52)$$
$$\langle \Gamma_5\, 6\, 6\, -\,||\, J \,||\, \Gamma_5\, 6\, 6\, - \rangle = -\tfrac{3}{16},$$

while those within the orthonormal set built from the Γ_2 representations for $l = 6$ and $l = 7$ are the eigenvalues of that matrix

$$\langle \Gamma_5\, 6\, 6| \,||\, J \,||\, \Gamma_5\, 6\, 6\rangle = \tfrac{1}{2},$$
$$\langle \Gamma_5\, 6\, 7| \,||\, J \,||\, \Gamma_5\, 6\, 7\rangle = -3. \qquad (53)$$

When the parent representations $|\Gamma_1\, l\rangle$ or $|\Gamma_2\, l\rangle$ are degenerate, which occurs for the first time for $l = 12$ or $l = 15$, our Γ_4 or Γ_5 will be indexed by l and the supplementary label ν of Γ_1 or Γ_2, and their coefficients will not be as simple as those of the natural set. However, the fictitious spin coefficients for our set will depend only on l, and not on ν.

In Butler's[5] tables, the Γ_4 representations used in (51) are labeled by 1 and 0, respectively, and the Γ_5 representations used in (53) by 0 and 1. There is no relation between these labels and the properties of the representations.

V. QUANTIZATION ALONG A TERNARY AXIS

With a ternary axis of quantization[17] Γ_1 representations are in the $|\hat{0}\, +\rangle$ subspace of D_3 and they are given by formula (6) with $m \equiv 0(3)$ and $a^{\alpha}_{lm} = (-)^{l+m} a^{\alpha}_{l-m}$. Therefore formula (7) may be used to generate orthonormal sets of Γ_6, Γ_4, and Γ_8 representations. As for Γ_2 representations, they belong to space $|\hat{0}\, -\rangle$ and are given by formula (8) with $m \equiv 0(3)$ but with $b^{\beta}_{lm} = (-)^{l+m+1} b^{\beta}_{l-m}$. Therefore orthonormal sets of Γ_7, Γ_5 and of components $\pm\hat{\tfrac{1}{2}}$ of Γ'_8 are obtained with formula (9), with a summation over $m \equiv 0(3)$. The phases $\epsilon_{i\hat{\sigma}}$ are found to be equal to $(-)^{j+s}$ for Γ_5 representations and $(-)^{j+s+\hat{\sigma}-1/2}$ for Γ_7 and components $\pm\hat{\tfrac{1}{2}}$ of Γ'_8 representations. As for components $|\Gamma'_8\ \pm\hat{\tfrac{3}{2}}\rangle$, they both belong to the same subspace $|\hat{\tfrac{3}{2}}\rangle$ of D_3; in this case, the right-hand side of formula (9) for $s = \tfrac{3}{2}$ and $\hat{\sigma} = \pm\tfrac{3}{2}$ is in this subspace $|\hat{\tfrac{3}{2}}\rangle$ but is not a pure $|\Gamma'_8\ \pm\hat{\tfrac{3}{2}}\rangle$: the correct components are

$$|\Gamma'_8\, j\, l\, \beta\ \pm\hat{\tfrac{3}{2}}\rangle$$
$$= 2\sqrt{2l+1}\left\{\mp\frac{1}{3}\sum_{m\equiv 0(3)} (-)^{l-3/2+\mu} b^{\beta}_{lm} \times\begin{pmatrix} l & \tfrac{3}{2} & j \\ m & \pm\tfrac{3}{2} & -\mu \end{pmatrix}|j\mu\rangle + \frac{2\sqrt{2}}{3}\sum_{m\equiv 0(3)} (-)^{l-3/2+\mu} b^{\beta}_{lm} \times\begin{pmatrix} l & \tfrac{3}{2} & j \\ m & \mp\tfrac{3}{2} & -\mu \end{pmatrix}|j\mu\rangle\right\}. \qquad (54)$$

The two components of a Γ_3 representation are defined by

$$|\Gamma_3\, l\, \gamma\ \pm\hat{1}\rangle = \sum_{m\equiv\pm 1(3)} c^{\pm\gamma}_{lm}|lm\rangle, \qquad (55)$$

with $c^{+\gamma}_{lm} = (-)^{l+m} c^{-\gamma}_{l-m}$. The $|\pm\hat{\tfrac{1}{2}}\rangle$ components of a Γ''_8 representation are given by

$$|\Gamma''_8\, j\, l\, \gamma\ \pm\hat{\tfrac{1}{2}}\rangle = \sqrt{2(2l+1)}\sum_{m\equiv\pm 1(3)} (-)^{l+j+\mu} \times c^{\pm\gamma}_{lm}\begin{pmatrix} l & \tfrac{1}{2} & j \\ m & \mp\tfrac{1}{2} & -\mu \end{pmatrix}|j\mu\rangle \qquad (56)$$

and the $|\hat{\tfrac{3}{2}}\rangle$ components by

$$|\Gamma''_8\, j\, l\, \gamma\ \pm\hat{\tfrac{3}{2}}\rangle = \sqrt{2(2l+1)}\left\{-\frac{1}{\sqrt{3}}\sum_{m\equiv\pm 1(3)} (-)^{l+j+\mu} c^{\pm\gamma}_{lm} \times\begin{pmatrix} l & \tfrac{1}{2} & j \\ m & \pm\tfrac{1}{2} & -\mu \end{pmatrix}|j\mu\rangle \pm\frac{\sqrt{2}}{\sqrt{3}}\sum_{m\equiv\mp 1(3)} (-)^{l+j+\mu} c^{\mp\gamma}_{lm} \times\begin{pmatrix} l & \tfrac{1}{2} & j \\ m & \mp\tfrac{1}{2} & -\mu \end{pmatrix}|j\mu\rangle\right\}. \qquad (57)$$

VI. PRESENTATION OF TABLES

The coefficients listed in Tables IV–VIII are expressed as signed square roots of rational numbers but, for legibility, the square root symbol is omitted. Tables V–VIII contain the coefficients of all the parent representations $\Gamma_1, \Gamma_2, \Gamma_3$ up to and including $l = 26$ for Γ_1 and Γ_2, $l = 25$ for Γ_3. Latin letters $a, b, c, \ldots$ are used to label degenerate sets. Coefficients of representations are listed as the signed square root of the quotient of two integers: a numerator, depending on the coefficient, and a denominator, common to all coefficients. Every integer is factorized in prime numbers; in the numerators, every prime factor greater than $2j$ is enclosed in parentheses, which means that it must be put outside the square root symbol. All quaternary components have been gathered in Table V, while ternary components, obtained from the quaternary components by formula (38) of Ref. 17, are split into three tables, one for each kind of parent.

We have an example as follows (Table V):

$$|\Gamma_3\, 16\, c\, \hat{0}\rangle = (1/(2^{10}.3\sqrt{3}))\{-\sqrt{5.7.17.19.23.29}\,|16\ 0\ +\rangle + 2.3\sqrt{2.3.5.13.23.29}\,|16\ 4\ +\rangle - 2.5\sqrt{7.13.29}\,|16\ 8\ +\rangle + 2.37\sqrt{2.5.7.29}\,|16\ 12\ +\rangle + 11\sqrt{7.31}\,|16\ 16\ +\rangle\}, \qquad (58)$$
$$|\Gamma_3\, 16\, c\, \hat{2}\rangle = \frac{1}{2.3}\{\sqrt{29}|16\ 10\ +\rangle + \sqrt{7}|16\ 14\ +\rangle\}.$$

These tables allow all representations $\Gamma_4, \Gamma_5, \Gamma_6, \Gamma_7, \Gamma_8, \Gamma'_8, \Gamma''_8$ to be constructed up to and including $j = 25$. For example, with

$$|\Gamma_1\, 6\rangle = \frac{\sqrt{7}}{4}|6\ 4\rangle - \frac{1}{2\sqrt{2}}|6\ 0\rangle + \frac{\sqrt{7}}{4}|6\ -4\rangle, \qquad (59)$$

we obtain Γ_6 representations for $j = \tfrac{11}{2}$ and $j = \tfrac{13}{2}$. They are

TABLE V. Γ_1, Γ_2, and Γ_3 representations up to $l = 26$ for Γ_1 and Γ_2, $l = 25$ for Γ_3 (quaternary axis). Base vectors are defined by formula (2).[1] The two components of a Γ_3 are listed on two lines, in the order $|\Gamma_3\, \hat{0}\rangle$, then $|\Gamma_3\, \hat{2}\rangle$, like base vectors on top of the table.

l	Γ	Denominator	$[0+\rangle$ $[2+\rangle$	$[4+\rangle$ $[6+\rangle$	$[8+\rangle$ $[10+\rangle$	$[12+\rangle$ $[14+\rangle$	$[16+\rangle$ $[18+\rangle$	$[20+\rangle$ $[22+\rangle$	$[24+\rangle$ $[26+\rangle$
0	1	1	1						
2	3	1	1						
		1	1						
3	2	1	1						
4	1	$2^2.3$	7	5					
	3	$2^2.3$	-5	7					
		1	1						
5	3	1		1					
		1	-1						
6	1	2^3	-1	7					
	2	2^4	-11	5					
	3	2^3	7	1					
		2^4	5	11					
7	2	$2^3.3$	13	11					
	3	1		1					
		$2^3.3$	11	-13					
8	1	$2^6.3$	$3^2.11$	$2^2.7$	5.13				
	3a	2^9	-11.13	$2^2.7.13$	5				
		1	0	1					
	3b	$2^9.3$	$-3^2.5.7$	$-2^2.5.11$	7.11.13				
		1	1	0					
9	1	$2^3.3$		-17	7				
	2	2^4	-3	13					
	3	$2^3.3$		7	17				
		2^4	-13	-3					
10	1	$2^7.3$	-5.13	$2^2.3.11$	11.17				
	2	$2^8.3$	2.13.19	19	$-3.5.17$				
	3a	$2^7.3^3$	11.13.17	$2^2.3.5.17$	5				
		$2^8.3^3$	2.5.13.17	5.17	$3^5.19$				
	3b	$2^4.3^3$	5.11	$-2^2.3.13$	13.17				
		3^3	-1	2.13	0				
11	2	$2^7.3$	2.5.17	3^4	7.19				
	3a	1		1	0				
		2^9	2.5.7	7.17	-17.19				
	3b	1		0	1				
		$2^9.3$	2.17.19	$-3^2.5.19$	-5.7				

TABLE V. (*Continued.*)

l	Γ	Denominator	$[0+\rangle$ $[2+\rangle$	$[4+\rangle$ $[6+\rangle$	$[8+\rangle$ $[10+\rangle$	$[12+\rangle$ $[14+\rangle$	$[16+\rangle$ $[18+\rangle$	$[20+\rangle$ $[22+\rangle$	$[24+\rangle$ $[26+\rangle$
12	1a	$2^7.3.5^5$	$-2.11.17.23$	$3^2.7.13.17.23$	$-2.3.5^2.13.19$ $.23$	$7.11.13.19$			
	1b	$2^{10}.5^5$	$2.3^2.7^3.13.19$	$11.19(47)$	$2.3.5^2.7.11.17$	$3^2.17^3.23$			
	2	$2^8.3$	2.17	$-3.5^2.7$	11.19				
	3a	$2^{10}.3^5$	$-2.5^2.11.13.17$	$3^2.7^3.17$	$2.3.19^3$	$7.11.19.23$			
		$2^8.3^5$	$2.5^2.7.17$	3^9	$5^2.7.11.19$				
	3b	$2^9.3^5$	$2.7.13.19$	$3^2.5^2.11.19$	$2.3.7.11.17$	$-13^2.17.23$			
		3^5	-11.19	0	2.17				
13	1	$2^8.3$		$5^2.19$	$2^3.5$	-11.23			
	2	$2^8.3$	$-2.5.19$	11.17	17.23				
	3a	$2^8.17$		$-3^2.5.11$	$-2^3.11.19$	$-5.19.23$			
		$2^8.3.17$	2.17^3	$5.11.19$	$5.19.23$				
	3b	$2^5.3.17$		19.23	$-2^3.5.23$	$5^2.11$			
		2.17	0	-23	11				
14	1	$2^{11}.3$	$-2.5.7.17$	$3^2.11.13$	$2.3.13.19$	$5.19.23$			
	2	$2^{12}.3$	$17.19.23$	$3.11.23$	5	$-3^2.5.7.13$			
	3a	$2^{14}.3^3.5$	$2.7.11.13.17$	$3^8.5.7^2$	$-2.3^5.5.11.19$	$-11.13.19.23$			
		$2^{17}.3^3.5$	$5.11.17(47)$	$3.5.19.23^2$	$-7^4.11.19.23$	$3^2.7.11.13.19$ $.23$			
	3b	$2^{14}.5^2$	$2.7.13.17.19$	$5^3.11.19$	$2.3.5^3.17^2$	$-13.19^2.23$			
		$2^{17}.5^2$	$5.17.19.23^2$	$-3.5.11(89)$	$23(149)$	$7.13.17^2.23$			
	3c	$3^3.5^2$	19.23	0	0	$2.7.17$			
		$2^{14}.3^3.5^2$	$5.7.13.19.23$	$3.5.7.11.13.17$ $.23$	$7.13.17(37)$	$3^2.17(137)$			
15	1	2^8		23	$-2^3.3.7$	5.13			
	2a	$2^{11}.3.41.43$	$7.11.13^3.23$	$3^8.19.23$	$7.11.19^3$	$3^2.5.11.13.19$ $.29$			
	2b	$2^6.41.43$	$3^2.19.29$	$-7.11.13.29$	$3^2.13.23.29$	$-5.7.23$			
	3a	$2^8.11$		$3.7.23$	$2^3.11^2$	$3.5.7.13$			
		$2^{11}.3.11$	$7.13.19.23$	$-3^2.11.23$	$-5^2.7^3$	$-3^2.5.13.29$			
	3b	$2^3.11$		-5.13	0	23			
		$2^{10}.11$	5.19	$-5.7.11.13$	$-5.13.23$	$7.23.29$			
16	1a	$2^{14}.3.5^2.19$	$7^2.13.17.19.23$	$-2^3.3.7.23^3$	$2^2.5^5.19^2$	$2^3.3^4.13(37)$	$3^2.5.13.29.31$		
	1b	$2^8.3.5^2.19$	$2.3^4.19.31$	$3.7.13.17.31$	0	$-17.23.31$	$2.5.17.23.29$		
	2	$2^{11}.3$	$5^2.19$	$-5.7^2.13$	$-3.7.23$	$3.23.29$			
	3a	$2^{20}.3^6.11$	$5.13^6.17$	$-2^3.3^3.5.7.13$ $.17^2.19$	$-2^2.7^4.11^2.13$ $.19.23$	$2^3.5.11^2.17^2.19$ $.23$	$5^4.19.23.29.31$		
		$2^8.3^5.11$	$3(191)$	$-3.5.7^2.13.19$	$5^2.7.19.23$	$-5^2.19.23.29$			
	3b	$2^{17}.3^6.11$	$-7^4.13.17.19.23$	$-2^3.3^3.7.23(61)$	$2^2.5.11^2(73)$	$-2^3.5^4.11^2.13$	$5.13.29.31(107)$		
		$2^{11}.3^5.11$	$3.5.7^2.13.19.23$	$3.23(131)$	$-5.7^3.13$	$5.7^2.13.29$			
	3c	$2^{20}.3^3$	$-5.7.17.19.23$	$2^3.3^3.5.13.23$	$-2^2.5^2.7.13.29$	$2^3.5.7.29(37)$	$7.11^2.31$		

TABLE V. (*Continued.*)

l	Γ	Denominator	$[0+\rangle$ $[2+\rangle$	$[4+\rangle$ $[6+\rangle$	$[8+\rangle$ $[10+\rangle$	$[12+\rangle$ $[14+\rangle$	$[16+\rangle$ $[18+\rangle$	$[20+\rangle$ $[22+\rangle$	$[24+\rangle$ $[26+\rangle$
		$2^2.3^2$	.29 0	.29 0	29	7			
17	1	$2^{11}.3$		5.23.29	2.13.29	3^2	$-2.3.11.31$		
	2	$2^{12}.3$	$-3.7^2.23$	$3^2.11.13$	5^5	5.29.31			
	3a	$2^{11}.3.5^2.7$ .11.19		$5^5.7^2.23.29$	$-2.11^2.13.19^2$ .29	$3^2.5^4.7^2$	$2.3.11^3.19^2.31$		
		$2^{12}.5^2.7.11$ .19	$-5.23.29(71)$	$-3.5.11.13.29$ $.31^2$	$-3.29(673)$	$3^5.31(67)$			
	3b	$2^3.5^2.7$		0	3.11.31	0	13.29		
		$2^{13}.3.5^2.7$	$-3.5.11.13.23$ .31	$-3^2.5.23^2.31$	$11^3.13.31$	$-11.13.19^2.29$			
	3c	$2^4.11.19$		3^2	0	$-5.23.29$	0		
		$2^{13}.11.19$	3(383)	$-11.13^3.23$	5.23(73)	5.23.29.31			
18	1a	$2^9.3.7^3.13$	$-3^4.17.19.29$	$2.7^3.11.29$	$2^2.23.29(37)$	$-2.3^2.23^3$	3.5.11.17.23.31		
	1b	$2^{15}.7^3.13$	$-11.23(239)$	$2^3.7^3.17.19.23$	$-2^2.11.17.19^3$	$2^3.11^3.17.19.29$	$3.5^3.19.29.31$		
	2a	$2^{17}.61$	$2.3^5.13.23.31$	$2^2.5.7^2.11.31$	$2^2.7^3.13.31$	$-13.19^2.29$	$-5.7.11.13.17$ .29		
	2b	$2^{17}.3.61$	2.3.7.11.23.29	$2^2.3^4.5.7.13.29$	$-2^2.11.29(103)$	7.11.31(47)	$-3^2.5^3.17.31$		
	3a	$2^{16}.3.7.13$ .19.29	$3^2.11.13^2.19.23$ $.29^2$	$2^3.7.17.23(211)$	$2^2.5^2.11.17$ (313)	$2^3.3^6.5^2.11.17$ .29	$-3.5.7^2.13^2.29$ .31		
		$2^{17}.3.7.13$ .19.29	2.3.5.11.17.23 (211)	$-2^2.3^4.13.17^3$	$2^2.5.7.11.17$ (157)	$5^3.11.13^2.17.29$ .31	$3^2.5^2.7.29.31$ (83)		
	3b	$2^{16}.7.19.29$	$5.17.19.29^2.31$	$-2^3.3^4.5.7.11$ .31	$-2^2.3^2.5.17^2.23$ .31	$-2^3.5.23.29.31$	$3^3.7^2.11.17.23$ .29		
		$2^{17}.7.19.29$	$-2.3^5.5^2.11^2.31$	$2^2.5.11.13^3.23$ .31	$2^2.7.23.31(73)$	23.29(83)	5.7.11.17.23.29		
	3c	$2^3.3.13.29$	0	$-5^2.11.23$	$2^7.7$	$3^2.7.29$	0		
		$2^{13}.3.13.29$	$-2.3^5.5.7.23$	$-2^2.3^2.5^2.7.11$ .13	$2^2.5(239)$	$5.7.13^2.29.31$	$-3^2.11.17.29.31$		
19	1	$2^{10}.3$		3.5.31	$-2.3.7.31$	-11.29	2.17.29		
	2a	$2^{15}.3.13.41^3$	$2.5^2.7.11.23.29$ .37	$-2^2.5.17.29.37$ (347)	$2^2.7.11.17.37$ (131)	$3.5^5.13^2.17.31$ .37	$-5.7.31(3011)$		
	2b	$2^{12}.3.13.41^3$	$2.5.17.19^2.31$ (47)	$2^2.7.11.23.31$ (61)	$2^2.5.23.29.31$ (59)	$3^5.7.11.13^2.23$ .29	$11.17.23^3.29.37$		
	3a	$2^{10}.3.7^2.11$ .13		$3.5.11^2.13^2.31$	$2.3.7^5.31$	$-11^3.13^2.29$	$-2.7^4.17.29$		
		$2^{14}.7^2.11.13$	$-2.3.5.11.17.23$ .31	$2^2.3.7.31(137)$	$2^2.3.5.11.29.31$	$5^4.7.13^2.29$	$-3^7.17.29.37$		
	3b	$2^4.7^2$		11.29	0	3.5.31	0		
		$2^{18}.3.7^2$	$2.17.19^2.23.29$	$-2^2.5.7.11.29$	$2^2.3^2.5^2(107)$	$-3.5.7.11.17^2$ .31	$-3^2.5.11.17.31$ .37		
	3c	$2^3.11.13$		0	17.29	0	3.7.31		
		$2^{18}.3.11.13$	$2.3^4.5.7.11.23$ .29	$-2^2.3^2.17.29$	$-2^2.5.7^5.11.17$	$3.13^2.17.31$	$-7.31^3.37$		

TABLE V. (*Continued.*)

l	Γ	Denominator	$[0+\rangle$ $[2+\rangle$	$[4+\rangle$ $[6+\rangle$	$[8+\rangle$ $[10+\rangle$	$[12+\rangle$ $[14+\rangle$	$[16+\rangle$ $[18+\rangle$	$[20+\rangle$ $[22+\rangle$	$[24+\rangle$ $[26+\rangle$
20	1a	$2^{15}.3.13.23$	$2.3^4.7.23.31$	$-2.3^2.5.11.17$ $.19.31$	$2^3.11.17.19.31$	$3^6.17.19.29$	$2.11.19.29^3$	$-5.11.13.29.37$	
	1b	$2^{18}.13.23$	$2.7.11.17.19.23$ $.29$	$2.5.29(211)$	$2^3.3^2.29(43)$	$7^2.11^3.31$	$2.17.31(47)$	$5.13.17.19.31$ $.37$	
	2	2^{16}	$-2.5.23.29$	$2^2.13.17.29$	$2^2.3.5.11.17$	7.31	$-19.31.37$		
	3a	$2^{18}.3.5^2.7$ $.19$	$-2.5^2.7.11.13$ $.17.19.23$	$-2.5.13(887)$	$2^3.3^2.13.23^2$ $.31^2$	$11.13.29.31(67)$	$-2.3^2.13^3.17.29$ $.31$	$3^2.5.17.19.29$ $.31.37$	
		$2^{16}.3.5^2.7$ $.19$	$2.5.13.17.23$ $.29^2$	$2^2.3^2.5^4.7^2.19^2$	$-2^2.3.5.11.13$ $.17^2.29$	$-7.13.17.29.31$	$13.17.19.29.31$ $.37$		
	3b	$2^{29}.5^4.19.31$ $.37$	$-2.5^6.7.11.19$ $.23.29.31^3$	$-2.3^4.5^5.17.29$ $.31(61)$	$2^3.17.29.31$ $(123\,997)$	$-11.17(1427$ $.2543)$	$2.7^2(1553.2459)$	$5.13.19.37(137$ $.839)$	
		$2^9.5^4.19.31$ $.37$	$2.5.23.29.31$ (59)	0	$2^{10}.3.5.7^2.11$ $.17.31$	$-7(67.181)$	$3^6.19.37(41)$		
	3c	$2^{31}.3.5^2.13$ $.31$	$2.5^2.11.29.31$ $.37(313)$	$2.5.7.17^3.19.23$ $.29.31.37$	$2^3.7.17.19.23$ $.29.31.37(47)$	$7^5.11.17.19.23$ $.37$	$-2.7.19.23.37$ (43.67)	$-5.7.13.23$ $(49\,477)$	
		$2^{11}.3.5^2.13$ $.31$	$-2.5.7.19.29.31$ $.37$	0	0	$-19^3.23.37$	$7.23(271)$		
	3d	$2^{31}.3.5^4.7$ $.13.31.37$	$-2.5^6.7.17.19$ $.23.31(1459)$	$2.5^7.11.31$ $(54\,601)$	$-2^3.11.31$ $(1\,288\,037)$	$29(51\,590\,179)$	$2.11.17.29$ $(2\,550\,967)$	$-5.11.13.17^3.19$ $.29.37(163)$	
		$2^{11}.3.5^4.7$ $.13.31.37$	$-2.5.11.17.23$ $.31(67)$	0	$2^{14}.3.5.13^2.29$ $.31$	$7.11.17.29$ (2819)	$11^5.17.19.29.37$		
21	1a	$2^{16}.19.23$		$-2.17.29.37.41$	$2^5.5.13.37.41$	$-3.5^3.31.37.41$	$2^5.3^2.17.31.41$	$13.17.19.31$	
	1b	$2^{14}.3.19.23$		$2.3^4.5.13.29.31$	$2^5.17^3.31$	$-3.13.17$	$2^5.5.13.37$	$-5.7^2.19.37.41$	
	2a	$2^{16}.11.23.41$	$-2.7.17.19.31$ $.37^2$	$2^2.5^2.7.13^3.31$	$2^2.3^7.5^2.7.29$	$5^2.17.29(139)$	$5.7.13.17.19.29$ $.37$		
	2b	$2^{14}.3.11.23$ $.41$	$2.3^2.5^4.13.29$	$2^2.17.19^3.29$	$-2^2.3.13^3.17.19$ $.31$	$3^2.7.13.19^3.31$	$-5.31.37(97)$		
	3a	$2^{17}.5^2.7^2.17$		$2.31(223)$	$-2^5.5.7^2.13.17$ $.29.31$	$-3.5.17.29(167)$	$2^5.7^4.29.37$	$-3^4.13.19.29.37$ $.41$	
		$2^{17}.5^2.7^2.17$	$2.5^2.19.31(151)$	$2^2.13.17.31$ (193)	$-2^2.3.17.29$ (137)	$-5^2.7.29(139)$	$5^5.13.19.29.37$		
	3b	$2^{18}.5^2.7^2.17$ $.23$		$-2.3.13.19.29$ (1087)	$2^5.3^3.5.7^2.11^2$ $.17.19$	$5.13.17.19.31$ (191)	$2^5.3^3.7^2.13.19$ $.31.37$	$-3.31.37.41$ (467)	
		$2^{18}.3.5^2.7^2$ $.17.23$	$-2.3^2.5^2.11^4.13$ $.29$	$2^2.17.19.29$ (2237)	$2^2.3.13.17.19$ $.31(181)$	$-3^2.5^2.7.11^2.13$ $.19.31$	$-5^3.31.37(967)$		
	3c	$2^{18}.3.5.7^2$ $.17.23$		$2.3^2.29.31.37$ (113)	$-2^5.5.7^2.13.17$ $.31.37$	$3.5.17.37(977)$	$2^5.7^2.13^2(431)$	$13.19.41(859)$	
		$2^{18}.5.7^2.17$ $.23$	$-2.3.5^2.11^2.19$ $.29.31.37$	$2^2.3^3.13.17.29$ $.31.37$	$-2^2.17.37(1549)$	$-3.5^2.7.19^2.31^2$ $.37$	$3.5.13.19(139)$		
22	1a	$2^{14}.3.5^2.11$ $.17$	$2.5.11.19.23.31$	$2.7.13.31^3$	$-2^3.3.5^2.13.17$ $.29.31$	$3.7.29(53)$	$2.3^5.7.19.29.37$	$-3.13.19.29.37$ $.41$	
	1b	$2^{19}.3.5^2.11$ $.17$	$-2.3.5.7^3.11.13$ $.29.37$	$2.3.19^3.23.29$ $.37$	$2^3.5^2.7.17.19$ $.23.37$	$13^3.19.23.31.37$	$2.13.23.31(317)$	$7.23.31.41(83)$	
	2a	$2^{20}.3^2.5^2.31$	$2.3.5.7.13.29$	$2.17.19.29.31$	$7.13.17.19.37$	$13.19.37.43$	7.43	$-3^2.5^4.11.31^2$	

TABLE V. (*Continued.*)

1	Γ	Denominator	$[0+\rangle$ $[2+\rangle$	$[4+\rangle$ $[6+\rangle$	$[8+\rangle$ $[10+\rangle$	$[12+\rangle$ $[14+\rangle$	$[16+\rangle$ $[18+\rangle$	$[20+\rangle$ $[22+\rangle$	$[24+\rangle$ $[26+\rangle$
			.31.37.43	.37.43	.43			.41	
	2b	$2^{11}.3^2.5^2.31$	2.3.5.17.19.31	$-2^3.7.11^2.13.31$	$2^4.29(103)$	7.17.23².29	− 13.17.19.29.37	0	
	3a	$2^{19}.3^2.5.7^3$ .31.43	2.3.5.7.11.13 .23.29.31².41	2.3.19.29.41 (7907)	$2^3.7.17.19.41$ (149)	13.19.31.41 (317)	$-2.5^2.7^2.13.29^2$ .31.37.41	$-5^2.7.17^2.31.37$ (61)	
		$2^{20}.3^3.5.7^3$ .31.43	2.3.5.13.29.41 (67.83)	2.7.17.19.29.41 (383)	$-5^4.13.17.19.31$ .41	$-7.13.19.31.41$ (2719)	− 31.37.41 (10 789)	$3^4.7^3.11.31.37$ .43³	
	3b	$2^{14}.3.5.7^3$ .31	2.5.7.23.29.31² .37	$-2.11.13.19.29$ .37	$2^3.3.7.11.13.17$ .19.37	$-3.11.19.31.37$	$2.3^3.5^4.7^2.11$ .31	$3.5^2.7.11.13.31$ .41	
		$2^{19}.5.7^3.31$ (113)	2.5.11.29.37 .19.29.37	2.3.7.11.13.17 .31.37	$3.5^2.11.17.19$ .31.37	$3^3.7.11.17^2.19$	3.11.13.31(569) .43	$3^3.7^3.13.31.41$	
	3c	$2^{18}.3^2.7^3.23$ .31.43	2.3.7.11.17.19 .23³.31²	$-2.3.5^3.13.17$ (1949)	$-2^3.5.7.13.19^2$ .29(337)	5.17.29.31 (11 617)	$-2.5.7^2.17.19$ .23².29.31.37	$5^3.7.13.17.19$ .29.31.37.41	
		$2^{11}.3^3.7^3.23$ .31.43	− 2.3.17.19 (2447)	$2^3.5.7.11^2.13$ (139)	$2^4.5.29.31$ (2237)	$-5^3.7.17.29.31$ (79)	$5.13^3.17.19.29$ .31.37	0	
	3d	$2^{10}.3.7.23$ .31	0	2.3.29(47)	$-2^{11}.3^2.7.17$	− 13.31(127)	0	7.19.31.37.41	
		$2^4.3^2.7.23$ .31	− 3.5.13.19.29	7.11².17.29	2.13.17.31	2.7.13.31	$-2^3.19.31.37$	0	
23	1	$2^{16}.3$		$-2.3.7.29.31$	$2^5.5.17.19$	3.5.11.13.19	$2^5.37$	− 37.41.43	
	2a	$2^{18}.3.17^2.23$ .43	$2.3^2.7.19.29.31$ .37.41	$-2.5.7^5.13.31$ .37.41	− 3.5.11.13.17 .37.41	5.7.17².41(293)	13.19(103.151)	$-7.11.13^5.19.43$	
	2b	$2^{19}.3.17^2.23$ .43	$2.3^2.5^3.11.13$ .29(131)	2.11.19(53.199)	$3^5.7^5.17.19.31$	$7^2.11.13.17^2.19$ .31.37	$5^3.7.11^3.31.37$ .41	5.31.37.41.43 (101)	
	3a	$2^{16}.3^7.5^3.7^2$ .13.17.23.31		$-2.3^5.5^4.7^3.17^2$ .19.29.31.41	$-2^5.5.17.41(97$ .971)	$3.5^3.7^2.11^3.13$ .41(181)	$-2^5.19^3.37.41$ (113.163)	$17^2.19.37.43(59$ .307)	
		$2^{23}.3^7.5^3.7^2$ .13.17.23.31	− 2.3⁴.5.13.19 .29.31.41(113 .167)	$2.3^2.19^2.31.41^3$ (83.103)	$-3.7.11^3.17.41$ .43²(15 269)	$13.17^2.37.41^3$ (17 359)	$3^2.5.7^3.19.37$ (330 689)	$5^3.11.19.37.43$ (195 809)	
	3b	$2^{15}.3^7.5^3.7$ .11.13.17.23		$-2.3^5.5^4.7.17^2$ .29.31.37.43	$2^5.5.17.19.37$ .43³(67)	$-3.5^3.7^2.11.13$ .19.37³.43	$2^5.29^2.43^3(347)$	$11^2.17^2.41$ (33 179)	
		$2^{24}.3^7.5^3.7$ .11.13.17.23	$2.3^4.5.13.19^2$ .29.31.37.43 (241)	$-2.3^2.19.29^2$ .31³.37.43(59)	$-3.7.11^3.17.19$ .37.43(5743)	$-13.17^2.19.43$ (506 213)	$3^2.5.7.41.43$ (492 281)	$5^3.11.31^2.41$ (151.173)	
	3c	$2^{15}.3^3.5.11$ .13.23.31		$2.3^{11}.5^3.7.19$ .31	$2^5.17.29(3467)$	3.11.13.29 (1259)	$-2^5.5.19.29.37$ (107)	$5.11^2.19.29.37$ .41.43	
		$2^{18}.3^3.5.11$ .13.23.31	$2.3^4.5^2.13.19$ .31(101)	$2.3^2.5.7^2.29.31$ (521)	$-3.5.7.11^3.17$ .29(47)	$5.7^2.13.29.37$ (97)	$3^8.7.19.29.37$ .41	− 11.19.29.37.41 .43(71)	
	3d	$2^2.3^4.5^2.7^2$ .17.23.31		0	$-5.11.17.19.29^2$ .37	$2^9.3.5.7^2.13.19$ .37	11(3037)	$2^7.11.17^2.41.43$	
		$2^{24}.3^4.5^2.7^2$ .17.23.31	$2.3^4.5.11^3.13$ .29.31.37(97)	$-2.3^2.11.13^4.19$ .31.37(97)	− 3.7.17.19.37 (101 599)	$11^3.13.17^4.19$ (313)	$-3^2.5^3.7^7.11.41$ (149)	5.41.43(181 .509)	
24	1a	$2^{31}.3^4.7.43$ .47	$2.3^2.11.13.23^3$ .31.37(89)	$-2^3.3^5.5^2.31.37$ .43²(59)	$5^2.17.19.29.37$ (149)	$2^2.5^4.11.13.19$ .29.37(349)	2.5².29(251 941)	$2^2.5.11.29.41$ .43(8629)	$-5^2.11.23.29.41$ .43.47
	1b	$2^{32}.3^6.5^2.11$ .43.47	$2.3^2.5^2.17.23$ .29.31(26 099)	$-2^3.3^5.11.13.17$ .29.31.41².43²	11.13.19(523 .15761)	$-2^2.17.19.29^2$ (570 539)	2.11.13.17.37 (2 421 407)	$-2^2.5.13.17.31^2$ .37.41.43	$5^2.7^2.13.17.23$ $37^3.41.43.47$
	1c	$2^{32}.3^6.5^2.7$	$2.3^2.5^2.13.19$	$2^3.3^5.11.19.23$	11.17.23.31.41	$2^2.13.23.31.41^3$	$-2.11.19.23.31$	$2^2.5.19.23.31$	$5^2.19.31.37.43$

TABLE V. (*Continued.*)

1	Γ	Denominator	$[0+\rangle$ $[2+\rangle$	$[4+\rangle$ $[6+\rangle$	$[8+\rangle$ $[10+\rangle$	$[12+\rangle$ $[14+\rangle$	$[16+\rangle$ $[18+\rangle$	$[20+\rangle$ $[22+\rangle$	$[24+\rangle$ $[26+\rangle$
		.11.43.47	.29.41(53 .2939)	$.29.41.43^2$ (457)	(210 193)	(9511)	.37.41(101 .269)	.37.43(457)	.47(59 341)
	2a	$2^{18}.5.23.43$	$-2.11.29.31(83)$	$2.3^3.19.31(97)$	$3.5.7^2.11.19$ (53)	$-13.37(71)$	37.41(97)	$-5^2.11.23.37.41$ .43	
	2b	$2^{16}.3.5.23$ .43	$2.3^2.19.29.37$ .41	$-2.3.11.29^2.37$ .41	$3.5.11^2.31.37$ .41	$-11.13^3.19.31$ .41	$3^4.11.19^3.31$	$5^2.19.23.31.43$	
	3a	$2^{29}.3^7.5^6.23$ .31	$-2.3^2.5^2.11.13^3$ .19.23.29.31 (883)	$-2^3.3.13^2.19.29$ .31(223)	$5^2.7^2.17.41^2$ (489 799)	$2^2.3^4.11.13.31^2$ (139 939)	$-2.19.37(191$ .267 23)	$2^2.5.7^2.11.19$ .37.41.43 (1069)	$3^2.5^2.11.19.23$ .37.41.43.47 (359)
		$2^{24}.3^7.5^6.23$ .31	$-2.3^4.5^2.11.13^2$ .19.29.31.41^2	$2.3.5^2.17^2.31$ (230 059)	$3^{11}.5^7.11.23^2$ .31^2	13.19.37(79 .4133)	$-3^4.19.23^2.37$ .41(2531)	$7^2.11.13^2.19.23$ .37.41.43(311)	
	3b	$2^{24}.3^{10}.5^6.7$ .23^2.37	$-2.3^2.5^2.23.31$ .37.43 (1 042 091)	$2^3.3.11.13.31$ .37.43 (1 681 103)	$-5^2.11.13.17.19$ .29.37.43 (79 811)	$2^2.3^4.7^2.11^2.19$ .29.37.43 (7043)	$2.7^2.11.13.29$ .37^2.43 (141 307)	$2^2.5.13.19^2.29$ .41(59.73 .257)	$3^2.5^2.13.23.29$ .41.47(251 .463)
		$2^{13}.3^{10}.5^6.7$ .23^2.37	$2.3^4.5^2.7^2.13^3$ .31.37^3.43	$2^3.3.5^2.7^4.11$ .13.19.29.31 .37.43	0	11.29.43.47^2 (91 369)	$3^4.11.13.23^2$ .29^3.41.43 (53)	$2^4.13.23.29.31^2$ .41(1171)	
	3c	$2^{29}.3^4.5^5.7$ .31	$-2.3^2.5^2.11^2.13$ .23.29.31(127)	$2^3.3.11^3.29.31$ (2383)	$5^2.11.17.19$ (35 051)	$-2^2.3^4.7^2.13.19$ .31^2(1531)	$-2.7^2.11.37$ (265 987)	$2^2.5.37.41^3.43$ (53^2)	$3^2.5^2.23.37^3.41$ .43.47
		$2^{24}.3^4.5^5.7$ .31	$2.3^4.5^2.7^2.11^2$ .29.31(53)	$-2.3.5^2.7^2.11$ .19.31(2617)	$3^5.5^5.7^2.19.23^2$ .31^2	11.13.37 (155 891)	$-3^6.11.37.41$ (1823)	$-11^4.17^2.23.37$ .41.43	
	3d	$2^{24}.3^{10}.5^6$ .23^2.31.37	$2.3^2.5^2.11.13$ .23.29.31.37 .41(12 791)	$2^3.3.29.31.37$ .41(8 691 971)	$5^2.17.19.37.41^3$ (53.3329)	$2^2.3^4.11.13.19$ .31^2.37.41 (9151)	$2.11^2.37^2.41^3$ (79.593)	$2^2.5.11.43$ (84 615 281)	$-3^2.5^2.11.23.43$ .47(7 052 987)
		$2^5.3^{10}.5^6$ .23^2.31.37	$-2.3^4.5^2.11.13^2$ .29.31.37^3.41	$-2^3.3.5^2.7^2.11^2$ .19.29^2.31.37 .41	0	13.41(626 113)	$3^4.7^2.23^2(61$ .193)	$2^4.11^3.23.43$ (5701)	
25	1a	$2^{19}.3^3.5$		$2.3^4.5^2.7.31.37$	$3.11.17^2.19.37$	11.19(179)	$2^3.41(67)$	$-5^3.11.41.43$	$-11.23.41.43.47$
	1b	$2^{12}.3^3.5$		0	$-3.37.41.43$	$2^6.41.43$	$2^3.11.19.43$	$-2^6.5^3.19$	19.23.47
	2a	$2^{20}.3.41^2$	$-2.3^7.11.23.31$ .43	2.3.29.43(79)	$5.7.11^3.19.29$ .43	$3^3.7^2.11.29.37$ .43	$7.29.37^3.41$	3.11.23.29.37 .41.47	
	2b	$2^{19}.3.41^2$	$-2.7.11^2.29.31$ .37	2.7.11.23.37 .47^2	$-3.5.19.23.37$ (59)	7.23(659)	$3.11.23.31^2.41$ .43	$-7.13^2.41.43.47$	
	3a	$2^{19}.3.7.61$ .677		$-2.7.11.17^2.19$ .31.37^2	$3.7^2(61.677)$	$5^2.37(5197)$	$2^3.11.19.37.41$ (97)	$-3^2.5.19.31^2.37$ .41.43	$7^2.19.23.37.41$ .43.47
		$2^{20}.3.7.61$ .677	$-2.3.11^2.17^2.19$ .23.29.31	$-2.3^3.11^3.17^2$ .19.23^2	5.7(11 789)	$3.5^2.7^2.19.37$ (331)	$-7^3.11.19.37.41$ .43	$-3^5.19.23.37.41$ .43.47	
	3b	$2^{12}.3^2.5^2.7$ .17.29.61 .677		$-2.7.11.13^2.23$ .31.37.41.43 .47	0	$19.23^3.41.43.47$	$2^7.3^4.11.17^2.23$ .43.47	$-5.23.47(28211)$	$-2^{10}.3^2.5^4.7^2$ .17^2.29^2
		$2^{19}.3^2.5^2.7$ .17.29.61 .677	$2.3^9.5^2.29.31$ .37.41.43.47	$2.3^{11}.5^2.11.23$ .37.41.43.47	$5^3.7.19.23.37$ .41.43.47(53)	$3.7^2.23.41.43$ .47(2837)	$-7.11.23.47$ (14 767)	$-3(107.293.751)$	
	3c	$2^{11}.5^2.7.17$ .47^2		2.7.11.31	0	$3^2.19.37(139)$	$-2^{11}.11.17^2.37$ .41	$-5.37.41.43$ (103)	0
		$2^{22}.3.5^2.7$ .17.47^2	$2.5^2.23.29.31$ (4877)	$-2.5^2.11(61$ .3067)	$3^5.5^3.7.19$ (1381)	$-7^6.37(101)$	$3^7.7.11.37.41$ .43(53)	23.37.41.43.47 (59)	

TABLE V. (*Continued.*)

l	Γ	Denominator	[0+⟩ [2+⟩	[4+⟩ [6+⟩	[8+⟩ [10+⟩	[12+⟩ [14+⟩	[16+⟩ [18+⟩	[20+⟩ [22+⟩	[24+⟩ [26+⟩
	3d	$2^{11}.3^2.5.29$ .47^2		$-2.7.37.41(89)$	0	$7^2.11.17^2.19.31$ .41	$-2^{11}.3^2.23^2.31$	5.11.31.43(151)	0
		$2^{22}.3^2.5.29$ .47^2	$-2.3^3.5^2.11^3.23$ .29.37.41	$2.3^7.5^4.31.37$ .41	5.7.11.19.31.37 .41(79)	$-3.11^3.31.41$ (977)	$7^3.31.43(1201)$	$-3.11.23.31.43$ .47(229)	
26	1a	$2^{23}.3.11.13^2$ .19	$-2.3^3.7.13.23$.$29^3.31$	$2^3.5^9.17^2.31$	$-5^4.7.17.19$ (227)	$2^2.3.5^2.37$ (1069)	$2.5.11.13^2.37$ .41(89)	$2^2.5.7.23.37$.$41^3.43$	3.7.23.37.41.43 .47
	1b	$2^{20}.3.11.13^2$ .19	$-2.3^2.5.13.17^2$ .29.37.41	$2^3.3^5.7.23.37$ .41	$3^3.5.17.19.23$ .31.37.41	$2^2.5.7.23.31.41$ (53)	$2.3.7.11.13^2.23$ .31	$-2^2.3.31.43$ (109)	$5.7^6.31.43.47$
	2a	$2^{24}.3^6.31$	$2^2.3^{10}.29.31^3$	$3^3.5.7.11.17^2$.$19^2.23$	$5.11^2.19.23$ (251)	$2.5.7.11^2.23$.$31^2.37$	$2.3^4.23.37.41$ .43	$-7.19^2.37.41.43$ .47	$-5^2.7.13.17.37$ .41.43.47
	2b	$2^{17}.3^6.31$	0	$-3^3.11.37.41.43$ .47	7.19.37.41.43 .47	$2^3.31^2.41.43.47$	$2^3.3^4.5^3.7.47$	$-5.11^2.23(257)$	$5.7^4.13.17.23$

TABLE VI. Γ_1 representations up to $l = 26$, ternary axis. When a coefficient extends over more than one line, the continuation line starts with a dot symbol.

l	Denominator	[0+⟩ [15+⟩	[3+⟩ [18+⟩	[6+⟩ [21+⟩	[9+⟩ [24+⟩	[12+⟩
0	1	1				
4	3^3	-7	$2^2.5$			
6	3^5	$-2^5.3$	$-2.5.7$	-7.11		
8	3^6	3.11	$-2^3.5.7$	$2^5.13$		
9	3^6		$-2.13.17$	-7.17	$2^3.3.7$	
10	3^9	$2^7.5.13$	$-2^3.3.11$	-3.11^3	$-2.11.17.19$	
12*a*	$3^{11}.5^5$	11.17.23.(149)	$2^2.7.13.17.23.(31)$	$2^5.13.23.(131)$	$2^3.7^3.13.19.23$	2.7.11.13.19.(61)
12*b*	$3^{10}.5^5$	2.7.13.19	$2^3.11.19.(139)$	$-2^4.7.11.13^2.17.19$	$2^2.11.17.(337)$	$7^2.17.23$
13	3^9		$-2^2.5.17.19$	$5^2.13^2$	$2.3^2.5^2.11$	$-2^4.11.23$
14	3^{11}	$-2^6.5.7.17$	2.13.(41)	11.13.19	$-2^2.13.19.23$	$-2^5.5.19.23$
15	3^{12}	$-2.5^2.7.13.29$	$2^7.3.19.23$	$2^5.3.7.11.23$	$2.3^4.19^2$	$5.7^2.13$
16*a*	$3^{13}.5^2.19$	$-3^2.13.17^3.19.23$ $2^4.5.13^3.29.31$	$-2^4.3.5.7.11^2.13.23$	$2^6.3.23.(127)$	$2^4.3^2.5.(157)$	$2^3.13.(1381)$
16*b*	$3^{15}.5^2.19$	$2^3.3^2.7^2.11^2.19.31$ $-2.5.17.23.29.(59)$	$-2^9.3.5.7.17.31$	$2^5.3.13.17.31.(37)$	$-2.3^2.5.13.17^3.23.31$	17.23.31.(307)

TABLE VI. (*Continued.*)

l	Denominator	$[0+\rangle$ $[15+\rangle$	$[3+\rangle$ $[18+\rangle$	$[6+\rangle$ $[21+\rangle$	$[9+\rangle$ $[24+\rangle$	$[12+\rangle$
17	3^{11}	$-2^{10}.31$	$2.3.5.23.29$	$-3.5.11.13.29$	$2^2.3^4.29$	$2^5.(41)$
18*a*	$3^{16}.7^3.13$	$2^7.3.17.19.29.(73)$ $-2^3.3.5^3.11.23.31$ $.(43)$	$-2^2.3^2.5.7.29.(521)$ $2^2.7.11.17.23.31.(73)$	$3^2.7^2.11.13.23.29.37^2$	$2.5.23^3.29.(107)$	$-2^4.3.13^2.23.(89)$
18*b*	$3^{17}.7^3.13$	$-2^9.3.11.23.(101)$ $-2.3.5.17.19.29.31$ $.(113)$	$-2^4.3^2.5^5.7.11.17.19$ $.23$ $-7.19.29^5.31$	$-2^2.3^2.7^2.13.17.19$ $.(79)$	$-2^3.5^3.11.17.19.(73)$	$-2^{10}.3.11.13^2.17.19$ $.29$
19	3^{15}	$2^4.5.7.17.29$	$-2^9.3.5.23.31$ $-2^6.3.17.29.37$	$-2^4.3.7.31$	$2^2.3.5^4.11.31$	$11.29.(89)$
20*a*	$3^{18}.13.23$	$3.7^3.23.31.(229)$ $2^3.3.5.11.19.29.(109)$	$-2^2.3^2.5.11.19.31$ $.(61)$ $2^5.11.29.37.(293)$	$-2^9.3^2.11.13.17.19.31$	$-2^5.5^2.11.17.19.29.31$	$2.3.7^4.13^2.17.19.29$
20*b*	$3^{15}.13.23$	$2.3.7.11.17.19.23.29$ $2^2.3.5.7^2.17.29^2.31$	$2^3.3^2.5.17.29$ $2^4.17.19.31.37$	$-2^6.3^2.13.29.(43)$	$2^4.5^2.(1619)$	$-3.7^2.11.13^4.31$
21*a*	$3^{17}.19.23$	$-2.17.31.37^3.41$	$-2^7.3.5^2.17.29.37.41$ $3^2.5.13.17.19.31.41$	$-2^4.13.29.37.41.(47)$ $2^2.7.13.17.19.31.(61)$	$-2^5.3^2.37.41.(139)$	$-2^4.5^3.7^2.31.37.41$
21*b*	$3^{18}.19.23$	$2^5.5.13.(4943)$	$-2^3.3.5.7^2.11^2.13.29$ $.31$ $-2^4.3^2.19.37.(499)$	$5.17.29.31.(557)$ $2^6.5^3.7.19.37.41$	$-2.3^2.5.13.17.31$ $.(311)$	$-2^6.7^2.13.17$
22*a*	$3^{16}.5^2.11.17$	$2^5.3.5.11.19.23^3.31$ $-2^3.29.37.(1019)$	$2.3.7.19.31.(487)$ $2.7.13.19.29^3.37$	$13.17.31.(2003)$ $-2^2.5^2.13.19.29.37.41$ $.43$	$-2^3.7.11^2.13.17.19^2.29$	$-2^6.7.29.(2069)$
22*b*	$3^{21}.5^2.11.17$	$2^6.3.5.7.11.13.29.37$ $.41^2$ $-2^2.7.13.19.23.31$ $.(739)$	$2^2.3.13.23.29.37^3$ $.(53)$ $-23.31.43^2.(59.67)$	$-2.7.17.19.23.29.37$ $.(251)$ $-2.5^2.7.23.31.41.43$ $.(47)$	$-2^4.11^2.17.19.23.31$ $.37$	$2^5.13.19.23.31.37$ $.(71)$
23	3^{18}	$-2.13.37.(283)$	$-2^{10}.3.5.7.29.31$ $-2^5.3^2.5.7.37.41$	$2^8.19^3.31$ $2^7.11.37.41.43$	$2.3^2.17.19.(71)$	$-5.11.13.19.41^2$
24*a*	$2^9.3^{26}.7.43.47$	$-2.3^2.7^2.11.13.23.31$ $.37^3.(883)$ $2.5.23^2.29.(811.1409)$	$2^4.5^2.31.37.(2439257)$ $2^7.3.5^2.23^2.29.37^2.41$ $.(563)$	$-2^6.5^2.19.29.31.37$ $.(8867)$ $2^3.5^2.11.13^4.29.41.43$ $.(1193)$	$-2.3.5^2.11.17.19.29$ $.37^3.(1951)$ $-5^2.11.23.29.41.43.47$ $.(6311)$	$-5^2.11.13.19.29^3.37$ $.47^2.(89)$
24*b*	$2^{10}.3^{28}.5^2.11$ $.43.47$	$2.3^2.5^2.7^4.17.23.29$ $.31.(3163)$ $2.5^5.11.13.17.37$ $.(503.773)$	$2^4.5^2.11.13.17.29.31$ $.(1\,574\,791)$ $2^7.3.11.13.17.37.41$ $.(479.1669)$	$2^6.11.13.17.19.31$ $.(431.13\,789)$ $2^3.5^4.13.17.37.41.43$ $.(197.373)$	$2.3.5^4.13.19$ $.(38\,221\,921)$ $7^2.13.17.23.37.41.43$ $.47.(107.557)$	$17.19.47^2.(109)$ $.(505\,511)$
24*c*	$2^{10}.3^{28}.5^2.7.11$ $.43.47$	$2.3^2.5^2.7^4.13.19.29$ $.41.(479)$ $-2.5^9.11.13^2.19.23.31$	$-2^4.5^2.11.17^2.19.23$ $.29.41.(42\,391)$ $2^7.3.11.19^3.23.31.37$	$2^6.11.23.41.(2\,180\,183)$ $-2^3.5^4.19.23.31.37.43$	$-2.3.5^4.17.23.31.41$ $.(2\,733\,373)$ $-19.31.37.43.47.(71)$	$13.23.31.41.47^2.(59)$ $.(86\,719)$

TABLE VI. (*Continued.*)

l	Denominator	$[0+\rangle$ $[15+\rangle$	$[3+\rangle$ $[18+\rangle$	$[6+\rangle$ $[21+\rangle$	$[9+\rangle$ $[24+\rangle$	$[12+\rangle$
		.37.41.(521)	.(109.1093)	.(223.593)	.(60 223)	
25*a*	$3^{21}.5$	$2^2.5.(167)$	$-2^2.7^3.11.29.31.37$ $3^3.7.41.43.(239)$	$-2.31^2.37.(311)$ $-2.11.23.41.43.(83)$	$2^5.3^3.11.19.37.47^2$ $-2^5.11.23.41.43.47$	$-2^5.11^3.19.(73)$
25*b*	$3^{21}.5$	$2^3.5.11.19.23^2.41.43$	$-2.7.19.29.31.37.41$ $.43$ $2.3^3.7.11.13^2.19$	$-11.13^2.19.37.41.43$ $2^2.19^3.23.31^2$	$-2^4.3^3.37.41.43$ $-2^6.19.23.47.(97)$	$2^6.41.43.(347)$
26*a*	$3^{24}.11.13^2.19$	$-2^8.3^2.7.13.19^2.23.29$ $.31.47^2$ $-2^2.5.7.37.41.(73)$ $.(571)$	$-2^4.5^4.23.31.41^2$ $.(151)$ $-2.3.5^2.37.41.43$ $.(10\,103)$	$5.7.11.(523.3011)$ $-2^3.5.7.13^2.23^3.37.41$ $.43.47$	$2.3.5.17.19.(157561)$ $-2^7.7.23.37.41.43.47$ $.(173)$	$-2^7.5^2.37.(14867)$
26*b*	$3^{21}.11.13^2.19$	$2^9.3^2.5.13.19^2.29.37$ $.41$ $2.23.31.(238\,439)$	$-2^5.5.7^3.37.41.(359)$ $3.5.7^3.23.31.43.(163)$	$2.11.23.31.37.41$ $.(1613)$ $-2^2.13^2.31.43.47$ $.(1531)$	$-2^2.3.7.17.19.23.31$ $.37.41.(53)$ $2^6.5.7^2.31.43.47$ $.(157)$	$-2^4.5.7.23.31.41$ $.(233)$

TABLE VII. Γ_2 representations up to $l = 26$, ternary axis.

l	Denominator	$[0-\rangle$ $[15-\rangle$	$[3-\rangle$ $[18-\rangle$	$[6-\rangle$ $[21-\rangle$	$[9-\rangle$ $[24-\rangle$	$[12-\rangle$
3	3^2	5	2^2			
6	3^3		2.11	-5		
7	3^6	$-3.7.13$	$2^3.13$	$2^5.11$		
9	3^8	$-2^5.5.11$	$-2.3.7^3$	$-3.5^2.13$	$-2^3.13.17$	
10	3^6		$2^3.19$	$-5^2.19$	2.3.17	
11	3^9	11.13.17	$-2^2.3.5.7.17$	$2^7.3$	$2^9.19$	
12	3^9		$-2^9.17$	$-2^4.7^3$	$-2^2.3^2.19$	11.19.23
13	3^{12}	$2^7.7.13.19$	$2^2.5.11.19$	-11.17^3	$-2.17.(53)$	$-2^4.5^2.17.23$
14	3^9		$-2.19.23$	$5^2.11.23$	$-2^2.3^2.17^2$	$2^5.5.13$
15*a*	$3^9.41.43$	$-2.5.11.13.17.23$ $-2.11.13.19.29$	$2^5.3.7^3.11.23$	$-2^5.3^3.5^2.19.23$	$-2.7^3.11.19$	$5.7.11.13^3.19$
15*b*	$3^{14}.41.43$	$3^2.5.7.17^3.19.29$ $2^2.7^3.23.(257)$	$2^4.3.13.19.29.(79)$	$2^{10}.3.5^2.7.11.13.29$	$2^2.3^2.13.23.29.(53)$	$2.5.23.29.(191)$

TABLE VII. (*Continued*.)

l	Denominator	$[0-\rangle$ $[15-\rangle$	$[3-\rangle$ $[18-\rangle$	$[6-\rangle$ $[21-\rangle$	$[9-\rangle$ $[24-\rangle$	$[12-\rangle$
16	3^{11}	$2.23.29.31$	$2^7.3.5^2.7$	$-2^5.3.5.13$	$-2.3^4.13.23$	$-5.11^2.23$
17	3^{15}	$-2^6.3^2.17.19.23$ $-2^{10}.5.29.31$	$2.3.23.(113)$	$3.11.13.(47)$	$-2^2.3^2.5.(41)$	$-2^5.5.23^2.29$
18*a*	$2.3^{14}.61$	$2.11.13.29.(71)$	$-2^2.3.7.13.23.31$ $3.5.7.11.13.17.29$	$-3.5^3.11.31^3$	$2.3.13.31.(241)$	$-2^4.5.13.17^2.29.31$
18*b*	$2.3^{15}.61$	$2.7.11^2.17^2.31$	$2^2.3.11.23.29.(67)$ $-3.5.17.31.(229)$	$3.5.7.13.29.(97)$	$2.3.7^3.11.29^3$	$2^4.5.7.11.13^2$
19*a*	$3^{18}.13.41^3$	$-3.5.11.19.23.29.37$ $.(1031)$ $-2^9.3.31.37.(761)$	$-2^2.3^2.5^2.7.17.23.29$ $.37.(67)$ $2^{11}.5^3.7.31.(1217)$	$-2^7.3^2.5.17.29.37$ $.(433)$	$2^7.5.7.11.17.29.37$ $.(307)$	$2^5.3.5.7^3.11.17.31.37$ $.(43)$
19*b*	$3^{18}.13.41^3$	$-2^5.3.5^2.7.17.19.31$ $.(353)$ $2^4.3.5.7^3.11.17.23.29$ $.(127)$	$-2^7.3^2.5.11.31.(5507)$ $2^6.11.17.23.29.37$ $.(149)$	$2^4.3^2.7.11.23.31$ $.(3191)$	$-2^2.17^4.23.31.(239)$	$-3.23.29.(12613)$
20	3^{14}	$2^2.5.31^3$	$2^{11}.3.5.29$ $-2^4.3.19.31.37$	$-2^6.3.13.17.29$	$-2^4.3.7^2.17$	$11.13^2.17.31$
21*a*	$3^{17}.11.23.41$	$2^{11}.3.5.11.17.19.23$ $.31$ $-2.5^4.7^3.17.29.31^2$	$-2^7.3.7^3.17.31.(89)$ $-5.7.13.17.19.29.37$ $.(53)$	$-2^4.5^2.7.13.31.(89)$ $-2^2.5^2.13.17.19.29.37$ $.41$	$2^5.5^2.7^3.11^2.29.31$	$-2^4.5.7.29.(727)$
21*b*	$3^{20}.11.23.41$	$-2^7.3.5.7.11.13.23.29$ $.(47)$ $2^5.13.19.31.(67)$	$-2^3.3.5^2.13.19.29$ $.(463)$ $2^4.5^5.31.37.(101)$	$-5^2.17.19.29.(6529)$ $-2^6.7^5.13^2.31.37.41$	$-2.7^2.11^2.13.17.19$ $.(463)$	$-2^6.5.13.17.19.31$ $.(251)$
22*a*	$3^{19}.5^2.31$	$-2^2.11^2.13.19.43$ $.(241)$	$-2^2.3.7.13.29.31.37$ $.43$ $3^2.5^2.7.17^2.43.(53)$	$2.5^2.17.19.29.31.37$ $.43$ $-2.41.(3361)$	$-2^4.3^2.7.17^3.19.37.43$	$2^5.5^2.7.13^3.19.37.43$
22*b*	$3^{19}.5^2.31$	$-2^3.7.17^5.29.37$	$2.3.17.19.31.(2293)$ $-2.3^2.5^2.13.17.19.29$ $.37$	$5^2.7.13.31.(691)$ $2^2.7.13.17.19.29.37$ $.41.43$	$-2^3.3^2.13.29.41^4$	$-2^6.5^2.17.29.(457)$
23*a*	$3^{16}.17^2.23.43$	$3.5^2.7.11.19.23.29.31$ $.37.41$ $2^4.3^2.5.7^5.17^2.19.41$	$-2^3.3^3.13.19.29.31.37$ $.41$ $2^8.13.19.(397)$	$-2^5.3^4.5.7.13.31.37$ $.41$ $2^{10}.3^6.5.7.11.13.19$ $.43$	$-2^4.5.7.13.17.37.41$ $.(103)$	$2^3.3^2.7.11.37.41$ $.(337)$
23*b*	$3^{20}.17^2.23.43$	$2^3.3.5.13.23.29$ $.(9601)$ $-2.11.13.17^2.31.37$ $.(463)$	$2^6.3.5.7.11.29.(1997)$ $2^5.5.7^3.11.31.37.41$ $.(127)$	$-2^{12}.5^2.11.19.(3209)$ $2^7.31.37.41.43.(239)$	$2.11^3.17.19.31.41^2$ $.(59)$	$-5.13.19.31.(24391)$
24*a*	$2.3^{18}.5.23.43$	$2.5.7^4.37.(607)$	$-2^{10}.5^2.29.31.(131)$ $2^7.3^3.17^2.19^2.37.41$	$2^8.19.31.(1861)$ $-2^3.11.37.41.43.(349)$	$-2.3^3.11.17.19.(2011)$ $11^3.23.37.41.43.47$	$-11.13^3.19.(1013)$

TABLE VII. (*Continued.*)

l	Denominator	$[0-\rangle$ $[15-\rangle$	$[3-\rangle$ $[18-\rangle$	$[6-\rangle$ $[21-\rangle$	$[9-\rangle$ $[24-\rangle$	$[12-\rangle$
24*b*	$2.3^{21}.5.23.43$		$-2^{10}.5^4.11.19.29.37.41$	$-2^8.11.37.41.(2999)$	$-2.3^3.13^2.17.31.37.41.47^2$	$-13.31.37.41.(3529)$
		$-2.5.11^3.19.31.41.(151)$	$-2^{11}.3^3.11.19.31.(89)$	$2^3.19.31.43.(3083)$	$13^2.19.23.31.43.47.(71)$	
25*a*	$3^{23}.41^2$	$-2^{10}.3^2.5^2.11.13.23.31.43$	$2^6.7.11.31.43.(2539)$	$-2^5.29.43.(24631)$	$-2^{21}.3.11.19.29.43$	$2^5.11.19.29.37.43.(211)$
		$2^2.5.11^2.29.37.43.(167)$	$-3.5^2.7.17^2.29.37.41.47^2$	$-2.5^2.11^3.23.29^3.37.41$	$2^5.7^2.11.23.29.37.41.47$	
25*b*	$3^{24}.41^2$	$2^5.3^2.5^2.7.13.29.31.37.(59)$	$2.23.29.31.37.(79).(89)$	$7.11.13^2.17^2.23.37.(113)$	$-2^4.3.7.19^3.23.37.(61)$	$-2^8.7.19.23.(59.163)$
		$-2^7.5^3.7.11.23.(1471)$	$-2^5.3.11.23.41.43.(659)$	$-2^6.7.41.43.(2801)$	$-2^{10}.7.41.43.47.(443)$	
26*a*	$3^{24}.31$		$-2^4.11^2.17^2.31.(83)$	$-5.7.11.23.(607)$	$2.3^3.5.17.19.23.(659)$	$-2^7.23.31^2.37.(137)$
		$2^2.5.7.23.37.41.(701)$	$-2.3^3.23.37.41.43.(193)$	$2^3.5.7.19^2.37.41.43.47$	$2^7.5^2.7.37.41.43.47$	
26*b*	$3^{24}.31$		$-2^5.5.7.23.31.37.41.43.47$	$2.5^2.11.17^2.37.41.43.47$	$2^2.3^3.7.17.19.37.41.43.47$	$2^4.5.7.31^2.41.43.47$
		$-2.19^2.43.47.(1259)$	$-3^3.5.7.47.(3259)$	$-2^2.17^2.23.41^2.(61)$	$2^6.5.7^2.23.(2399)$	

TABLE VIII. $|\Gamma_3\,\hat{1}\rangle$ representations up to $l=25$, ternary axis.

l	Denominator	$[1\rangle$ $[-8\rangle$ $[16\rangle$ $[-23\rangle$	$[-2\rangle$ $[10\rangle$ $[-17\rangle$ $[25\rangle$	$[4\rangle$ $[-11\rangle$ $[19\rangle$	$[-5\rangle$ $[13\rangle$ $[-20\rangle$	$[7\rangle$ $[-14\rangle$ $[22\rangle$
2	3	2	1			
4	3^3	2^2	-2^4	-7		
5	3^4	2.3.7	3	2^4	$-2^2.5$	
6	3^4	$-2^2.3$	-3.5	2^5	-2.11	
7	3^6	2.3.11	$-2^5.11$	1	2^7	$-2.7.13$
8*a*	$2^3.3^4$	$-2.11.13$ -5^3	0	-7.13	$2^3.7$	$-2.3^2.5$
8*b*	$2^3.3^7$	$-2.5.7$ $7.11.13$	2^{11}	5.11^3	$2^3.5.11.13$	$-2.7.11.13$
9	3^7	$-2^2.11.13$ $2^5.17$	$-5^2.13$	$2^5.7$	-2.5	2^9

TABLE VIII. (*Continued.*)

l	Denominator	$\|1\rangle$ $\|-8\rangle$ $\|16\rangle$ $\|-23\rangle$	$\|-2\rangle$ $\|10\rangle$ $\|-17\rangle$ $\|25\rangle$	$\|4\rangle$ $\|-11\rangle$ $\|19\rangle$	$\|-5\rangle$ $\|13\rangle$ $\|-20\rangle$	$\|7\rangle$ $\|-14\rangle$ $\|22\rangle$
10*a*	3^{11}	$2^2.5.13.17$ $2^7.5.7^2$	$2.3.5.13.17$ $-11^2.19$	$-2^7.3.5.17$	$2^3.3.11^2.17$	$2.3.5.(41)$
10*b*	3^{11}	$-2.(61)$ $2^6.13.17$	$3.5^2.(31)$ $2.5.13.17.19$	$-2^4.3.13$	$-2^2.3.5.7^2.13$	$2^2.3.13.17$
11*a*	$2^2.3^9$	$-2^2.7.13$ $-5.7.17.19$	$2^4.5^3.7$ $-2^5.17.19$	$3.(89)$ $-2.11.17.19$	$2.3.7.13^2$	$2^3.5.7.17$
11*b*	$2^2.3^{10}$	$-2^2.5.13.17.19$ $-(139)$	$2^4.17.19$ $-2^5.5^3.7$	$-3.5.7.17.19$ $2.5.7^3.11$	$2.3.5.17.19$	$-2^3.11^2.19$
12*a*	3^{14}	$2^3.11.17.(31)$ $-2^3.19.(73)$	$2^5.7^3.17$ $-2^4.7^3.11.19$	$-3.7.17.(29)$ $2^4.7.11.19.23$	$-2^4.3.7.(31)$	$-2^2.19.23^2$
12*b*	3^{14}	$2^6.5^2.7.19$ $7.11.17.23^2$	$2^6.5^2.11.19$ $2^7.5^2.17$	$2^3.3.7^2.11.19$ $2.17.23.(29)$	$2.3.5^2.11.17.19$	$-2^5.7^3.11.17$
13*a*	$3^{12}.17$	$-2^2.5.17.(29)$ $-2^7.11.19$	$-2.17.(263)$ $-5.19.23.(31)$	$2^4.5.11.(43)$ $2^{10}.5.19.23$	$-2^4.5.11.13^2$ $-2^2.5.13.19.23$	$-2.7^3.11.19$
13*b*	$3^{11}.17$	$2.11.17.19.23$ $-2^6.5.7^2.23$	$-5.11.17.19.23$ $2.11.(41)$	$-2^7.19.23$ $-2^9.11$	$2^3.19^3.23$ $-2^3.5^4.11.13$	$2^2.5.7.23$
14*a*	$2^7.3^{10}.5$	$-2^3.3^4.5.11.13.17$ $-2^{10}.3^2.5.11.19$	$5^3.7^2.11.17$ $2^2.3^2.11.19.23$	$-2^9.3.5.13^2$ $2.11.19.23$	$-2.3.7^4.19$ $-2^3.3^2.11.13.19.23$	$-2^4.5.7.19.(29)$ $-3^2.7.11.13.19.23$
14*b*	$2^7.3^{13}.5^2$	$2^3.3^2.5.7^2.13.17.19$ $-2^{10}.3^2.5^3(29)$	$3^2.5^3.17.19.23^2$ $2^2.3^2.23.(953)$	$2^9.3.5^3.7^2.11.19$ $-2.3^2.23.(349)$	$-2.3.11^3.(97)$ $-2^3.13^3.23^3$	$2^4.3^2.5.7.11.(43)$ $7^3.13.23.(83)$
14*c*	$2^4.3^{10}.5^2$	$2^3.3^2.5.7.19.23$ 0	$-5^3.7.13.19.23$ $2^2.3^2.7^3.13.17$	0 $-2.7.13.17.(43)$	$-2.3.7.11.13.17.23$ $2^3.7.17.(47)$	$-2^4.5.11.13.17.23$ $-17.(101)$
15*a*	$3^{12}.11$	$2.3^2.13.17.19.23$ $-2^3.3^2.(41)$	0 $-2^5.3^2.5^2.7$	$-3.7.23^3$ $2^5.3^4.5.7.13$	$-2^5.3.5.7.11.23$ $-2^7.5.13^3$	$-2.3^4.23$ $-2^7.5.13.29$
15*b*	$3^{13}.11$	$2^4.3^2.5.7.17.19$ $3^2.5.7.13.23$	$2^4.3^2.5.11^2.19$ $-2^8.3^2.5.13.23$	$2^3.3.5^3.7^2.13$ $-2^2.3^2.23.29^2$	$2^2.3.5^4.11.13$ $2^{10}.7.23$	$-2^4.3^2.5.7^3.13$ $-2^4.7^3.23.29$
16*a*	$2^6.3^{20}.11$	$2^6.3^2.5.7^2.(1327)$ $-3^2.13.19.23.(2251)$ $-5^2.11^2.17^2.19.23.29$ $.31$	$-2^4.3.7^2.17^2.(37)$ $2^6.3^2.5^2.7.19.23.(73)$	$-2^5.3.5^3.7^3.11^2.13.19$ $2^2.5^2.7.19.23.(457)$	$2^2.3.5.13.19.(7349)$ $2^4.5.13^2.19.23^3.29$	$-2^6.3.7^2.13.19.23.31^2$ $2^4.5^2.7^2.19.23.29.31^2$
16*b*	$2^3.3^{20}.11$	$-2^4.3^2.13.19.23.(53)$ $3^2.5.(44\,971)$ $5^5.11^2.13.29.31$	$2^4.3.5.13.19.23.(43)$ $2^6.3^2.5.7.13.17^2.29^2$	$2^3.3.7.11^2.23.(157)$ $2^2.5.7.13.(5503)$	$2^2.3.23.(3169)$ $-2^4.5^2.13.29.(431)$	$-2^4.3.5^3.(37.71)$ $2^4.5.13.29.(367)$
16*c*	$2^6.3^{17}$	$-2^6.3^2.5.7.19.23.29$ $-3^2.5^2.7.13.17^2.29$ $-7.23^2.31.(83)$	$-2^4.3.5^4.7.19.23.29$ $-2^6.3^2.29.(139)$	$2^5.3.5.13^3.23.29$ $2^2.29.(2213)$	$2^2.3.5.7.13^3.23.29$ $-2^4.5.7.17^2.(79)$	$2^6.3.5^2.7.13.29$ $-2^4.7.(3209)$
17*a*	$3^{14}.5^2.7.11.19$	$-2^3.3.5.19.23.29.(41)$ $-2^6.3.13.29.(631)$ $-2^6.7^2.11.31.(139)$	$3.5^3.11^2.13^2.23.29$ $7^2.29.(1279)$ $-2^4.5^2.7^4.11.17.31$	$2^5.3.5.19^4.23.29$ $2.29.(19\,813)$	$-2.3.5.7^2.11^3.13.23.29$ $2.(47.2129)$	$-2^2.3.5.13.29.(101)$ $31.(37.439)$

TABLE VIII. (*Continued.*)

l	Denominator	$[1\rangle$ $[-8\rangle$ $[16\rangle$ $[-23\rangle$	$[-2\rangle$ $[10\rangle$ $[-17\rangle$ $[25\rangle$	$[4\rangle$ $[-11\rangle$ $[19\rangle$	$[-5\rangle$ $[13\rangle$ $[-20\rangle$	$[7\rangle$ $[-14\rangle$ $[22\rangle$
17*b*	$2^3.3^{16}.5^2.7$	$-2^3.3.5.11.13.19.23.31$ $-2^{12}.3.11.31$ $2^8.7^2.13.17^2.29$	$-3.5^3.11.13.17^2.23.31$ $-2^2.7^2.11.13.31.(53)$ $2^4.5^2.7^4.13.17.29$	$-2^9.3.5.11.13.23.31$ $-2.11.13.31.(157)$	$-2.3.5.7^2.23.31.(53)$ $2^3.11.13.29.31.(71)$	$-2^4.3.5.11.31.(509)$ $11.13.29.(1447)$
17*c*	$2^3.3^{13}.11.19$	$-2^3.3.17^2.19.(59)$ $-2^{10}.3.5^3.13.23$ $-2^8.5.11.23.29.31$	$3.11^2.(971)$ $-2^2.5.23.(503)$ $2^4.5.11.17.23.29.31$	$-2^9.3^7.19^2$ $2.5^3.23.(167)$	$2.3^3.11^3.13$ $-2^3.3^2.5.23^3.29$	$-2^4.3.13.23.(113)$ $-3^2.5.13^2.23.29.31$
18*a*	$2.3^{16}.7.13.19.29$	$-2^4.3.11.23.(4951)$ $2^7.3.5^2.11.17.(509)$ $-2^7.5.29^3.31$	$-2^3.3.5.11^3.13^2.17.23$ $-5.7.11^5.17.(47)$ $-2.5^6.7.29^3.31$	$-2^4.3.7.17.23.(1297)$ $2^4.5^5.7.11.17.19^2.29$	$2^6.3.11^2.17.(569)$ $2^2.5^2.11^3.17.23^2.29.31$	$2.3.5^2.13.17.(59.71)$ $5.11.17.29.31.(751)$
18*b*	$2.3^{13}.7.19.29$	$-2^{12}.3^3.5.17.31$ $-2^7.3^3.5.23.31$ $2^7.11^3.17.23.29$	$2^7.3.5^4.13^2.31$ $-3^4.7.23.31.(67)$ $2.3^4.5.7.11.17.23.29$	$2^4.3.5^3.7.11^3.31$ $-2^4.7.19^2.23.29.31$	$2^{14}.3.5.11.23.31$ $-2^2.5.23.29.(41)$	$-2.3.5^3.11.13.23.31$ $23.29.(37).(43)$
18*c*	$3^{16}.13.29$	$-2.3.5^2.7.17.23.(109)$ $-2^6.3.7^3.(97)$ $2^8.5.7.11.17.29.31$	$-3.5.7.13^2.17^2.23$ $2.5.19^2.(1069)$ $-2^2.11.17.29.31.(61)$	$-2^{13}.3.5^2.11.23$ $-2^5.5.11^2.29.(61)$	$2^3.3.5^2.7.11.(43)$ $2^3.7.29.31.(41)$	$2^2.3.7.11.13.19^2$ $2.5.7.29.31^3$
19*a*	$3^{15}.7^2.11.13$	$2^3.3^4.5.7.11.17.23.31$ $2.3.7.31.(3391)$ $-2.17^5.29$	$2^5.3.5.11.17^3.23.31$ $2^4.3.5.11.29.31.(41)$ $-2^{10}.3.17.29.(107)$	$3.5.31.(151)$ $-2^2.11.29.31.(191)$ $2^2.3.5^2.7^2.17.19.29.37$	$2^2.3.5^2.31.(47)$ $2^2.7.11^3.29.(131)$	$-2^2.3.7.31.(2459)$ $-2^8.5^2.7.29.37^2$
19*b*	$2^3.3^{16}.7^2$	$-2^8.3.7.17.23.29$ $-5.7.11.29.(97)$ $-3.5^3.11.17.29^2.31$	$-2^6.3^2.13^2.17.23.29$ $2^7.5^2.(191)$ $-2^9.5.11.17.31$	$2^7.3^2.11.29.(47)$ $2.3.5.(53.97)$ $-2.5.7^2.11.17.19.31.37$	$2.3^2.5.11.29.(457)$ $2.3.5^3.7.31.(47)$	$-2^{13}.3^2.5.7.11.29$ $-2^7.3.5^5.7.11.31$
19*c*	$2^3.3^{18}.11.13$	$-2^8.3.5.11.23.29.(53)$ $17.29.(61.251)$ $-3.7.31.(9791)$	$-2^6.3^2.5^3.7.11.23.29$ $-2^7.5.7.11.17.(41)$ $2^9.7^3.31.(79)$	$2^9.3^2.5.7.17.23^2.29$ $-2.3.7.11.17.(1223)$ $2.7^7.19.31.37$	$-2.3^2.5^2.7.17.29.(109)$ $-2.3.11^3.17.31.(53)$	$2^9.3^2.17^3.29$ $-2^7.3.17.19^4.31$
20*a*	$3^{15}.5^2.7.19$	$2^3.3^3.5.11.13.17.19.23$ $2^5.3^2.13.(113)$ $-2^5.5^4.13.17.29.31$	$-2^5.3.5.13.17.23.29^2$ $-2^6.3^2.5.11.13.17^2.29$ $2^2.3^2.13.17.29.31.37$	$-2.3^2.5^5.13.(103)$ $-2^6.11.13.29.31.(47)$ $2^6.3^2.17.19.29.31.37$	$2^6.3^2.5.13.(281)$ $2^6.3.11^2.13.29.31$ $-5^3.17.19.29.31.37$	$2^4.7.(359)$ $-2^4.3.7.13.17.29.31$
20*b*	$2^{12}.3^{18}.5^4.19.31.37$	$2^4.3.5.11.19.23.29.31.(167.179)$ $-17.29.31.(4\,228\,141)$ $-5^2.(83.234\,932\,9)$	$2^{12}.3.5.7^2.13^2.23.29.31.(307)$ $2^{13}.5.11.17.31.(61).(1283)$ $2^5.7^2.37.(43.197.257)$	$2^2.3^2.5.17.23^2.29.31.(26\,821)$ $-2.11.17.(50\,545\,553)$ $-2.13.19.37.(382\,727)$	$2.3^2.5.17.29.31.(101).(2393)$ $2.3.17.(1637.687\,37)$ $2^3.5.13^5.19.37.(19\,577)$	$2^5.7.13.17.29.31.(437\,861)$ $-2^{13}.3.7.(2\,199\,151)$
20*c*	$2^{14}.3^{19}.5^2.13.31$	$2^4.3.5^3.7.11.29.31.37.(727)$ $7.13^2.17.19.23.29.31.37.(293)$ $5^2.7.19.23.37.(11\,279)$	$-2^{12}.3.5^3.7.11^2.19.29.31.37$ $2^{13}.5.7.11.17.19.23.31.37.(41)$ $2^5.7^3.19.23.(223^2)$	$-2^2.3^2.5.7.17.19.23.29.31^3.37$ $2.7.11.17.19.23.37.(5077)$ $2.7^5.13.23.(67.463)$	$2.3^2.5.7.17.19.23.29.31.37.(367)$ $-2.3.7^3.13^2.17.19.23.37.(571)$ $2^3.5.7.13.23.(311).(547)$	$2^5.13.17^3.19.23.29.31.37$ $2^{13}.3.19.23.37.(47).(199)$

TABLE VIII. (*Continued.*)

l	Denominator	[1⟩ [−8⟩ [16⟩ [−23⟩	[−2⟩ [10⟩ [−17⟩ [25⟩	[4⟩ [−11⟩ [19⟩	[−5⟩ [13⟩ [−20⟩	[7⟩ [−14⟩ [22⟩
20*d*	$2^{14}.3^{19}.5^4.7.13.31.37$	$-2^4.3.5.17.19.23.31.(814\,399)$ $-11.13^2.31.(607).(21\,929)$ $-5^2.11^5.17.29.(275\,201)$	$2^{12}.3.5.11.17.23.31.(241.859)$ $2^{13}.5.29.31.(226843)$ $2^5.11.17.29.37.(137).(41\,851)$	$-2^2.3^2.5.11.31.(47).(103.11\,923)$ $2.29.(23197.27611)$ $-2.11.13.17.19.29.37.(934\,981)$	$2.3^2.5.11.31.(17\,904\,347)$ $-2.3.11.13^4.29.(659).(1811)$ $-2^3.5.11.13.17.19.29.37.(122\,861)$	$-2^5.7.11.13.31.(1\,034\,381)$ $-2^{13}.3.7.11.17.29.(225\,353)$
21*a*	$2.3^{16}.5^2.7^2.17$	$-2^4.3.5.19.23.31.(73)$ $-2^7.5.13.17.19^2.29.31$ $-2^7.29^5.37$	$-2^3.3.5^2.13^2.19.31.(53)$ $17.29.(43).(673)$ $2.5.19.29.37.(1069)$	$-2^5.3^4.7^2.31.(137)$ $-2^2.11.17^3.23^2.29$ $-2^{10}.7^2.13.19.29.37$	$-2^5.3^4.13.17.31.(101)$ $2^4.3^5.5.7^2.29.(89)$ $2^4.5^2.7^2.13.19.29.37.41$	$-2.5.7.13.17.31.(487)$ $3^5.7.11^2.17^2.29$
21*b*	$2^2.3^{19}.5^2.7^2.17.23$	$-2^2.3.5.13^3.23.29.(1567)$ $-2^{13}.5.17.19.(5279)$ $-2^7.13.19.31.37.(1301)$	$2.3.5^2.13.29.(49\,937)$ $13.17.19.29^2.31.(251)$ $-2.5.13.31.37.(2699)$	$2^7.3^2.7^2.13.19^3.29.37^2$ $-2^2.11.13^5.17.19.29^2.31$ $2^{10}.7^4.31.37.(419)$	$2^3.3^2.17.19.29.(6113)$ $-2^4.3.5.7^2.13.19^3.31$ $-2^4.5^2.7^2.31.37.41.(43)$	$-2.5.7.13^4.17.19.29.(61)$ $3.7.11^2.13^3.17^2.19.23^2.31$
21*c*	$2^2.3^{19}.5.7^2.17.23$	$-2^2.3.5.19.23.29.31.37.(359)$ $2^9.5.13.17.31.37.(89)$ $2^7.(763613)$	$2.3.5^2.19.29.31.37.(61)$ $17.37.(115\,589)$ $-2.5^3.19.(151.701)$	$-2^{11}.3^2.7^2.29.31.37.(43)$ $2^2.11.17.37.(2609)$ $2^{10}.7^2.13.19.(853)$	$2^3.3^2.13.17.29.31.37.(211)$ $2^4.3.5.7^2.19^2.37.(373)$ $2^4.5^2.7^2.13.19.41.(541)$	$2.5.7.13.17.29.31.37.(211)$ $3.7.13^2.17^2.23^2.37$
22*a*	$3^{22}.5.7^3.31.43$	$-2^3.3.5.7.13.29.41.(18\,311)$ $-2^{10}.7.17.19.41^3.(151)$ $2^6.13.31.37.41.(6521)$	$-2.3.5^5.13.29.41.(3209)$ $2^2.11^2.13.17.19.31.41.(1423)$ $-2^2.31.37.41.(1627)$	$2^6.3.19.29.41.(23\,039)$ $-2.11.13.17.19.31.41.(67.109)$ $-2^7.7^2.11^2.31.37.(2857)$	$2^2.19.29.41.(257.661)$ $-2^3.7.13.19.31.41.(27\,073)$ $-2^5.5^2.7.31.37.(89).(229)$	$-2^6.7^3.17.19.29^2.41.43^2$ $-7.13^3.19.31.41.(7019)$ $-2^6.7^5.11.31.37.43.(47)$
22*b*	$2.3^{18}.5.7^3.31$	$2^5.3.5.7.11^3.29^3.37$ $-2^{14}.7.11.13.17.19.37$ $2^6.11.23^2.31.(683)$	$2^3.3.5^3.11.29.37.(269)$ $11^3.17.19.31.37.(67)$ $-2^4.11.13.31.(10\,273)$	$-2^{10}.3.11^3.13.19.29.37$ $-2^3.17.19.31.37.(1103)$ $2.7^2.11.13^3.29^2.31.41$	$-2^4.11.13.19^5.29.37$ $-2.7.11.19.31.37.(757)$ $2^7.5^2.7.11.13.17^2.31.41$	$-2^4.7^3.11^3.13.17.19.37$ $-2^2.7.11.19.31.37.(197)$ $7^3.13.31.41.43.(79)$
22*c*	$2.3^{22}.7^3.23.31.43$	$2^{13}.3.7.17.19^3.(1487)$ $-2^8.5.7.13.29.(255\,617)$ $2^{10}.5.17.19.29.31.37.(47)$	$-2^9.3.17.19.23^2.(10\,141)$ $5.11^2.29^3.31.(1217)$ $-2^4.5.13.17.19.29.31.37.(1571)$	$-2^8.3.5.13.17.(79).(2347)$ $2^3.5.11^3.29.31.(49\,669)$ $2.5.7^2.13.17.19.29.31.37.41.(127)$	$2^{10}.5.13.17.(133\,831)$ $-2.5.7.17.29.31.(69\,899)$ $-2^7.5.7.13.17.19.29.31.37.41$	$2^4.5.7^3.13.23^4.29.43^2$ $-2^2.5.7.17.29.31.(5233)$ $-5^3.7^3.11.13.17.19.29.31.37.41.43$
22*d*	$3^{21}.7.23.31$	$-2^2.3.5.7.13.19.29.(1009)$ $2^5.7.11^2.17.(463)$ $-2^{13}.13.19.31.37.(47)$	$-3.5.11^2.13.19^3.23^2.29$ $2.13.17.31.(2371)$ $-2^3.19.31.37.(2833)$	$-2^5.3.29.(89.229)$ $2^2.5^2.11.13^3.17.31$ $2^2.7^2.19.31.37.41.(79)$	$-2.11^2.29.(35\,221)$ $-2^2.7.13.31.41^2.(631)$ $-2^6.5^2.7.19.31^3.37.41$	$2^5.5^2.7.17.23^2.(419)$ $2.7.13.31.(18\,793)$ $2.7.11.13^2.19.31.37.41.43$

TABLE VIII. (*Continued.*)

l	Denominator	$\|1\rangle$ $\|-8\rangle$ $\|16\rangle$ $\|-23\rangle$	$\|-2\rangle$ $\|10\rangle$ $\|-17\rangle$ $\|25\rangle$	$\|4\rangle$ $\|-11\rangle$ $\|19\rangle$	$\|-5\rangle$ $\|13\rangle$ $\|-20\rangle$	$\|7\rangle$ $\|-14\rangle$ $\|22\rangle$
23*a*	$2^5.3^{28}.5^3.7^2.13$ $.17.23.31$	$2^3.3^2.5.11.13.19.29$ $.31.41.(996\ 109)$ $3.5.17.41.(673)$ $.(2\ 589\ 131)$ $-3.19.37.41$ $.(86\ 106\ 541)$ $-2^4.5.11.19.23.37^3.43$ $.(167.1709)$	$-2^5.3^2.5.7^2.13^3.17^2$ $.19.29.31.41.(1567)$ $2^4.3.7.11.17.41.(53)$ $.(59.648\ 91)$ $-2^4.3.5.7.19.37.41^3$ $.(948\ 061)$	$2^2.7.19^3.29.31.37^2.41$ $.(89\ 083)$ $-2^4.3.7.11.13.41$ $.(72\ 030\ 397)$ $2^2.3.5^2.17^2.19^3.37$ $.(73.182\ 89)$	$2^3.29.31.41$ $.(431\ 877\ 583)$ $2^2.5.13.41.(907)$ $.(1\ 327\ 789)$ $2^3.3.19.37.43.(317)$ $.(108\ 529)$	$2^3.3.5.13^2.17.31.41$ $.(7\ 248\ 113)$ $2^6.7^2.13.31^2.37.41$ $.(665\ 029)$ $2^4.5.11.13^4.19.37.43$ $.(296\ 753)$
23*b*	$2^6.3^{28}.5^3.7.11$ $.13.17.23$	$2^3.3^2.5.11.13.29.31$ $.37.43.(178\ 859)$ $3.5.17.19.37.43.(211)$ $.(26\ 317)$ $-3.41^2.43.(202\ 540\ 087)$ $2^4.5.11.23.41.(337)$ $.(158\ 567)$	$2^5.3^2.5.7^2.13.17^2.23^2$ $.29.31.37.43.(53)$ $2^4.3.7.11.17.19^3.37$ $.43.(26\ 987)$ $-2^4.3.5.7.43.(71)$ $.(464\ 927)$	$-2^2.7.29.31.37.43$ $.(19\ 701\ 901)$ $-2^4.3.7.11.13.19.37$ $.43.(557.3109)$ $-2^2.3.5^2.17^2.19^2.41$ $.43.(247\ 781)$	$2^3.19.29.31.37.43$ $.(2593.2699)$ $-2^2.5.13.19.37.43$ $.(83.179.829)$ $-2^3.3.41.(15877)$ $.(203\ 641)$	$-2^3.3.5.13^2.17.19.31$ $.37.43.(102\ 563)$ $2^6.7^2.13.19.23^2.43$ $.(438\ 049)$ $2^4.5.11.13^2.41$ $.(26\ 968\ 589)$
23*c*	$3^{24}.5.11.13.23$ $.31$	$2.3^2.5^4.11.13.19.31$ $.(1979)$ $-2^2.3.17^3.29.(139)$ $.(647)$ $2^8.3.5.19.29.37$ $.(13\ 679)$ $-2^4.11.19.23.29.37.41$ $.43.(97)$	$-2^5.3^2.13.19.31.(83)$ $.(439)$ $-2^8.3.5.7.11.17.29$ $.(59.109)$ $2^4.3.7.19.29.37.(131)$ $.(367)$	$-5^5.7.19.31.(10529)$ $-2^6.3.5.7.11.13.29$ $.(31\ 013)$ $-2^{10}.3.5.19.29.37.41$ $.(1033)$	$-2^5.5.31.(101111)$ $2^{18}.13.29.(3779)$ $2^3.3.5^3.19.29.37.41$ $.43.(179)$	$2.3.7^2.13^2.17.29.31$ $.(5573)$ $2^{10}.5.13.29.31^2.37$ $.(349)$ $2^{12}.5^2.11.13^2.19.29$ $.37.41.43$
23*d*	$2^6.3^{25}.5^2.7^2.17$ $.23.31$	$2^3.3^2.5.13.29.31.37$ $.(59.4523)$ $-3.5.11.17.19.37$ $.(1\ 465\ 193)$ $-3.11^3.(271.212\ 131)$ $-2^4.5.13^4.23.41^3.43$ $.(757)$	$2^5.3^2.5.7^2.11.13.17^4$ $.29.31.37.(83)$ $2^4.3.7.17.19.37$ $.(1237.5167)$ $2^4.3.5.7.11$ $.(38\ 640\ 971)$	$2^2.7.11.29.31.37.(89)$ $.(3163)$ $2^4.3.7.13.19.37.(73)$ $.(23\ 117)$ $-2^2.3.5^2.11.17^2.41$ $.(947.1093)$	$-2^3.11.19.23^2.29.31$ $.37.(3923)$ $2^2.5.11.13^3.19.37^3$ $.(6779)$ $-2^3.3.11.41.43$ $.(11\ 850\ 907)$	$2^3.3.5.11.17.19.31^3$ $.37.(15\ 511)$ $-2^6.7^2.11.13.19.31^2$ $.(59.701)$ $-2^4.5.41.43.(61)$ $.(311.853)$
24*a*	$2^8.3^{28}.5^6.23.31$	$-2.3^2.7^4.11.13.19.23$ $.29.31(31\ 013)$ $-2^5.3.11^2.17$ $.(167\ 846\ 051)$ $3.5^{10}.19^3.37.(6029)$ $-2^2.11.19.23.37.41.43$ $.47.(185\ 821)$	$-2^5.3^2.11.19.29.31$ $.(1669.2017)$ $2^4.3.5.11$ $.(1\ 084\ 097\ 879)$ $-2^2.3.19.37.41.(107)$ $.(397.509)$	$5^8.19.29.31^3.(101)$ $.(157)$ $2^{12}.3.11.(82\ 479\ 329)$ $-2^2.3.19.37.41.43$ $.(97.255\ 637)$	$-2^7.5^9.11^2.19.31$ $.(15\ 877)$ $2^2.11.13.37.(199)$ $.(2\ 226\ 769)$ $-2.3.5.11.19.37.41.43$ $.(347.113\ 11)$	$-2.3.5^2.(103.113)$ $.(684\ 799)$ $2^6.13.19.37$ $.(140\ 378\ 699)$ $2^4.5^4.11.19.23.37.41$ $.43.(33\ 329)$
24*b*	$2^3.3^{31}.5^6.7.23^2$ $.37$	$2.3^2.11^2.23.31.37.43$ $.(33\ 347\ 857)$ $2^7.3.11.13.17.19.29$ $.37.43.(606\ 521)$ $-3.5^6.11.13.29.43$ $.(4\ 433\ 771)$ $2^2.7^2.13.23.29.41.47$	$-2^5.3^2.11^2.13.31.37$ $.43.(137.413\ 87)$ $-2^4.3.5.13.19.29.37$ $.43.(1453.2731)$ $2^2.3.11.13^3.29.41.43$ $.(101.801\ 77)$	$5^6.11.13.31.37^3.43$ $.(70\ 321)$ $2^8.3.11^2.13.19.29.37$ $.43.(294\ 223)$ $-2^2.3.11.13.29.41$ $.(739\ 283\ 693)$	$-2^7.5^9.11.13.29.31$ $.37^3.43.(79)$ $-2^2.19.29.43.(53)$ $.(42\ 347\ 779)$ $-2.3.5.7^2.13^3.29.41$ $.(149.999\ 91)$	$-2.3.5^2.11.13.19.29$ $.37.43.(2\ 352\ 433)$ $-2^6.11.29.43.(683)$ $.(289\ 099)$ $-2^4.5^4.11^2.13.17^2.23$ $.29.41.(89\ 671)$

TABLE VIII. (*Continued.*)

l	Denominator	$[1\rangle$ $[-8\rangle$ $[16\rangle$ $[-23\rangle$	$[-2\rangle$ $[10\rangle$ $[-17\rangle$ $[25\rangle$	$[4\rangle$ $[-11\rangle$ $[19\rangle$	$[-5\rangle$ $[13\rangle$ $[-20\rangle$	$[7\rangle$ $[-14\rangle$ $[22\rangle$
		.(1019.150 91)				
24c	$2^8.3^{25}.5^5.7.31$	$-2.3^2.13.23.29.31$.(189 391)	$-2^5.3^2.29.31.41^2$.(211 879)	$-5^6.11.29.31.(223)$.(2003)	$-2^7.5^5.11^3.31.(73)$.(109)	$-2.3.5^2.11.19$.(6 669 947)
		$2^5.3.11.17.19.(97)$.(53 549)	$2^4.3.5.19.37^2$.(1 225 099)	$2^{12}.3.17^2.19.(173)$.(193)	$2^2.13.19.37$.(19 418 807)	$-2^6.11.13.37^3$.(123 049)
		$3.5^6.11.37.(97)$.(7079)	$2^2.3.11.37.41.(97)$.(70 181)	$2^2.3.11.37.41.43$.(130 643)	$-2.3.5.7^2.37.41.43$.(79.727)	$-2^4.5^4.23.37.41.43$.(163.191)
		$-2^2.7^2.23.37.41.43.47$.(67.587)				
24d	$2^3.3^{31}.5^6.23^2$ .31.37	$2.3^2.11.13.23.29.31$ .37.41.(1 460 923)	$-2^5.3^2.11.13^4.29.31$ .37.41.(61 057)	$5^6.11^2.19^2.29.31.37^3$ .41.(239)	$-2^9.5^7.31.37^3.41$.(21 269)	$2.3.5^2.11^2.17^2.19.37$ .41.(179 623)
		$-2^3.3.17.19.37.41$.(71.553 591 3)	$-2^4.3.5.11.19.37.41$.(195 606 139)	$-2^6.3.11.19.37.41$.(107.258 114 1)	$2^2.7^2.11.13.19.41$.(101.708 833 9)	$-2^6.13.41.(463)$.(3169.101 41)
		$-3.5^6.41.(619.1319)$.(2609)	$-2^2.3.19^2.(59.839)$.(650 933)	$2^2.3.43.(50\ 329)$.(353 263)	$-2.3.5.11.43.(1249)$.(1553.7877)	$-2^4.5^4.11.19^4.23.43$.(83.1249)
		$-2^2.11.23.43.47$.(4337.175 81)				
25a	$3^{23}.7.61$	$-2^3.3^2.19.23.29.31$.(4657)	$-2.5^2.19.23.29.31.41^2$.(179)	$-2^6.7.11.19.31$.(18 973)	$-2^2.5.11.19.31$.(202 549)	$2^3.11.13^2.29^2.31^2$.(239)
		$-2^5.3.41^2.(4729)$	$3.5.7.11^2.43^2.(11\ 897)$	$-2.7.(787.8443)$	$-2.5^2.13.19.37$.(68 711)	$19.37.(448\ 157)$
		$-2^8.11.19.37.41$.(7673)	$-2^2.3.5^2.7.11.17^2.19$.$23^2.37.41$	$2.3.19.31^2.37.41.43$.(89)	$2^5.5.7^2.19.37.41.43$.(467)	$19.23.37.41.43.47$.(601)
		$2^9.7^2.11^2.19.23.37.41$ .43.47	$2^3.5^2.7^4.19.23.37.41$ .43.47			
25b	$2.3^{24}.5^2.7.17$ .29.61	$-2^5.3^2.29.31.37.41.43$ .47.(2281)	$-2^3.29.31.37.41.43.47$.(19 463)	$-2^{10}.5^2.7.11^3.13^2.23$ .31.37.41.43.47	$-2^4.5^3.11.23.31.37.41$ .43.47.(487)	$2^3.5^4.11.19.23.37.41^3$ .43.47^3
		$2^7.3.17^2.19^3.23.37.41$ .43.47.(67)	$-3.5.7.17^2.19.23.37$ .41.43.47.(1753)	$2^3.7.19.23.37.41.43$ .47.(55 501)	$2.11^2.13.23.41.43.47$.(50 047)	$2^2.23.41.43.47.(107)$.(17 291)
		$2^8.5^2.11.23.43.47$.(35 221)	$-2^4.3.7.11.23^3.43.47$.(62 617)	$-2.3.23.47.(193)$.(1 212 251)	$-2^7.5.7^2.11^2.23^3.47$.(7159)	$5^2.19^2.(223.542947)$
		$2^{11}.7^2.(1423.100\ 91)$	$2^3.5^2.7^2.(30\ 563\ 453)$			
25c	$2^4.3^{24}.5^2.7.17$ $.47^2$	$-2^3.3^2.23.29.31.(421)$.(4789)	$-2.23.29.31.(7\ 181\ 087)$	$2^6.5^2.7^3.11.19^2.29^2$ .31.(59)	$2^2.5^3.11.31.(61)$.(83 443)	$2^3.5^4.11.19.(1\ 003\ 273)$
		$-2^9.3.17^4.19.29^2$.(757)	$3.5.7.17^2.19.(59)$.(6043)	$2^3.7.19.(19\ 414\ 651)$	$-2.11^2.13.37$.(1 488 419)	$-2^2.37.(1759.378\ 11)$
		$-2^6.5^2.11.37^3.41$.(5297)	$2^4.3.7.11.37.41$.(483 827)	$-2.3.37.41.43.(1367)$.(1559)	$-2^7.5.7^2.11^2.37.41.43$.(2857)	$-5^2.23^3.37.41.43.47$.(1637)
		$2^{11}.7^2.23.37.41.43.47$.(613)	$-2^3.5^2.7^2.23.37.41.43$ .47.(1697)			
25d	$2^4.3^{24}.5.29.47^2$	$2^3.3^2.11.19^2.23.29.37$ .41.(743)	$-2.11.23.29.37.41$.(46 589)	$-2^6.5^2.7.37.41^3$.(1373)	$2^2.5.11^2.37.41$.(41 897)	$2^3.5^2.19.31.37.41$.(571)
		$-2^9.3.11.13^2.19.31.37$ .41.(353)	$3.5.7.11.19.31.37.41$.(2269)	$2^3.7.11.19^3.31^3.37.41$ $.43^2$	$-2.11.13^3.31.41$.(5197)	$-2^2.11.31.41.(900\ 701)$
		$-2^6.5^2.31.(182\ 519)$	$-2^4.3.7.31.(1\ 473\ 847)$	$2.3.11.31.43.(211)$.(2053)	$2^7.5.11.31.43.(86\ 771)$	$5^2.11.23.31.43.47$.(17 341)
		$-2^{11}.11.23.31.43.47$.(181)	$-2^3.5^2.7^2.11.23.31.43$ .47.(1201)			

$$|\Gamma_6 j\, 6\, \hat{\tfrac{1}{2}}\rangle = \sqrt{2.13}\left[\frac{\sqrt{7}}{4}\begin{pmatrix}6 & \frac{1}{2} & j\\ 4 & \frac{1}{2} & -\frac{9}{2}\end{pmatrix}|j\,\tfrac{9}{2}\rangle - \frac{1}{2\sqrt{2}}\begin{pmatrix}6 & \frac{1}{2} & j\\ 0 & \frac{1}{2} & -\frac{1}{2}\end{pmatrix}|j\,\tfrac{1}{2}\rangle + \frac{\sqrt{7}}{4}\begin{pmatrix}6 & \frac{1}{2} & j\\ -4 & \frac{1}{2} & \frac{7}{2}\end{pmatrix}|j\,-\tfrac{7}{2}\rangle\right]. \quad (60)$$

Tables of analytic formulas for Clebsch–Gordan coefficients with an angular momentum up to 5 are in Ref. 18; a very concise table of $3jm$ coefficients with an angular momentum up to 2 can be found in Appendix C of Ref. 19. Using them for $j = \frac{11}{2}$ and $\frac{13}{2}$, we get

$$|\Gamma_6\, \tfrac{11}{2}\, 6\, \pm\hat{\tfrac{1}{2}}\rangle = \mp\frac{\sqrt{7}}{4\sqrt{3}}|\tfrac{11}{2}\ \pm\tfrac{9}{2}\rangle \pm\frac{1}{2\sqrt{2}}|\tfrac{11}{2}\ \pm\tfrac{1}{2}\rangle \mp\frac{\sqrt{5.7}}{4\sqrt{3}}|\tfrac{11}{2}\ \mp\tfrac{7}{2}\rangle,$$

$$|\Gamma_6\, \tfrac{13}{2}\, 6\, \pm\hat{\tfrac{1}{2}}\rangle = \frac{\sqrt{11}}{4}|\tfrac{13}{2}\ \pm\tfrac{9}{2}\rangle - \frac{1}{2\sqrt{2}}|\tfrac{13}{2}\pm\tfrac{1}{2}\rangle + \frac{\sqrt{3}}{4}|\tfrac{13}{2}\ \mp\tfrac{7}{2}\rangle. \quad (61)$$

With the same Γ_1, we also obtain 3 Γ_4 representations the components of which in $|\hat{0}\ -\rangle$ subspace are

$$|\Gamma_4\, 5\, 6\, \hat{0}\rangle = -\frac{\sqrt{5.7}}{2\sqrt{2.11}}|5\,4\rangle + \frac{3}{2\sqrt{11}}|5\,0\rangle - \frac{\sqrt{5.7}}{2\sqrt{2.11}}|5\ -4\rangle,$$

$$|\Gamma_4\, 6\, 6\, \hat{0}\rangle = \frac{1}{\sqrt{2}}|6\,4\rangle - \frac{1}{\sqrt{2}}|6\ -4\rangle, \quad (62)$$

$$|\Gamma_4\, 7\, 6\, \hat{0}\rangle = \frac{\sqrt{3.11}}{4\sqrt{5}}|7\,4\rangle - \frac{\sqrt{7}}{2\sqrt{2.5}}|7\,0\rangle + \frac{\sqrt{3.11}}{4\sqrt{5}}|7\ -4\rangle,$$

and 4 Γ_8 representations the components of which in $|\pm\hat{\tfrac{1}{2}}\rangle$ are

$$|\Gamma_8\, \tfrac{9}{2}\, 6\, \pm\hat{\tfrac{1}{2}}\rangle = \pm\frac{\sqrt{7}}{2\sqrt{2.11}}|\tfrac{9}{2}\ \pm\tfrac{9}{2}\rangle \mp\frac{3}{2\sqrt{11}}|\tfrac{9}{2}\pm\tfrac{1}{2}\rangle \pm\frac{3\sqrt{7}}{2\sqrt{2.11}}|\tfrac{9}{2}\mp\tfrac{7}{2}\rangle,$$

$$|\Gamma_8\, \tfrac{11}{2}\, 6\, \pm\hat{\tfrac{1}{2}}\rangle = \frac{-19}{4\sqrt{3.11}}|\tfrac{11}{2}\ \pm\tfrac{9}{2}\rangle + \frac{\sqrt{7}}{2\sqrt{2.11}}|\tfrac{11}{2}\pm\tfrac{1}{2}\rangle + \frac{5\sqrt{5}}{4\sqrt{3.11}}|\tfrac{11}{2}\mp\tfrac{7}{2}\rangle, \quad (63)$$

$$|\Gamma_8\, \tfrac{13}{2}\, 6\, \pm\hat{\tfrac{1}{2}}\rangle = \pm\frac{\sqrt{11}}{2\sqrt{2.5}}|\tfrac{13}{2}\ \pm\tfrac{9}{2}\rangle \pm\frac{1}{2\sqrt{5}}|\tfrac{13}{2}\pm\tfrac{1}{2}\rangle \mp\frac{3\sqrt{3}}{2\sqrt{2.5}}|\tfrac{13}{2}\mp\tfrac{7}{2}\rangle,$$

$$|\Gamma_8\, \tfrac{15}{2}\, 6\, \pm\hat{\tfrac{1}{2}}\rangle = \frac{3\sqrt{11}}{4\sqrt{2.5}}|\tfrac{15}{2}\ \pm\tfrac{9}{2}\rangle - \frac{\sqrt{7}}{2\sqrt{2.5}}|\tfrac{15}{2}\ \pm\tfrac{1}{2}\rangle + \frac{\sqrt{3.11}}{4\sqrt{2.5}}|\tfrac{15}{2}\mp\tfrac{7}{2}\rangle.$$

Naturally, the corresponding components $|\Gamma_4 \pm\hat{1}\rangle$ and $|\Gamma_8 \pm\hat{\tfrac{3}{2}}\rangle$ are also obtained. Butler's[5] tables include these Γ_6 and Γ_4 representations, as well as the Γ_8 representations for $j = \frac{11}{2}$ and $j = \frac{15}{2}$ but not the two other Γ_8; the Γ_8 representations for $j = \frac{9}{2}$ published by Butler are of our Γ_8'' type because they can be built from the Γ_3 representations for $j = 4$ and $j = 5$, while his Γ_8 representations for $j = \frac{13}{2}$ are of our Γ_8' type because they can be built from the Γ_2 representations for $j = 6$ and $j = 7$.

ACKNOWLEDGMENTS

We thank M. Kibler for his interest in our work and for useful comments, and a referee for calling our attention to the work of Butler.[5]

[1] R. Conte, J. Raynal, and E. Soulié, "Determination of point group harmonics for arbitrary j by a projection method. I. Cubic group, quantization along an axis of order 4," J. Math. Phys. **25**, 1176 (1984).
[2] L. Michel, "Invariants polynomiaux des groupes de symétries moléculaire et cristallographique," in *Proceedings of the 5th International Colloquium on Group Theoretical Methods in Physics*, edited by R. T. Sharp and B. Kolman (Academic, New York, 1977), pp. 75–91.
[3] J. Patera, R. T. Sharp, and P. Winternitz, "Polynomial irreducible tensors for point groups," J. Math. Phys. **19**, 2362 (1978) (integer j); Erratum in Y. Saint-Aubin, "Fonctions génératrices et bases d'intégrité pour les sous-groupes finis du groupe de Lorentz 0(3,1)," Can. J. Phys. **58**, 1075 (1980); P. E. Desmier and R. T. Sharp, "Polynomial tensors for double point groups," J. Math. Phys. **20**, 74 (1978) (half-integer j).
[4] T. Molien, "Uber die Invarianten der linearen Substitutionsgruppen," Sitzungber. K. Preuss. Akad. Wiss **2**, 1152 (1897).
[5] P. H. Butler, *Point Group Symmetry Applications, Methods and Tables* (Plenum, New York, 1981), pp. 80 and 527–531.
[6] M. Kibler, "Ionic and paramagnetic energy levels: Algebra," J. Mol. Spectros. **26**, 111 (1968); "Energy levels of paramagnetic ions: Algebra. II," Int. J. Quantum Chem. **3**, 795 (1969).
[7] J. Griffith, *The Irreducible Tensor Method for Molecular Symmetry Groups* (Prentice-Hall, Englewood Cliffs, NJ, 1962), p. 4.
[8] Y. Tanabe and H. Kamimura, "On the absorption spectra of complex ions. IV. The effect of the spin-orbit interaction and the field of lower symmetry on d electrons in cubic field," J. Phys. Soc. Jpn. **13**, 394 (1958).
[9] S. E. Harnung and C. E. Schäffer, "Phase-fixed 3-Γ symbols and coupling coefficients for the point groups," Struct. Bonding **12**, 201 (1972).
[10] P. H. Butler, "Coupling coefficients and tensor operators for chains of groups," Philos. Trans. R. Soc. London Ser. A **227**, 545 (1975); M. R. Kibler and P. A. M. Guichon, "Clebsch–Gordan coefficients for chains of groups of interest in quantum chemistry," Int. J. Quantum Chem. **10**, 87 (1976); M. R. Kibler, "Group theory around ligand field theory," in *Proceedings of the 5th International Colloquium on Group Theoretical Methods in Physics*, edited by R. T. Sharp and B. Kolman (Academic, New York, 1977), pp. 161–172.
[11] J. Moret-Bailly, "Sur l'interprétation des spectres de vibration-rotation des molécules à symétrie tétraédrique ou octaédrique," Cah. Phys. **26**, 311 (1961).
[12] A. Abragam and B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions* (Clarendon, Oxford, 1970), pp. 632 and 857.
[13] J. Griffith, *The Theory of Transition Metal Ions* (Cambridge U.P., Cambridge, 1971), p. 393.
[14] S. L. Altmann and A. P. Cracknell, "Lattice harmonics. I. Cubic group," Rev. Mod. Phys. **37**, 19 (1965).
[15] Y. Ayant, E. Bélorizky and J. Rosset, "Quelques propriétés des niveaux cristallins Γ_8 en symétrie cubique," J. Phys. Radium **23**, 201 (1962).
[16] D. Descamps and Y. Merle d'Aubigné, "Electron paramagnetic resonance of the Er^{3+} ion in magnesium oxide," Phys. Lett. **8**, 5 (1964).
[17] J. Raynal. and R. Conte, "Determination of point group harmonics for arbitrary j by a projection method. III. Cubic group, quantization along a ternary axis," J. Math. Phys. **25**, 3375 (1984).
[18] D. A. Varshalovich, A. N. Moskaliev, and V. K. Khersonskii, *Quantum Theory of Angular Momenta* (Nauka, Leningrad, 1975), p. 228 (in Russian).
[19] A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1959), p. 911 (in French).

On a labeling for point group harmonics. II. Icosahedral group

Jacques Raynal
Service de Physique Théorique, CEN-SACLAY, 91191 Gif-sur-Yvette Cedex, France

(Received 17 July 1984; accepted for publication 4 April 1985)

The expressions for the Γ_2, Γ_5, Γ_6, Γ_8, and Γ_9 representations are given in terms of those of Γ_1 representations for neighboring angular momenta. The coefficients of the Γ_4 representations are expressed in terms of those of Γ_7 representations. Therefore, with an arbitrary choice of orthonormal sets of Γ_1, Γ_3, and Γ_7 representations, orthonormal sets of other kinds of representations are well defined and can be labeled with the labels of parent representations. All Clebsch–Gordan coefficients are expressed in terms of those between parent representations (and a few others). Tables of all nondegenerate Γ_1, Γ_3, and Γ_7 representations are given, for the axes of quantization of order 5. With some degenerate Γ_3 and Γ_7 representations, which are also given, any representation of integer or half-integer angular momentum up to 27 can be obtained using some usual Clebsch–Gordan coefficients of SU(2).

I. INTRODUCTION

In a first series of articles,[1–3] we gave expressions for point group harmonics in terms of reduced rotation matrix elements. The first of these articles described a projection method to obtain the coefficients of cubic harmonics quantized on an axis of order 4, the second one described the application of this method to the icosahedral group, and the last one to the cubic group with a quantization axis of order 3. We wanted to give tables of coefficients in terms of integer as far as possible. However, we noticed relations between different kinds of representations, which did not seem to be known until now. An account of these relations and a study of some of their consequences have been given[4] for the cubic group. Similar relations for the icosahedral group are presented here.

For the icosahedral group, we use the same axes as in the previous work.[2] We choose one of the six rotation axes of order 5 as the z axis and another one in the x-z plane. As the rotation of π around the y axis is an element of the group, all the coefficients are real; this is an advantage over other notations.[5] The five kinds of representations for integer values of the angular momentum j and the four kinds of representations for half-integer values have been presented[2] and will not be repeated here.

As for the cubic group, the number $n_i(j)$ of representations Γ_i obtained for a given value of the angular momentum j can be derived from the characters of SU(2) and of the double icosahedral group. They are given in Table I with their generating functions which we define by

$$g_i(x) = \sum_j x^{2j} n_i(j). \tag{1}$$

TABLE I. Multiplicities and their generating functions. The second column gives the multiplicity for a given value of j, using the a's quoted at the bottom of the table. The generating functions of the last column use x^2 instead of x in order to simplify the relations between multiplicities for different values of j, integer or half-integer.

	$n(j)$	$g(x)$
Γ_1	$\{a_1 + 15a_2 + 20a_3 + 12a_4 + 12a_5\}/60$	$\dfrac{1+x^{30}}{(1-x^{12})(1-x^{20})}$
Γ_2	$\{a_1 - 5a_2 + 2a_4(1+\sqrt{5}) + 2a_5(1-\sqrt{5})\}/20$	$\dfrac{x^2(1+x^{12})}{(1-x^8)(1-x^{10})}$
Γ_3	$\{a_1 - 5a_2 + 2a_4(1-\sqrt{5}) + 2a_5(1+\sqrt{5})\}/20$	$\dfrac{x^6}{(1-x^4)(1-x^{10})}$
Γ_4	$\{a_1 + 5a_3 - 3a_4 - 3a_5\}/15$	$\dfrac{x^6(1+x^2)}{(1-x^6)(1-x^{10})}$
Γ_5	$\{a_1 + 3a_2 - 4a_3\}/12$	$\dfrac{x^4}{(1-x^4)(1-x^6)}$
Γ_6	$\{a_1 + 10a_3 + 3(1+\sqrt{5})a_4 - 3(1-\sqrt{5})a_5\}/30$	$\dfrac{x(1+x^{18})}{(1-x^{10})(1-x^{12})}$
Γ_7	$\{a_1 + 10a_3 + 3(1-\sqrt{5})a_4 - 3(1+\sqrt{5})a_5\}/30$	$\dfrac{x^7}{(1-x^6)(1-x^{10})}$
Γ_8	$\{a_1 - 5a_3 + 3a_4 - 3a_5\}/15$	$\dfrac{x^3(1+x^8)}{(1-x^6)(1-x^{10})}$
Γ_9	$\{a_1 - 2a_4 + 2a_5\}/10$	$\dfrac{x^5}{(1-x^2)(1-x^{10})}$

$a_1 = 2j+1 \quad a_2 = \sin(a_1\pi/2) \quad a_3 = \sin(a_1\pi/3)/\sin(\pi/3) \quad a_4 = \sin(a_1\pi/5)/\sin(\pi/5) \quad a_5 = \sin(2a_1\pi/5)/\sin(2\pi/5)$

 0022-2488/85/102441-16$02.50

Among them, $g_1(x)$ has been known for a long time.[6] Explicit values of the multiplicities for the 15 first integer and half-integer angular momenta are given in Table II. They will be of great help in our demonstration.

The coefficients of the Γ_1 representations for which the multiplicity is unity are given later in Table V. They are quite simple but not so simple as for the cubic group.[4] Their square is the product of prime numbers smaller than $2j$, divided by a large power of 5 and sometimes by 2 and/or 3; often a larger prime number appears but this can be taken out of the root: the first of them is 41 for $j = 20$ and the largest of them is 586 471 for $j = 59$, which is the last nondegenerate Γ_1 representation which exists. The first Γ_1 representations with multiplicity 2 appear for $j = 30$; no rotation could be found between these two representations, which lead to simple coefficients. However, with the coefficients of the Γ_1 representations for $j<29$, we can express the coefficients of any kind of representation for $j<27$ except for the Γ_3, Γ_4, and Γ_7 representations. The coefficients of the Γ_7 representations with multiplicity 1 are given in Table VI and those of the two components of the few nondegenerate Γ_3 representations in Tables VII and VIII; they are as simple as those of the Γ_1 representations. The coefficients of the Γ_4 representations can be expressed with those of the Γ_7 representations. The Tables IX–XI give the coefficients of the degenerated Γ_7 and Γ_3 representations needed to reach $j = 27$ for any kind of representation: their values are quite arbitrary whereas those of Tables V–VIII are fixed to within a sign.

TABLE II. Multiplicities of the irreducible representations for the first 15 integer and half-integer angular momenta. The last lines give the period of the multiplicity and the increment. When the angular momentum is increased by the period, the multiplicity is increased by the increment. This table has been shortened: there is one Γ_1 representation for $l = 15,16,18,20,21,22,24,25,26,27$, and 28 and none for $l = 17,19,23$, and 29.

l	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	Γ_9	j
0	1	0	0	0	0					
						1	0	0	0	1/2
1	0	1	0	0	0					
						0	0	1	0	3/2
2	0	0	0	0	1					
						0	0	0	1	5/2
3	0	0	1	1	0					
						0	1	0	1	7/2
4	0	0	0	1	1					
						0	0	1	1	9/2
5	0	1	1	0	1					
						1	0	1	1	11/2
6	1	1	0	1	1					
						1	1	1	1	13/2
7	0	1	1	1	1					
						0	0	1	2	15/2
8	0	0	1	1	2					
						0	1	1	2	17/2
9	0	1	1	2	1					
						1	1	1	2	19/2
10	1	1	1	1	2					
						1	0	2	2	21/2
11	0	2	1	1	2					
						1	1	2	2	23/2
12	1	1	1	2	2					
						1	1	1	3	25/2
13	0	1	2	2	2					
						0	1	2	3	27/2
14	0	1	1	2	3					
						1	1	2	3	29/2
Per.	30	10	10	15	6	15	15	15	5	
Inc.	1	1	1	2	1	1	1	2	1	

In next section, we show that the coefficients of the Γ_2, Γ_5, Γ_6, Γ_8, and Γ_9 representations for a given angular momentum j can be obtained by multiplying the coefficients of a Γ_1 representation for a neighboring angular momentum l by a Clebsch–Gordan coefficient of SU(2). These sets of representations are orthonormal and each representation can be labeled by the angular momentum l and the extra label α of the parent Γ_1 representation from which it is built. The remaining Γ_3, Γ_4, and Γ_7 representations have dimension 3, 4, and 2, respectively. From the decomposition into irreducible representations of the direct product of two representations[7] given in Table III we see that the Γ_4 and Γ_9 representations can be obtained from the Γ_7 representations. Then, only the Γ_3 representations are left; the Γ_9 representations can also be obtained from them. So, when the Γ_1, Γ_7, and Γ_3 representations are known and labeled, all the other kinds of representations are known and labeled. There are three labelings of Γ_9 representations, as was also the case for the Γ_8 representations for the cubic group.[4] To distinguish them, we shall denote them as Γ_9, Γ_9', and Γ_9'' representations when the parent representations are Γ_1, Γ_7, and Γ_3 representations, respectively. In the two cases of degeneracy which appear for an angular momentum smaller than or equal to 8, these representations differ from those published by Mc Lellan[5] but agree with those published by Butler.[8] As for cubic harmonics, Butler's criteria of simplicity lead to the same results at those presented here.

As in the case of cubic harmonics, with three components of three representations

$$|\Gamma_i j\alpha\hat{m}\rangle = \sum_{\mu\equiv\hat{m}(5)} x^{\alpha}_{j\mu}\,|j\mu\rangle, \tag{2}$$

one can define an icosahedral $3jm$ symbol by

$$\begin{pmatrix} \Gamma_i j\alpha & \Gamma_{i'} j'\alpha' & \Gamma_{i''} j''\alpha'' \\ \hat{m} & \hat{m}' & \hat{m}'' \end{pmatrix} = \sum_{\mu,\mu',\mu''} x^{\alpha}_{j\mu} x^{\alpha'}_{j'\mu'} x^{\alpha''}_{j''\mu''} \begin{pmatrix} j & j' & j'' \\ \mu & \mu' & \mu'' \end{pmatrix}. \tag{3}$$

We consider that as a sum of $3jm$ symbols of SU(2) weighted by the coefficients of icosahedral harmonics. In fact, since all our coefficients $x^{j\alpha}_{\mu}$ are real, this icosahedral $3jm$ symbol coincides with the f symbol defined by Kibler.[9] Its dependence on the magnetic quantum numbers $\hat{m}$ is eliminated by dividing it by a similar expression obtained with the "basic" representation

$$|\gamma_i q_i \hat{m}\rangle = \sum_{\mu} x_{q_i\mu}\,|q_i\mu\rangle, \tag{4}$$

which is the representation of the same kind as Γ_i but with the lowest possible angular momentum. The values of the angular momenta q_i can be found in Table II. In fact, when $\Gamma_{i''}$ is not unique in the product of Γ_i and $\Gamma_{i'}$, the dependence on magnetic quantum numbers is a linear combina-

tion of the icosahedral $3jm$ symbols obtained with the "basic" representations and those obtained with the "next basic" representation $\gamma'_{i''}$ (next lowest angular momentum) in the third position. In the case of two Γ_9 representations and a Γ_5 representation, one even needs a "third basic" Γ_5 or Γ_9 representation, $\gamma''_{i''}$. Writing

$$\begin{aligned}
&\begin{pmatrix} \Gamma_i j\alpha & \Gamma_{i'} j'\alpha' & \Gamma_{i''} j''\alpha'' \\ \hat{m} & \hat{m}' & \hat{m}'' \end{pmatrix} \\
&\quad = (\Gamma_i j\alpha \quad \Gamma_{i'} j'\alpha' \quad \Gamma_{i''} j''\alpha'')^{(1)} \\
&\qquad \times \begin{pmatrix} \gamma_i q_i & \gamma_{i'} q_{i'} & \gamma_{i''} q_{i''} \\ \hat{m} & \hat{m}' & \hat{m}'' \end{pmatrix} \\
&\qquad + (\Gamma_i j\alpha \quad \Gamma_{i'} j'\alpha' \quad \Gamma_{i''} j''\alpha'')^{(2)} \\
&\qquad \times \begin{pmatrix} \gamma_i q_i & \gamma_{i'} q_{i'} & \gamma'_{i''} q'_{i''} \\ \hat{m} & \hat{m}' & \hat{m}'' \end{pmatrix} \\
&\qquad + (\Gamma_i j\alpha \quad \Gamma_{i'} j'\alpha' \quad \Gamma_{i''} j''\alpha'')^{(3)} \\
&\qquad \times \begin{pmatrix} \gamma_i q_i & \gamma_{i'} q_{i'} & \gamma''_{i''} q''_{i''} \\ \hat{m} & \hat{m}' & \hat{m}'' \end{pmatrix},
\end{aligned} \tag{5}$$

we define what we call "icosahedral 3Γ symbols," which are reduced matrix elements. All the 3Γ symbols are unity between basic representations. They are the ratio of reduced coefficients $\bar{f}(( ))$ of Kibler or of the $3jm$ factors tabulated by Butler by the reduced coefficient $\bar{f}(( ))$ or $3jm$ factor for the "basic representations." The reason why we adopt such definitions is that they allow us to use the Racah algebra of SU(2) and that the results can be easily translated by anybody into his own notation. The icosahedral $3jm$ symbols (3) are invariant by a circular permutation of their arguments; for an odd permutation, there is a phase $(-)^{j+j'+j''}$. The same holds for the icosahedral 3Γ symbols if the kind of the third representation does not change; in the other cases the following relations hold for a circular permutation:

$$\begin{aligned}
&(\Gamma_9\Gamma_9\Gamma_2) = (\Gamma_2\Gamma_9\Gamma_9), \\
&(\Gamma_9\Gamma_9\Gamma_2)^{(2)} = \tfrac{1}{2}\sqrt{11}(\Gamma_2\Gamma_9\Gamma_9)^{(2)};
\end{aligned} \tag{6}$$

$$\begin{aligned}
&(\Gamma_9\Gamma_9\Gamma_3) = (\Gamma_3\Gamma_9\Gamma_9) - \tfrac{1}{3}\sqrt{\tfrac{5}{6}}(\Gamma_3\Gamma_9\Gamma_9)^{(2)}, \\
&(\Gamma_9\Gamma_9\Gamma_3)^{(2)} = -\tfrac{1}{3}\sqrt{\tfrac{77}{6}}(\Gamma_3\Gamma_9\Gamma_9)^{(2)};
\end{aligned} \tag{7}$$

$$\begin{aligned}
&(\Gamma_5\Gamma_5\Gamma_4) = (\Gamma_4\Gamma_5\Gamma_5) - \tfrac{1}{2}\sqrt{2}(\Gamma_4\Gamma_5\Gamma_5)^{(2)}, \\
&(\Gamma_5\Gamma_5\Gamma_4)^{(2)} = \tfrac{1}{2}\sqrt{\tfrac{7}{2}}(\Gamma_4\Gamma_5\Gamma_5)^{(2)};
\end{aligned} \tag{8}$$

$$\begin{aligned}
(\Gamma_8\Gamma_9\Gamma_4) &= (\Gamma_4\Gamma_8\Gamma_9) - \tfrac{1}{8}\sqrt{15}(\Gamma_4\Gamma_8\Gamma_9)^{(2)} \\
&= (\Gamma_9\Gamma_4\Gamma_8) + \tfrac{3}{2}\sqrt{\tfrac{3}{110}}(\Gamma_9\Gamma_4\Gamma_8)^{(2)}, \\
(\Gamma_8\Gamma_9\Gamma_4)^{(2)} &= \tfrac{3}{8}\sqrt{7}(\Gamma_4\Gamma_8\Gamma_9)^{(2)} = \tfrac{3}{2}\sqrt{\tfrac{7}{22}}(\Gamma_9\Gamma_4\Gamma_8)^{(2)};
\end{aligned} \tag{9}$$

$$\begin{aligned}
&(\Gamma_9\Gamma_9\Gamma_4) = (\Gamma_4\Gamma_9\Gamma_9) + \tfrac{1}{4}\sqrt{\tfrac{5}{6}}(\Gamma_4\Gamma_9\Gamma_9)^{(2)}, \\
&(\Gamma_9\Gamma_9\Gamma_4)^{(2)} = -\tfrac{1}{4}\sqrt{\tfrac{21}{2}}(\Gamma_4\Gamma_9\Gamma_9)^{(2)};
\end{aligned} \tag{10}$$

$$\begin{aligned}
&(\Gamma_8\Gamma_9\Gamma_5) = (\Gamma_5\Gamma_8\Gamma_9) = (\Gamma_9\Gamma_5\Gamma_8), \\
&(\Gamma_8\Gamma_9\Gamma_5)^{(2)} = \tfrac{3}{2}\sqrt{\tfrac{3}{5}}(\Gamma_5\Gamma_8\Gamma_9)^{(2)} \\
&\qquad = -\tfrac{3}{5}\sqrt{2}(\Gamma_9\Gamma_5\Gamma_8)^{(2)};
\end{aligned} \tag{11}$$

$$\begin{aligned}
&(\Gamma_9\Gamma_9\Gamma_5) = (\Gamma_5\Gamma_9\Gamma_9), \\
&(\Gamma_9\Gamma_9\Gamma_5)^{(2)} = -\sqrt{\tfrac{3}{10}}(\Gamma_5\Gamma_9\Gamma_9)^{(2)} \\
&\qquad + \tfrac{1}{5}\sqrt{21}(\Gamma_5\Gamma_9\Gamma_9)^{(3)}, \\
&(\Gamma_9\Gamma_9\Gamma_5)^{(3)} = -\tfrac{1}{2}\sqrt{\tfrac{77}{15}}(\Gamma_5\Gamma_9\Gamma_9)^{(2)} - \tfrac{1}{5}\sqrt{\tfrac{22}{3}}(\Gamma_5\Gamma_9\Gamma_9)^{(3)}.
\end{aligned} \tag{12}$$

This comes from the choice of replacing always the third basic representation by the next basic to define the icosahedral 3Γ symbols when there are at least two of them. Relations (6)–(12) are deduced from the icosahedral $3jm$ symbols between basic and next basic representations. Note that there is only the $(\Gamma_4\Gamma_4\Gamma_4)^{(2)}$ icosahedral 3Γ symbol because the icosahedral $3jm$ symbols between three basic Γ_4 representations vanish. In the third section, the icosahedral 3Γ symbols are expressed in terms of those between parent representations.

There are many applications of these results. In the last section we define a fictitious spin for Γ_2, Γ_4, Γ_5, Γ_6, and Γ_8 representations and we give tables of coefficients of Γ_1, Γ_7, and Γ_3 representations needed to obtain all representations for angular momenta up to 27. For the tables already published, those of Cohan[10] relate to integer angular momentum up to $j = 14$ in decimal values and those of Mc Lellan[5] and Butler[8] go up to $j = 8$ in terms of integers.

II. EXPRESSION FOR OTHER REPRESENTATIONS IN TERMS OF THE Γ_1, Γ_7, AND Γ_3 REPRESENTATIONS

For a given value of the angular momentum l, the expression for a Γ_1 representation is

$$|\Gamma_1\, l\, \alpha\, \hat{0}\, +\rangle = \sum_{m\equiv 0(5)} a^{\alpha}_{lm} |lm\rangle, \tag{13}$$

where $a^{\alpha}_{l-m} = (-)^{l+m} a^{\alpha}_{lm}$. This notation is not the same as in a previous work[2]: the sum over m is extended to negative values and $m = 0$ is taken into account with a vanishing coefficient for an odd value of l. The index α labels the $n_1(l)$ different orthonormal Γ_1 representations which exist for this angular momentum.

We shall show below that

$$\begin{aligned}
|\Gamma_i\, j\, l\alpha\, \hat{\sigma}\rangle &= \sqrt{(2s+1)(2l+1)} \sum_{m\equiv 0(5)} (-)^{l-s+\mu} \\
&\quad \times a^{\alpha}_{lm} \begin{pmatrix} l & s & j \\ m & \hat{\sigma} & -\mu \end{pmatrix} |j\mu\rangle,
\end{aligned} \tag{14}$$

for a given value of j and all the values of (l,α) for which the $3jm$ coefficient exists, is a complete orthonormal set of Γ_6, Γ_2, Γ_8, Γ_5, or Γ_9 representations when $s = \frac{1}{2}$, 1, $\frac{3}{2}$, 2, or $\frac{5}{2}$, respectively. This result is trivial when written between basic representations because $|s\sigma\rangle$ for $s = \frac{1}{2}$ to $\frac{5}{2}$ are the basic Γ_6, Γ_2, Γ_8, Γ_5, and Γ_9 representations of the icosahedral group as seen in Table II.

Among the other kinds of representation, the Γ_7 representations have the lowest dimension. We shall note by l the angular momentum of the parent representation and by j that of the derived representation, independent of their val-

ue. For a given half-integer angular momentum l, the two components of a Γ_7 representation are

$$|\Gamma_7\, l\beta\ \pm\tfrac{3}{2}\hat{}\rangle = \sum_{m=\pm(3/2)(5)} b^{\pm\beta}_{lm}|j\,m\rangle, \tag{15}$$

with $b^{-\beta}_{l-m} = (-)^{l-m} b^{+\beta}_{lm}$. Then

$$|\Gamma_i\, j\, l\beta\ \pm\hat{\sigma}\rangle = \epsilon_{i,\hat{\sigma}}\sqrt{(2s+1)(2l+1)}$$
$$\times \sum_{m=\pm(3/2)(5)} (-)^{l-s+\mu} b^{\pm\beta}_{lm}$$
$$\times \begin{pmatrix} l & s & j \\ m & \sigma & -\mu \end{pmatrix} |j\mu\rangle \tag{16}$$

is a complete orthonormal set of Γ_4 representations when $s=\frac{1}{2}$ and of Γ_9 representations (which we note as Γ'_9 representations) when $s=1$ if $\hat{\sigma}\neq\frac{5}{2}$. The phase $\epsilon_{i\hat{\sigma}}$ turns out to be $(-)^{1/2+\sigma}$ for a Γ_4 representation and $(-)^{\sigma}$ for a Γ'_9 representation. The component $|\Gamma'_9\ \pm\frac{3}{2}\hat{}\rangle$ can be obtained with

$$|\Gamma'_9\, j\, l\beta\pm\tfrac{3}{2}\hat{}\rangle$$
$$= \sqrt{3(2l+1)}\left\{\frac{1}{\sqrt{10}}\sum_{m=\pm(3/2)(5)}(-)^{l-1+\mu}\right.$$
$$\times b^{\pm\beta}_{lm}\begin{pmatrix} l & 1 & j \\ m & \pm 1 & -\mu \end{pmatrix}|j\mu\rangle$$
$$\pm\frac{3}{\sqrt{10}}\sum_{m=\mp(3/2)(5)}(-)^{l-1+\mu}$$
$$\left.\times b^{\mp\beta}_{lm}\begin{pmatrix} l & 1 & j \\ m & \mp 1 & -\mu \end{pmatrix}|j\mu\rangle\right\}. \tag{17}$$

The Γ_3 representations are not given by these formulas. They have a component in the $|\hat{0}-\rangle$ subspace

$$|\Gamma_3\, l j\hat{0}-\rangle = \sum_{m=\hat{0}(5)} c^{\gamma}_{lm}|l\,m\rangle, \tag{18}$$

with $c^{\gamma}_{l-m} = (-)^{l+m+1}c^{\gamma}_{lm}$, and a component in the $|\pm\hat{2}\rangle$ subspaces

$$|\Gamma_3\, l j\pm\hat{2}\rangle = \sum_{m=\pm 2(5)} d^{\pm\gamma}_{lm}|l\,m\rangle, \tag{19}$$

with $d^{-\gamma}_{l-m} = (-)^{l+m+1}d^{+\gamma}_{lm}$. With them, a third set of Γ_9 representations can be obtained: we denote them as Γ''_9 representations. They are

$$|\Gamma''_9\, j\, l\gamma\pm\tfrac{1}{2}\hat{}\rangle = \epsilon_{\pm 1/2}\sqrt{2(l+1)}\sum_{m=\hat{0}(5)}(-)^{l-1/2+\mu}$$
$$\times c^{\gamma}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \pm\frac{1}{2} & -\mu \end{pmatrix}|j\mu\rangle,$$

$$|\Gamma''_9\, j\, l\gamma\pm\tfrac{3}{2}\hat{}\rangle = \epsilon_{\pm 3/2}\sqrt{2(2l+1)}\sum_{m=\pm 2(5)}(-)^{l-1/2+\mu}$$
$$\times d^{\pm\gamma}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \mp\frac{1}{2} & \mu \end{pmatrix}|j\mu\rangle, \tag{20}$$

$$|\Gamma''_9\, j\, l\gamma\pm\tfrac{5}{2}\hat{}\rangle = \sqrt{2(2l+1)}\left[\mp\frac{1}{\sqrt{5}}\sum_{m=\pm 2(5)}(-)^{l-1/2+\mu}\right.$$
$$\times d^{\pm\gamma}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \pm\frac{1}{2} & \mu \end{pmatrix}|j\mu\rangle$$
$$+\frac{2}{\sqrt{5}}\sum_{m=\mp 2(5)}(-)^{l-1/2+\mu}$$
$$\left.\times d^{\mp\gamma}_{lm}\begin{pmatrix} l & \frac{1}{2} & j \\ m & \mp\frac{1}{2} & \mu \end{pmatrix}|j\mu\rangle\right].$$

The extra phase is $\epsilon_{\sigma} = (-)^{1/2+\sigma}$. Here also, the components $\pm\frac{3}{2}\hat{}$ are a mixture of coupling to $\frac{3}{2}$ and $-\frac{5}{2}$. The Γ'_9 and Γ''_9 representations do not appear in as simple a form as for the Γ_9 representations.

The kind of representation given by formulas (14), (16), (17), and (20) is straightforward. The completeness of the sets of representations can be deduced from the generating functions when their orthogonality is proved. For example, the completeness of the Γ_9, Γ'_9, and Γ''_9 representations comes from the relations

$$g_9(x) = (1/x^5 + 1/x^3 + 1/x + x + x^3 + x^5)\, g_1(x)$$
$$- 1/x^5 - 1/x^3 - 1/x - x - x^3$$
$$= (1/x^2 + 1 + x^2)\, g_7(x) = (1/x + x)\, g_3(x), \tag{21}$$

where the terms subtracted correspond to representations forbidden by triangular relations.

To show that these sets of representations are orthonormal let us consider the scalar product of two components. For simplicity, we discard the components $\pm\frac{3}{2}\hat{}$ of Γ'_9 and Γ''_9 representations. Such a scalar product is

$$\langle\Gamma_{i_1}\, j\, l_1\, \nu_1\, \hat{\sigma}_1|\Gamma_{i_2}\, j\, l_2\, \nu_2\, \hat{\sigma}_2\rangle$$
$$= \epsilon_1\epsilon_2(-)^{l_1+j-\sigma_1}\sqrt{(2s_1+1)(2s_2+1)(2l_1+1)}$$
$$\times\sum_{L,M}(2L+1)\begin{pmatrix} s_2 & s_1 & L \\ \sigma_2 & -\sigma_1 & M \end{pmatrix} N(L,M), \tag{22}$$

with

$$N(L,M) = \sqrt{2l_2+1}\sum_{m_1\,m_2}(-)^{l_2-m_2}\, x^{\nu_1}_{l_1 m_1}\, x^{\nu_2}_{l_2 m_2}$$
$$\times\begin{pmatrix} l_1 & l_2 & L \\ m_1 & -m_2 & M \end{pmatrix}, \tag{23}$$

where x stands for a, b, c, or d; ν for α, β, or γ; and ϵ is the extra phase which appears in (16) and (20). When the parent states are both Γ_1 representations, the vector

$$|\ \ \rangle = \sum_{M} N(L,M)|L\,M\rangle \tag{24}$$

is also a Γ_1 representation. Therefore, the values of L are 0 or 6, 10,...; $L=0$ is the only value allowed in the sum of (22) because $L\leqslant s_1+s_2$. When the parent representations are both Γ_7 representations, the vector (24) is a Γ_1 or a Γ_3 representation: the first nonzero value of L is 3, but it is not allowed in the sum of (22). Similarly, with two parent Γ_3 representations, the first nonzero value of L is 2 with (24) being a Γ_5 representation, but this value is not allowed in the sum of (22).

So, with two parent representations of the same kind,

$$\langle\Gamma_{i_1}\, j\, l_1\, \nu_1\, \hat{\sigma}_1|\Gamma_{i_2}\, j\, l_2\, \nu_2\, \hat{\sigma}_2\rangle = \delta_{s_1 s_2}\,\delta_{\sigma_1\sigma_2}\,\delta_{l_1 l_2}\,N(0,0),$$
$$N(0,0) = \sum x^{\nu_1}_{l_1 m}\, x^{\nu_2}_{l_1 m} = \delta_{\nu_1\nu_2}. \tag{25}$$

With two parent representations of different kinds, there is no term in (22) except when computing the overlaps between Γ_9, Γ'_9, and Γ''_9 representations. With a Γ_9 and a Γ'_9 representation, the vector (24) is a Γ_7 representation with

$L = \frac{7}{2}$. Therefore, the $N(L,M)$ are the components $b^{\pm}_{7/2,M}$ of the basic Γ_7 representation multiplied by $\sqrt{2l_2+1}$ and by the icosahedral $3jm$ symbol between the parent representations and this basic Γ_7 representation. Using the icosahedral 3Γ symbol (5), the overlap is

$$\langle \Gamma_9 j\, l_1\, \alpha\, \hat{\sigma} | \Gamma'_9\, j\, l_2\, \beta\, \hat{\sigma}\rangle = (-)^{l_1+j-1/2}\sqrt{\tfrac{8}{9}}\sqrt{(2l_1+1)(2l_2+1)} \times \begin{Bmatrix} \frac{5}{2} & 1 & \frac{7}{2} \\ l_2 & l_1 & j \end{Bmatrix} (\Gamma_1\, l_1\, \alpha \quad \Gamma_7\, l_2\, \beta \quad \Gamma_7\, 7/2). \tag{26}$$

The overlap between a Γ_9 and a Γ''_9 representation can be obtained in the same way:

$$\langle \Gamma_9 j\, l_1\, \alpha\, \hat{\sigma} | \Gamma''_9\, j\, l_2\, \gamma\, \hat{\sigma}\rangle = (-)^{l_1+j-1/2}\sqrt{6(2l_1+1)(2l_2+1)} \times \begin{Bmatrix} \frac{5}{2} & \frac{1}{2} & 3 \\ l_2 & l_1 & j \end{Bmatrix} (\Gamma_1\, l_1\, \alpha \quad \Gamma_3\, l_2\, \gamma \quad \Gamma_3\, 3). \tag{27}$$

The overlap between a Γ'_9 and a Γ''_9 representation involves the "basic" Γ_8 representation

$$\langle \Gamma'_9\, j\, l_1\, \beta\, \hat{\sigma} | \Gamma''_9\, j\, l_2\, \gamma\, \hat{\sigma}\rangle = (-)^{l_1+j}\sqrt{(2l_1+1)(2l_2+1)} \times \begin{Bmatrix} 1 & \frac{1}{2} & \frac{3}{2} \\ l_2 & l_1 & j \end{Bmatrix} (\Gamma_7\, l_1\, \beta \quad \Gamma_3\, l_2\, \gamma \quad \Gamma_8\, 3/2). \tag{28}$$

As seen here, some results involve other basic representations besides the Γ_1, Γ_7, or Γ_3 representations. For cubic harmonics, any product of the parent Γ_1, Γ_2, or Γ_3 representations is a Γ_1, Γ_2, or Γ_3 representation. Here, a product of parent representations can be any kind of representation, i.e., the set of parents is not a closed set.

Some other relations can be found between different kinds of representations, but they are not as simple as those given above. For example, there is a relation between the generating functions of Γ_7 and of Γ_1 representations:

$$g_7(x) = (1/x^7 + x^7)\, g_1(x) - 1/x^7 - x^5. \tag{29}$$

For j larger than or equal to $\frac{7}{2}$, the Γ_7 representations can be obtained from the Γ_1 representations for $l = j \pm \frac{7}{2}$. The Γ_7 representations obtained from the Γ_1 representations with $l = j - \frac{7}{2}$ are orthogonal to those obtained from the Γ_1 representations with $l = j + \frac{7}{2}$. However the Γ_7 representations obtained from several Γ_1 representations with the same l are not orthonormal: their overlap can be expressed with the icosahedral 3Γ symbol $(\Gamma_1\, l\alpha_1\, \Gamma_1\, l\alpha_2\, \Gamma_1 6)$; consequently, if these Γ_1 representations are chosen such that $(\Gamma_1\, l\alpha_1\, \Gamma_1\, l\alpha_2\, \Gamma_1 6) = \text{const} \times \delta_{\alpha_1,\alpha_2}$, the Γ_7 representations are orthogonal but their norm depends upon the values of $(\Gamma_1\, l\alpha\, \Gamma_1\, l\alpha\, \Gamma_1 6)$.

III. THE $3jm$ SYMBOLS OF THE ICOSAHEDRAL GROUP

As for the cubic group, the icosahedral $3jm$ symbols defined by (3) can be expressed in terms of those involving the parent Γ_1, Γ_7, or Γ_3 representations. As long as we avoid the components $\pm\hat{\frac{5}{2}}$ of Γ'_9 and Γ''_9 representations, this expression is

TABLE III. Decomposition of a product of irreducible representations.

Γ_1 Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	Γ_9
Γ_1 Γ_1	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	Γ_9
Γ_2 Γ_2	$\Gamma_4 + \Gamma_5$	$\Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_2 + \Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_6 + \Gamma_8$	Γ_9	$\Gamma_6 + \Gamma_8 + \Gamma_9$	$\Gamma_7 + \Gamma_8 + 2\Gamma_9$
Γ_3 Γ_3	$\Gamma_1 + \Gamma_3 + \Gamma_5$	$\Gamma_2 + \Gamma_4 + \Gamma_5$	$\Gamma_2 + \Gamma_3 + \Gamma_4 + \Gamma_5$	Γ_9	$\Gamma_7 + \Gamma_8$	$\Gamma_7 + \Gamma_8 + \Gamma_9$	$\Gamma_6 + \Gamma_8 + 2\Gamma_9$
Γ_4 Γ_4	$\Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_1 + \Gamma_2 + \Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_2 + \Gamma_3 + \Gamma_4 + 2\Gamma_5$	$\Gamma_7 + \Gamma_9$	$\Gamma_6 + \Gamma_9$	$\Gamma_8 + 2\Gamma_9$	$\Gamma_6 + \Gamma_7 + 2\Gamma_8 + 2\Gamma_9$
Γ_5 Γ_5	$\Gamma_2 + \Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_2 + \Gamma_3 + \Gamma_4 + 2\Gamma_5$	$\Gamma_1 + \Gamma_2 + \Gamma_3 + 2\Gamma_4 + 2\Gamma_5$	$\Gamma_8 + \Gamma_9$	$\Gamma_8 + \Gamma_9$	$\Gamma_6 + \Gamma_7 + \Gamma_8 + 2\Gamma_9$	$\Gamma_6 + \Gamma_7 + 2\Gamma_8 + 3\Gamma_9$
Γ_6 Γ_6	$\Gamma_6 + \Gamma_8$	$\Gamma_7 + \Gamma_9$	$\Gamma_8 + \Gamma_9$	$\Gamma_1 + \Gamma_2$	Γ_4	$\Gamma_2 + \Gamma_5$	$\Gamma_3 + \Gamma_4 + \Gamma_5$
Γ_7 Γ_7	Γ_9	$\Gamma_6 + \Gamma_9$	$\Gamma_6 + \Gamma_7 + \Gamma_8 + 2\Gamma_9$	Γ_4	$\Gamma_1 + \Gamma_3$	$\Gamma_3 + \Gamma_5$	$\Gamma_2 + \Gamma_4 + \Gamma_5$
Γ_8 Γ_8	$\Gamma_6 + \Gamma_8 + \Gamma_9$	$\Gamma_8 + 2\Gamma_9$	$\Gamma_6 + \Gamma_7 + \Gamma_8 + 2\Gamma_9$	$\Gamma_2 + \Gamma_5$	$\Gamma_3 + \Gamma_5$	$\Gamma_1 + \Gamma_2 + \Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_2 + \Gamma_3 + 2\Gamma_4 + 2\Gamma_5$
Γ_9 Γ_9	$\Gamma_7 + \Gamma_8 + 2\Gamma_9$	$\Gamma_6 + \Gamma_7 + 2\Gamma_8 + 2\Gamma_9$	$\Gamma_6 + \Gamma_7 + 2\Gamma_8 + 3\Gamma_9$	$\Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_2 + \Gamma_4 + \Gamma_5$	$\Gamma_2 + \Gamma_3 + 2\Gamma_4 + 2\Gamma_5$	$\Gamma_1 + 2\Gamma_2 + 2\Gamma_3 + 2\Gamma_4 + 3\Gamma_5$

$$\begin{pmatrix} \Gamma_{i_1} j_1 l_1 \nu_1 & \Gamma_{i_2} j_2 l_2 \nu_2 & \Gamma_{i_3} j_3 l_3 \nu_3 \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} = \sum_{k,L,M} A(s_1 s_2 s_3;\sigma_1 \sigma_2 \sigma_3;kLM) \sum_{L'} B(j_1 j_2 j_3 s_1 s_2 s_3 l_1 l_2;L'kL)$$
$$\times C(\Gamma_{p_1} l_1 \nu_1 \hat{m}_1, \Gamma_{p_2} l_2 \nu_2 \hat{m}_2, L'; \Gamma_{p_3} l_3 \nu_3 \hat{m}_3; LM), \tag{30}$$

where

$$A(s_1 s_2 s_3;\sigma_1 \sigma_2 \sigma_3;kLM) = (-)^{s_3+\sigma_3} \epsilon_1 \epsilon_2 \epsilon_3 \sqrt{2k+1} \begin{pmatrix} s_1 & s_2 & k \\ \sigma_1 & \sigma_2 & \kappa \end{pmatrix} \begin{pmatrix} s_3 & k & L \\ -\sigma_3 & \kappa & M \end{pmatrix} \tag{31}$$

is almost all the dependence on magnetic quantum numbers,

$$B(j_1 j_2 j_3 s_1 s_2 l_1 l_2;L'kL) = (-)^{k+L+j_3-l_3}(2L+1)\sqrt{(2k+1)}\sqrt{2L'+1}$$
$$\times\sqrt{(2s_1+1)(2s_2+1)(2s_3+1)(2l_1+1)(2l_2+1)(2l_3+1)}$$
$$\times \begin{Bmatrix} j_3 & k & L' \\ L & l_3 & s_3 \end{Bmatrix} \begin{Bmatrix} j_1 & j_2 & j_3 \\ s_1 & s_2 & k \\ l_1 & l_2 & L' \end{Bmatrix} \tag{32}$$

gathers the dependence on angular momenta and

$$C(\Gamma_{p_1} l_1 \nu_1 \hat{m}_1, \Gamma_{p_2} l_2 \nu_2 \hat{m}_2, L'; \Gamma_{p_3} l_3 \nu_3 \hat{m}_3; LM)$$
$$= \sqrt{2L'+1} \sum_{m_1 m_2 m_3 M'} (-)^{L'+M'} x^{\nu_1}_{l_1 m_1} x^{\nu_2}_{l_2 m_2} x^{\nu_3}_{l_3 m_3} \begin{pmatrix} l_1 & l_2 & L' \\ m_1 & m_2 & M' \end{pmatrix} \begin{pmatrix} L' & l_3 & L \\ -M' & m_3 & M \end{pmatrix}, \tag{33}$$

the Γ_p's are the parent representations of the Γ_i's and x^{ν}_{lm} are their coefficients. The possible values of (k,L) are limited by the triangular relations $|s_1 - s_2| \leqslant k \leqslant s_1 + s_2$ and $|k - s_3| \leqslant L \leqslant k + s_3$. The maximum values are 5 for k and 7 for L.

Except in some cases involving Γ'_9 and Γ''_9 representations, the number of different (k,L) values of the sum (29) is exactly the number of different basic icosahedral $3jm$ symbols (see Table III). For example, if the three parent representations are Γ_1 representations, the coefficients (33) are those of a Γ_1 representation. Therefore, the possible values of L are 0, 6, 10, ... The value $L = 0$ is always allowed by the triangular relations; for the icosahedral $3jm$ symbols of $\Gamma_2\Gamma_9\Gamma_9$, $\Gamma_5\Gamma_8\Gamma_9$, $\Gamma_5\Gamma_9\Gamma_9$, the value $L = 6$ is possible; for $\Gamma_5\Gamma_9\Gamma_9$, there are two possible values of k with $L = 6$. With an arbitrary combination of parent representations, the other possible values of L are $\frac{1}{2}$, 2, 3, $\frac{7}{2}$, 5, and 1, which appears only for $\Gamma_3\Gamma''_9\Gamma''_9$.

When $L = 0$, the coefficient (33) is the $3jm$ symbol between the parent states and $L' = l_3$, $k = s_3$. If the parent representations are Γ_1 representations, the coefficient (31) is essentially the $3jm$ symbol between basic representations and

$$(\Gamma_{i_1} j_1 l_1 \nu_1 \quad \Gamma_{i_2} j_2 l_2 \nu_2 \quad \Gamma_{i_3} j_3 l_3 \nu_3) = \sqrt{(2s_1+1)(2s_2+1)(2s_3+1)(2l_1+1)(2l_2+1)(2l_3+1)}$$
$$\times \begin{Bmatrix} j_1 & j_2 & j_3 \\ s_1 & s_2 & s_3 \\ l_1 & l_2 & l_3 \end{Bmatrix} (\Gamma_{p_1} l_1 \nu_1 \quad \Gamma_{p_2} l_2 \nu_2 \quad \Gamma_{p_3} l_3 \nu_3). \tag{34}$$

For $L \neq 0$ the expression (33) can be rewritten as

$$C(\Gamma_{p_1} l_1 \nu_1 \hat{m}_1, \quad \Gamma_{p_2} l_2 \nu_2 \hat{m}_2, \quad \Gamma_{p_3} l_3 \nu_3 \hat{m}_3; LM)$$
$$= \sqrt{2L'+1} \sum_{i',\nu'} (-)^{j_{i'}+\hat{m}'} \begin{pmatrix} \Gamma_{p_1} l_1 \nu_1 & \Gamma_{p_2} l_2 \nu_2 & \Gamma_{i'} L' \nu' \\ \hat{m}_1 & \hat{m}_2 & \hat{m}' \end{pmatrix} \begin{pmatrix} \Gamma_{i'} L' \nu' & \Gamma_{p_3} l_3 \nu_3 & \Gamma_i L \\ -\hat{m}' & \hat{m}_3 & \hat{M} \end{pmatrix} x_{L,M}, \tag{35}$$

where Γ_i stands for Γ_2, Γ_8, Γ_5, Γ_3, Γ_7, Γ_3, and Γ_1 when $L = 1, \frac{3}{2}, 2, 3, \frac{7}{2}, 5$, and 6, respectively, and the $x_{L,M}$ are the components of these representations. These limitations come from the fact that $L = 6$ can be obtained only with three Γ_1 parent representations, $L = 5$ can be obtained also with two Γ_1 and one Γ_5 representation, etc. The phase comes from the transformation of the components under a rotation of π around $0y$ (see Ref. 2): $j_{i'}$ is 0 for Γ_1 and Γ_5 representations, 1 for Γ_3, and $\frac{3}{2}$ for Γ_7 and Γ_8 representations. The sum over i' extends to all kinds of representations, which can be obtained by the product of the Γ_{p_1} and the Γ_{p_2} representations and by the product of the Γ_{p_3} and the Γ_i representations. This sum reduces to one term when at least one parent representation is a Γ_1 representation. For $\Gamma_3\Gamma'_9\Gamma''_9$, $\Gamma_4\Gamma'_9\Gamma''_9$, and $\Gamma'_9\Gamma''_9\Gamma''_3$ the $\Gamma_{i'}$ representations are Γ_7 and Γ_8, for $\Gamma_3\Gamma''_9\Gamma'_9$, they are Γ_3 and Γ_5. In these four cases there is only one set of values (k,L): the number of different (k,L,i') values is always the number of basic icosahedral $3jm$ symbols.

As the products of any two representations involved in

TABLE IV. Values of D used in formulas (36) and (37). The number in parentheses refers to the 3Γ symbol when it is not unique. When there is more than one term, values are given in sequence without repetition of the 3Γ symbols of the first columns (for example, $\Gamma_4\Gamma_8\Gamma_9$). The square of the coefficients and their signs are given.

$i_1i_2i_3$	k	L	i	i'	Coefficient
$\Gamma_1\Gamma_1\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_2\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_3\Gamma_3$	0	0	Γ_1	Γ_3	7
$\Gamma_1\Gamma_4\Gamma_4$	$\frac{1}{2}$	0	Γ_1	Γ_7	$2^7/7$
$\Gamma_1\Gamma_5\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_6\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_7\Gamma_7$	0	0	Γ_1	Γ_7	2^3
$\Gamma_1\Gamma_8\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_1\Gamma_9\Gamma_9$	$\frac{5}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_2\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_2\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_2\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_3\Gamma_4$	1	$\frac{3}{2}$	Γ_8	Γ_3	$-2^4.3$
$\Gamma_2\Gamma_3\Gamma_5$	1	3	Γ_3	Γ_3	7^2
$\Gamma_2\Gamma_4\Gamma_3$	$\frac{3}{2}$	$\frac{3}{2}$	Γ_8	Γ_7	$2^7.3/7$
$\Gamma_2\Gamma_4\Gamma_4$	$\frac{1}{2}$	0	Γ_1	Γ_7	$-2^5/7$
$\Gamma_2\Gamma_4\Gamma_5$	$\frac{3}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	$2^{10}/7$
$\Gamma_2\Gamma_5\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_5\Gamma_3$	3	3	Γ_3	Γ_1	7
$\Gamma_2\Gamma_5\Gamma_4$	3	$\frac{7}{2}$	Γ_7	Γ_1	$-2^7/7$
$\Gamma_2\Gamma_5\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_6\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_6\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_7\Gamma_9$	1	$\frac{7}{2}$	Γ_7	Γ_7	2^6
$\Gamma_2\Gamma_8\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_8\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_8\Gamma_9$	$\frac{5}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_9\Gamma_7$	$\frac{7}{2}$	$\frac{7}{2}$	Γ_7	Γ_1	2^3
$\Gamma_2\Gamma_9\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_9\Gamma_9$	$\frac{5}{2}$	0	Γ_1	Γ_1	1
$\Gamma_2\Gamma_9\Gamma_9(2)$	$\frac{7}{2}$	6	Γ_1	Γ_1	$-3.13/2^2$
$\Gamma_3\Gamma_3\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_3\Gamma_3\Gamma_3$	0	0	Γ_1	Γ_3	7
$\Gamma_3\Gamma_3\Gamma_5$	0	2	Γ_5	Γ_5	5^2
$\Gamma_3\Gamma_4\Gamma_2$	$\frac{1}{2}$	$\frac{3}{2}$	Γ_8	Γ_8	$2^6.3/7$
$\Gamma_3\Gamma_4\Gamma_4$	$\frac{1}{2}$	0	Γ_1	Γ_7	$2^6.11/7^2$
$\Gamma_3\Gamma_4\Gamma_5$	$\frac{1}{2}$	$\frac{3}{2}$	Γ_8	Γ_8	$2^6/7$
$\Gamma_3\Gamma_5\Gamma_2$	2	3	Γ_3	Γ_3	7^2
$\Gamma_3\Gamma_5\Gamma_3$	2	2	Γ_5	Γ_3	5.7
$\Gamma_3\Gamma_5\Gamma_4$	2	$\frac{3}{2}$	Γ_8	Γ_3	2^4
$\Gamma_3\Gamma_5\Gamma_5$	2	3	Γ_3	Γ_3	-7^2
$\Gamma_3\Gamma_6\Gamma_9$	$\frac{1}{2}$	3	Γ_3	Γ_3	-7^2
$\Gamma_3\Gamma_7\Gamma_7$	0	0	Γ_1	Γ_7	2^3
$\Gamma_3\Gamma_7\Gamma_8$	0	$\frac{3}{2}$	Γ_8	Γ_8	2^4
$\Gamma_3\Gamma_8\Gamma_7$	$\frac{3}{2}$	$\frac{3}{2}$	Γ_8	Γ_3	$-2^2.7$
$\Gamma_3\Gamma_8\Gamma_8$	$\frac{3}{2}$	3	Γ_3	Γ_3	-7^2
$\Gamma_3\Gamma_8\Gamma_9$	$\frac{3}{2}$	3	Γ_3	Γ_3	7^2
$\Gamma_3\Gamma_9\Gamma_6$	$\frac{5}{2}$	3	Γ_3	Γ_3	-7^2
$\Gamma_3\Gamma_9\Gamma_8$	$\frac{5}{2}$	3	Γ_3	Γ_3	7^2
$\Gamma_3\Gamma_9\Gamma_9$	$\frac{5}{2}$	3	Γ_3	Γ_3	-7^2
$\Gamma_3\Gamma_9\Gamma_9(2)$	$\frac{5}{2}$	3	Γ_3	Γ_3	$5.7^2/(2.3^3)$
	$\frac{5}{2}$	5	Γ_3	Γ_3	$7^3.11/(2.3^3)$
$\Gamma_4\Gamma_4\Gamma_1$	0	0	Γ_1	Γ_1	$2^4/7$
$\Gamma_4\Gamma_4\Gamma_2$	1	0	Γ_1	Γ_1	$-2^2/7$

TABLE IV. (*Continued*).

$i_1i_2i_3$	k	L	i	i'	Coefficient
$\Gamma_4\Gamma_4\Gamma_3$	0	0	Γ_1	Γ_3	$2^3.11/7$
$\Gamma_4\Gamma_4\Gamma_4(2)$	1	$\frac{3}{2}$	Γ_8	Γ_3	$-2^{11}/(3.7)$
$\Gamma_4\Gamma_4\Gamma_5$	1	3	Γ_3	Γ_3	$2^2.3.11$
$\Gamma_4\Gamma_5\Gamma_2$	$\frac{5}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	$2^{10}/7$
$\Gamma_4\Gamma_5\Gamma_3$	$\frac{3}{2}$	$\frac{3}{2}$	Γ_8	Γ_7	$2^7/7$
$\Gamma_4\Gamma_5\Gamma_4$	$\frac{5}{2}$	3	Γ_3	Γ_7	$-2^5.3.11/7$
$\Gamma_4\Gamma_5\Gamma_5$	$\frac{3}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	$-2^9/(5.7)$
	$\frac{5}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	$-2^9.3^2/(5.7)$
$\Gamma_4\Gamma_5\Gamma_5(2)$	$\frac{3}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	$-2^8.3^2/(5.7)$
	$\frac{5}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	$2^{14}/(3^2.5.7)$
$\Gamma_4\Gamma_6\Gamma_7$	0	0	Γ_1	Γ_7	2^3
$\Gamma_4\Gamma_6\Gamma_9$	1	$\frac{7}{2}$	Γ_7	Γ_7	$-2^{10}/7$
$\Gamma_4\Gamma_7\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	-1
$\Gamma_4\Gamma_7\Gamma_9$	$\frac{1}{2}$	3	Γ_3	Γ_3	7.11
$\Gamma_4\Gamma_8\Gamma_8$	2	$\frac{7}{2}$	Γ_7	Γ_7	$-2^{10}/7$
$\Gamma_4\Gamma_8\Gamma_9$	1	$\frac{7}{2}$	Γ_7	Γ_7	$2^5.5/7$
	2	$\frac{7}{2}$	Γ_7	Γ_7	$2^5.3^3/7$
$\Gamma_4\Gamma_8\Gamma_9(2)$	1	$\frac{7}{2}$	Γ_7	Γ_7	$2^7.3/7$
	2	$\frac{7}{2}$	Γ_7	Γ_7	$-2^7.5/7$
$\Gamma_4\Gamma_9\Gamma_6$	3	$\frac{7}{2}$	Γ_7	Γ_7	$-2^{10}/7$
$\Gamma_4\Gamma_9\Gamma_7$	3	3	Γ_3	Γ_7	$2^3.11$
$\Gamma_4\Gamma_9\Gamma_8$	2	$\frac{7}{2}$	Γ_7	Γ_7	$2^6/7$
	3	$\frac{7}{2}$	Γ_7	Γ_7	$2^6.3.5/7$
$\Gamma_4\Gamma_9\Gamma_8(2)$	2	$\frac{7}{2}$	Γ_7	Γ_7	$2^{13}.3/(5.7.11)$
	3	$\frac{7}{2}$	Γ_7	Γ_7	$-2^{11}/(7.11)$
$\Gamma_4\Gamma_9\Gamma_9$	2	$\frac{7}{2}$	Γ_7	Γ_7	$-2^9/(3.7)$
	3	$\frac{7}{2}$	Γ_7	Γ_7	$-2^9.5/(3.7)$
$\Gamma_4\Gamma_9\Gamma_9(2)$	2	$\frac{7}{2}$	Γ_7	Γ_7	$2^6.5/7$
	3	$\frac{7}{2}$	Γ_7	Γ_7	$-2^8/7$
$\Gamma_5\Gamma_5\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_5\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_5\Gamma_3$	3	3	Γ_3	Γ_1	7
$\Gamma_5\Gamma_5\Gamma_4$	3	$\frac{7}{2}$	Γ_7	Γ_1	$-2^7/7$
$\Gamma_5\Gamma_5\Gamma_4(2)$	4	$\frac{7}{2}$	Γ_7	Γ_1	$2^7/3^2$
$\Gamma_5\Gamma_5\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_5\Gamma_5(2)$	4	6	Γ_1	Γ_1	$5.13/3^2$
$\Gamma_5\Gamma_6\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_6\Gamma_9$	$\frac{5}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_7\Gamma_8$	2	$\frac{7}{2}$	Γ_7	Γ_7	2^6
$\Gamma_5\Gamma_7\Gamma_9$	2	$\frac{7}{2}$	Γ_7	Γ_7	2^6
$\Gamma_5\Gamma_8\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_8\Gamma_7$	$\frac{7}{2}$	$\frac{7}{2}$	Γ_7	Γ_1	2^3
$\Gamma_5\Gamma_8\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_8\Gamma_9$	$\frac{5}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_8\Gamma_9(2)$	$\frac{7}{2}$	6	Γ_1	Γ_1	$-3.13/2^2$
$\Gamma_5\Gamma_9\Gamma_6$	$\frac{1}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_9\Gamma_7$	$\frac{7}{2}$	7/2	Γ_7	Γ_1	2^3
$\Gamma_5\Gamma_9\Gamma_8$	$\frac{3}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_9\Gamma_8(2)$	$\frac{9}{2}$	6	Γ_1	Γ_1	$-2.13/5$
$\Gamma_5\Gamma_9\Gamma_9$	$\frac{5}{2}$	0	Γ_1	Γ_1	1
$\Gamma_5\Gamma_9\Gamma_9(2)$	$\frac{7}{2}$	6	Γ_1	Γ_1	$-3.13/2^2$
$\Gamma_5\Gamma_9\Gamma_9(3)$	$\frac{9}{2}$	6	Γ_1	Γ_1	$3.13/5$
$\Gamma_6\Gamma_6\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_6\Gamma_6\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_6\Gamma_7\Gamma_4$	$\frac{1}{2}$	0	Γ_1	Γ_7	2^3
$\Gamma_6\Gamma_8\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_6\Gamma_8\Gamma_5$	2	0	Γ_1	Γ_1	1

TABLE IV. (*Continued*).

$i_1 i_2 i_3$	k	L	i	i'	Coefficient
$\Gamma_6\Gamma_9\Gamma_3$	3	3	Γ_3	Γ_1	7
$\Gamma_6\Gamma_9\Gamma_4$	3	$\frac{7}{2}$	Γ_7	Γ_1	$-2^7/7$
$\Gamma_6\Gamma_9\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_7\Gamma_7\Gamma_1$	0	0	Γ_1	Γ_1	Γ_1
$\Gamma_7\Gamma_7\Gamma_3$	0	0	Γ_1	Γ_3	7
$\Gamma_7\Gamma_8\Gamma_3$	$\frac{3}{2}$	$\frac{3}{2}$	Γ_8	Γ_7	2^5
$\Gamma_7\Gamma_8\Gamma_5$	$\frac{3}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	-2^6
$\Gamma_7\Gamma_9\Gamma_2$	$\frac{5}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	-2^6
$\Gamma_7\Gamma_9\Gamma_4$	$\frac{5}{2}$	3	Γ_3	Γ_7	$2^3.11$
$\Gamma_7\Gamma_9\Gamma_5$	$\frac{5}{2}$	$\frac{7}{2}$	Γ_7	Γ_7	2^6
$\Gamma_8\Gamma_8\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_8\Gamma_8\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_8\Gamma_8\Gamma_3$	3	3	Γ_3	Γ_1	7
$\Gamma_8\Gamma_8\Gamma_4$	3	$\frac{7}{2}$	Γ_7	Γ_1	$-2^7/7$
$\Gamma_8\Gamma_8\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_8\Gamma_9\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_8\Gamma_9\Gamma_3$	3	3	Γ_3	Γ_1	7
$\Gamma_8\Gamma_9\Gamma_4$	3	$\frac{7}{2}$	Γ_7	Γ_1	$-2^7/7$
$\Gamma_8\Gamma_9\Gamma_4(2)$	4	$\frac{7}{2}$	Γ_7	Γ_1	$2^7/3^2$
$\Gamma_8\Gamma_9\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_8\Gamma_9\Gamma_5(2)$	4	6	Γ_1	Γ_1	$5.13/3^2$
$\Gamma_9\Gamma_9\Gamma_1$	0	0	Γ_1	Γ_1	1
$\Gamma_9\Gamma_9\Gamma_2$	1	0	Γ_1	Γ_1	1
$\Gamma_9\Gamma_9\Gamma_2(2)$	5	6	Γ_1	Γ_1	3.13/11
$\Gamma_9\Gamma_9\Gamma_3$	3	3	Γ_3	Γ_1	7
$\Gamma_9\Gamma_9\Gamma_3(2)$	5	5	Γ_3	Γ_1	7
$\Gamma_9\Gamma_9\Gamma_4$	3	$\frac{7}{2}$	Γ_7	Γ_1	$-2^7/7$
$\Gamma_9\Gamma_9\Gamma_4(2)$	4	$\frac{7}{2}$	Γ_7	Γ_1	$2^7/3^2$
$\Gamma_9\Gamma_9\Gamma_5$	2	0	Γ_1	Γ_1	1
$\Gamma_9\Gamma_9\Gamma_5(2)$	4	6	Γ_1	Γ_1	$5.13/3^2$
$\Gamma_9\Gamma_9\Gamma_5(3)$	5	6	Γ_1	Γ_1	$-5.13/11$

the $3jm$ symbols of (35) give only one representation of the third kind, the 3Γ symbols are unique and we can use them to factorize the dependence on magnetic quantum numbers. We obtain

$$
\begin{aligned}
&(\Gamma_{i_1} j_1 l_1 \nu_1 \quad \Gamma_{i_2} j_2 l_2 \nu_2 \quad \Gamma_{i_3} j_3 l_3 \nu_3)^{(n)} \\
&\quad = \sum_{L,k,i'} D^{(n)}(i_1,i_2,i_3;kLi') \\
&\qquad \times \sum_{L'} \sqrt{(2L'+1)}\, B(j_1 j_2 j_3 s_1 s_2 l_1 l_2; L'kL) \\
&\qquad \times \sum_{\nu'} (\Gamma_{p_1} l_1 \nu_1 \quad \Gamma_{p_2} l_2 \nu_2 \quad \Gamma_{i'} L' \nu') \\
&\qquad \times (\Gamma_{i'} L' \nu' \quad \Gamma_{p_3} l_3 \nu_3 \quad \Gamma_i L), \qquad (36)
\end{aligned}
$$

where $(\quad)^{(n)}$ is one of the 3Γ symbols defined by (5) and

$$
\begin{aligned}
&\sum_n D^{(n)}(i_1 i_2 i_3; kLi') \begin{pmatrix} \gamma_{i_1} & \gamma_{i_2} & \gamma_{i_3}{}^{(n)} \\ \hat{\sigma}_1 & \hat{\sigma}_2 & \hat{\sigma}_3 \end{pmatrix} \\
&\quad = (-)^{j'_i + \hat{m}' + s_3 - \sigma_3} \epsilon_1 \epsilon_2 \epsilon_3 \\
&\qquad \times \begin{pmatrix} s_1 & s_2 & k \\ \sigma_1 & \sigma_2 & \kappa \end{pmatrix} \begin{pmatrix} s_3 & k & L \\ \sigma_3 & \kappa & M \end{pmatrix} \\
&\qquad \times \begin{pmatrix} \gamma_{p_1} & \gamma_{p_2} & \gamma_{i'} \\ \hat{m}_1 & \hat{m}_2 & \hat{m}' \end{pmatrix} \\
&\qquad \times \begin{pmatrix} \gamma_{i'} & \gamma_{p_3} & \gamma_i \\ -\hat{m}' & \hat{m}_2 & \hat{M} \end{pmatrix} x_{L,M}\,. \qquad (37)
\end{aligned}
$$

These coefficients have been computed for all icosahedral $3jm$ symbols and are given in Table IV. However, due to the length of the table, the Γ'_9 and Γ''_9 representations have not been included in this table; the coefficients which vanish are not included. In most cases, there is only one couple of values (k,L) for each 3Γ symbol. When one of the first two parent representations is a Γ_3 or Γ_7 representation, two of them are present but a permutation can avoid these difficulties. The intermediate representations Γ'_i are Γ_1, Γ_3, or Γ_7 representations, except for $(\Gamma_3\Gamma_3\Gamma_5)$ for which it is a Γ_5, and $(\Gamma_3\Gamma_4\Gamma_2)$, $(\Gamma_3\Gamma_4\Gamma_5)$, $(\Gamma_3\Gamma_7\Gamma_8)$ for which it is a Γ_8 representation, but this disappears by circular permutation.

So, all the icosahedral $3jm$ symbols and the 3Γ symbols can be expressed with two 3Γ symbols: one involving two parent representations and a neighboring Γ_1, Γ_3, or Γ_7 representation of the third representation; the second involving this third representation, the intermediate representation, and one of the two lowest Γ_1, Γ_3, or the lowest Γ_7, Γ_8, or Γ_5 representations.

IV. CONCLUSIONS

As in the case of cubic harmonics, the labeling system presented here generates some properties. One of the most important is the definition of a "fictitious spin" for Γ_2, Γ_4, Γ_5, Γ_6, and Γ_8 representations. Table III shows that the product of these representations with a vector that is a Γ_2 representation includes a representation of the same kind. For Γ_9 representations, there are two "fictitious spins" and for Γ_3 and Γ_7 representations, there are none.

The matrix elements of the vector **J** between two representations are the icosahedral $3jm$ symbol of these two representations with the Γ_2 representation for $j=1$ (the parent representation of which is the Γ_1 representation with $l=0$) multiplied by the reduced matrix element $\langle j_1 \| \mathbf{J} \| j_3 \rangle = \delta_{j_1 j_3} \times \sqrt{j_1(j_1+1)(2j_1+1)}$. The fictitious spin is proportional to the overlap between parent representations; therefore, it is diagonal. For the Γ_2, Γ_5, Γ_6, and Γ_8 representations, when divided by its value for the basic representation, it is

$$\alpha = \tfrac{1}{2}\,[j(j+1) + s(s+1) - l(l+1)]/s(s+1). \qquad (38)$$

However, for a Γ_4 representation, when divided by the value for the basic representation ($j=3$, $s=\frac{1}{2}$, $l=\frac{7}{2}$) we obtain

$$\alpha = -\tfrac{1}{3}\,\{j(j+1) + \tfrac{3}{4} - l(l+1)\}. \qquad (39)$$

In Table V we give the components of all the nondegenerate Γ_1 representations. We could not obtain simple choices of the doubly degenerate representations for $l=30$, 36, 40, etc. This table can be used up to $l=29$ to build the other representations. Values for higher values of l are given because they are well-defined constants.

Table VI gives the coefficients of nondegenerate Γ_7 representations, and Tables VII and VIII give the coefficients

TABLE V. Coefficients of nondegenerate Γ_1 representations. The coefficient is the square root of the quotient of prime numbers given in the table by those of the second column, with the sign given in the table. Numbers between parentheses are not in the square root. Base vectors are symmetrized as in a previous work.[2] There is no Γ_1 representation for missing values of l smaller than 30; there are two of them for missing values larger than or equal to 30.

l	Den	$[0+\rangle$ $[20+\rangle$	$[5+\rangle$ $[25+\rangle$	$[10+\rangle$ $[30+\rangle$	$[15+\rangle$ $[35+\rangle$
0	1	1			
6	5^2	11	2.7		
10	3.5^4	13.19	$-2.3.11.19$	2.11.17	
12	5^5	$3^2.7.17$	$2^2.11.13$	2.3.13.19	
15	2.5^5		$-5.23.29$	2.3.11.29	7.11.13
16	$2.3.5^7$	$2^5.5.19.31$	$-3.5.13.17.31$	$-2.7.17.23.31$	$3^2.17.23.29$
18	5^7	5.11.17.23	$2^4.3^2.5.19$	2.3.7.11.19	2.19.29.31
20	5^9	5.7.23.29 2.11.13.31.37	$-2^2.11.17.19.29$	2.17.19.(41)	$-2^4.11.19.31$
21	5^7	17.19.37	$-29.31.41$	$-2^2.13.41$	3.13.17.41
22	3.5^{10}	2.3.5.11.19.31.37 13.19.23.29.41	$-7.13.23.31.37$	$-2.7.17.23.29.37$	$-23.29.(103)$
24	5^{10}	$5.7^2.13.23.29$ 2.31.37.41.43	$2^4.3.11.(59)$	$2^2.19^3.31$	$2^3.3^2.11.31.37$
25	3.5^{10}	$3^2.7^2.11.41$	$-7^2.31.37.43$ 2.5.11.23.41.47	2.3.7.11.19.37.43	$-7^2.43.(61)$
26	3.5^{12}	$2.3^3.13.29.31.41$ $-5.37.43.(89)$	$-5.7.11.23^3.41$ $2.7^2.17.37.43.47$	$2^7.5.7.19.23.41$	3.5.23.37.(139)
27	$2.3.5^{10}$	$2^6.17.19.23.43$	$-17.19.37.41.47$ 5.7.13.19.23.43	$-2.41^3.47$	2.3.17.19.47
28	$2.3.5^{12}$	$2^7.3^2.7.31.37.43$ $-2^3.3^2.5.13.29.41.47$	$-5.11.13.29.37.43$ 7.17.29.41.47.53	$-2.3.5.13.17.23.29.43$	$-2.5.23.29^3.41$
31	$2.3.5^{12}$	2.13.47.61.(79)	$2^3.7.37.41.43.61$ $-2.3.5^2.47.53.61$	$-5.7.11.19.23.43.61$ $-3.19.29.47.53.59$	$-2.5.7.17.31^2.61$
32	$2.3.5^{15}$	$2^5.11.17.29.37.41.47$ $2.3^2.7.31.43.(463)$	$-2^3.3^2.31.41.47.(103)$ $-2.3.7^4.19.31.43.53$	$3^4.5.7.19.23.31.47$ 7.19.29.43.53.59.61	2.3.5.31.43.(953)
33	2.5^{12}	$-2.7^2.13.19^3$	$2^2.3^2.41.43.47.53$ $-2^3.3.5^2.7.29.37^2$	5.47.53.(79) $-3.7.29.31.59.61$	$2^8.5.7.19.53$
34	3.5^{15}	$3.7^2.11.17.19.31.37.43$ $-2^4.13.19.29.41.47.53$	$-2^3.7.13.43.59^2$ $-2^2.23^2.41.47.53.59$	$-2^4.3^2.5.7.11^3.29.41$ 2.31.41.47.53.59.61	$-2.3.5.29.41.47.(79)$
35	$2.3.5^{15}$	$-2^3.3^2.11.19.29.41^2.59$	7.41.43.47.53.59 $5^2.13.(3001)$	$-2.3.7.11.13.29.47.53.59$ $-2^4.3.11^2.31.61$	$2.3^2.13.29.53.59.61^2$ $-2^3.7.11.17.23.31.61.67$
37	3.5^{15}	$-17.31.53.61.67$	$-7.11.29.43.47.61.67$ $-2.5^2.7.13.29.53.59.67$	$2^5.3.7^2.23.31.61.67$ $2^2.29.53.59.(79)$	$3.17.31.47^2.61.67$ 2.3.17.23.29.53.59.71
38	3.5^{17}	2.3.13.19.37.41.43.47.53 $7^3.31.(6269)$	$-3.17.47.53.(1933)$ 2.7.17.59.61.(563)	$-2.11^2.17.23.29.31.53$ $-2^3.3^2.7.47^2.59.61.67$	$3^5.7^3.11^2.23.29.31$ 2.7.23.59.61.67.71.73
39	$2.3.5^{15}$	$-2^6.19.23.(79)$	$-2.3.7.11.23.31.47.53.59$ $5^2.7^3.17.23.31.61$	$-2^3.7.17.53.59.(107)$ 2.3.31.61.67.(283)	$-2^2.3.29.59.(659)$ 31.37.61.67.71.73

TABLE V. (*Continued*).

l	Den	$[0+\rangle$ $[20+\rangle$ $[40+\rangle$	$[5+\rangle$ $[25+\rangle$ $[45+\rangle$	$[10+\rangle$ $[30+\rangle$ $[50+\rangle$	$[15+\rangle$ $[35+\rangle$ $[55+\rangle$
41	3.5^{17}	$-5.11.19.23.31.53^2.71$ $2^2.3^3.13.19.37.67.73.79$	$-3.5.7.47.53.59.61.71$ $2^3.5^3.13.17.23.43^2.71$	$2.11.41^2.53.59.61.71$ $-2^3.3.5.41^2.67.79^2$	$-3.7^2.13.29.31.59.61.71$ $2.7.11.19.37.67.73$
43	3.5^{17}	$2^2.5.7.13^3.59.67.73$ $3^2.37.41.59.71.79.83$	$-2.5.7.13.37.53.61.67.73$ $-2^4.5^3.17.19.23.59.73$	$3.19.61.67.73.(89)$ $-5.19.29^2.43^2.59.71$	$2^{11}.3.7.31.61.67.73$ $2.3.7.37.59.71.(109)$
44	3.5^{20}	$2.3^2.5.7.11.23.41.43.47.53.59$ $2.3^2.5.13.37.61.(3623)$ $3^2.41.61.67.71.73.79.83$	$-2.5.53.59.(37\,573)$ $2^8.5.7.17.37.43^2.61.67$	$-3.7.13^2.17.19.37.47^2.59$ $3.5.19.59^2.61.67.71.73$	$2^5.3^2.13^2.29.31.37.79^2$ $-2.61.67.71.73.79.(103)$
47	2.5^{20}	$-2^5.7.13.17.37.41.71.79^2$ $2^4.11.29.43.73.79.83$	$-3^4.11.13.53.59.61.67.71$ $2.3.5^3.37.41.(293)$ $2.13.23.29.43.73.79.83.89$	$2.41.59.61.67.71.(179)$ $2^2.3.7^2.31^2.41.73.(173)$	$-7.11.17.29.31.37.41.67.71$ $-2^3.3^3.7.17.73.79.(271)$
49	$2.3.5^{20}$	$2.23.37.73.79.(2243)$ $2^2.3.13^3.41.43.71.83.89$	$-2.29.41.43.59.61.67.73.79$ $2.3.5^3.7^2.13.23.29.71.79$ $3.23.31.41.43.47.71.83.89$	$2^3.19.47^2.61.67.73.79$ $-2^2.3.19.29.71.(8329)$	$2^3.13.31.37.43^2.67.73.79$ $-3.7.17.29.41.71.83.(233)$
53	3.5^{22}	$-2^5.19.37.41.43.83.(571)$ $-3.13.19.79.89.(10\,133)$	$-2^2.31.47.59.61.67.71.73.83$ $-2.5.7.13.29.31.41.43.59^2.83$ $2^4.3^3.7^2.11.47.79.89.97$	$7.11.23.67.71.73.83.(173)$ $3^2.29.31.43.79.(2917)$ $2.3.5.7.17.47.79.89.97.101.103$	$-2^3.3.17.23.41.43.71.73.83$ $2^{13}.3.17.19.23.31.59^2.79$
59	2.5^{25}	$-2^3.7.13^3.19.37.41.43.47.83.89$ $2^5.5.19.97.(586\,471)$	$-2.11.17^2.53.67.71.73.79.83.89$ $-2^5.5^2.7.43.47.89.(8069)$ $-3^2.13.89^2.97.101.103.(239)$	$2^2.5.23.71.73.79.83.89.(271)$ $-2^2.3.7.29.31.47.(139.239)$ $2.3.5^2.11.53.97.101.103.107.109$	$3.5.7.37.47.79.83.89.(239)$ $2.3.5.7^3.23.(557.757)$ $2.19.37.53.97.101.103.107.109.113$

TABLE VI. Coefficients of nondegenerate Γ_7 representations. Same presentation as in Table V.

l	Den	$\lvert 3/2\rangle$ $\lvert 33/2\rangle$	$\lvert -7/2\rangle$ $\lvert -37/2\rangle$	$\lvert 13/2\rangle$ $\lvert 43/2\rangle$	$\lvert -17/2\rangle$ $\lvert -47/2\rangle$	$\lvert 23/2\rangle$ $\lvert 53/2\rangle$	$\lvert -27/2\rangle$ $\lvert -57/2\rangle$
7/2	2.5	7	3				
13/2	2.5^2	13	2.13	-11			
17/2	$2.3.5^3$	11.17	$-2.7.17$	-13.17	$2^3.13$		
19/2	2.5^4	2.3.7.13	2^2	7^3	$-3.7.17$		
23/2	$2.3.5^5$	-17.23	3.7.17.23	$2.3^4.23$	$2^3.19.23$	2.7.11.19	
25/2	$2.3.5^5$	$2^3.3.11.19$	$13^2.19$	$2^2.7.17$	$-2.13^2.17$	$-11.17.23$	
27/2	$2.3.5^6$	2.3.5.17.19	$2^3.5.11.19$	$-5.11.(29)$	5.(47)	$2.7^2.23$	$2.3^3.13.23$
29/2	$2.3.5^5$	$-2.5.17.23$	3.5.13.23	$-2.5.13.19$	0	$17^2.19$	$2.3^2.7.19$
31/2	2.5^6	$3^2.5.13.19$	-5.7	5.31^2	$-3.5.23$	3.7.13.23	13.23.29
33/2	$2.3.5^7$	2.3.5.19.23 $-2^3.11.29.31$	$3^4.5.7.23$	$-2^7.3^2.5.13$	$2.3^3.13.17^2$	3.(89)	$2.13^2.29$
35/2	$2.3.5^7$	$-2.5.23.29$ $2^2.3.11.17.31$	$2^3.5.7.23.29$	5.11.13.29	$-23^2.29$	2.3.(109)	$-2.3.23^2.31$
39/2	2.5^9	5.11.17.23.31 $2.3.7^2.17.29$	5.31.(43) $-2.17.19.29.37$	$-2.7.11^2.13.31$	$2^7.31$	$-2.3.11.23^2.29$	$-2^{11}.7.29$
41/2	$2.3.5^8$	2.3.5.19.23.29 $-2.17^2.31.37$	$-2^3.3^2.5.7^2.29$ 13.19.31.37	$2^2.13.17.29$	$2^3.17.(43)$	$-2^2.11.17.31$	$-3.31.(47)$
45/2	$2.3.5^{10}$	$-3^2.5.7.13.29.31$ $-2^{15}.3.37$	$-2.3.29.31.43^2$ $2^3.13^2.37.41$	$3^4.7.17.19.31$ 2.7.11.37.41.43	$-2^3.7.17^3.19$	$2^2.11.13^3.19$	$2.3^2.7.13.19.23^2$
51/2	$2.3.5^{11}$	$2^2.3.7.23.29.31.37$ $-2.3.11.41.(131)$	2.31.37.(79) 3.41.43.(71)	$-2^2.7.11.37.41^2$ 7.23.41.43.47	$-2^2.3.7.17.19.37$ $2^2.7.23.41.43.47$	$-19.23^2.(71)$	$-3^5.7.47^2$

TABLE VII. Coefficients of component $|\hat{0}-\rangle$ of nondegenerate Γ_3 representations. Same presentation as in Table V.

l	Den	$[0-\rangle$	$[5-\rangle$	$[10-\rangle$	$[15-\rangle$
3	1	1			
5	5^2	-7	2.3^2		
7	5^3	$2^4.3$	-7.11		
8	1		1		
9	5^3	$3^2.11$	2.13		
10	5^3		$2^2.17$	3.19	
11	5^5	$-2.3.11.13$	-2.3	7.17.19	
12	5^3		2.3.19	-11	
14	5^3		3.11	$2^2.23$	
16	2.5^5		5.13.23	$2.3.7^3$	3.29.31

TABLE VIII. Coefficients of component $|\hat{2}\rangle$ of nondegenerate Γ_3 representations. Same presentation as in Table V.

l	Den	$\lvert 2\rangle$	$\lvert -3\rangle$	$\lvert 7\rangle$	$\lvert -8\rangle$	$\lvert 12\rangle$	$\lvert -13\rangle$
3	5	3	-2				
5	5	3	2				
7	$2^2.5^3$	-2.7	7^3	11.13			
8	$2^2.5^3$	$-2.11.13$	-3.13	$3^2.7$	$2^4.7$		
9	5^4	2^7	$-2.3.7$	$2.3^2.13$	13.17		
10	2.5^4	13.17	$-2^5.17$	-2.13^2	-3.7^2		
11	2.5^4	(23)	$2.3^2.7$	$2^4.17$	-17.19		
12	2.5^5	17.19	2.3.17.19	$-2^6.7$	7.17^2	$-2.3.11.23$	
14	$2^2.5^5$	$-2^2.5.17.19$	$-3.5.19$	$-2.5.7.11$	$2^2.5^3$	2.3.13.23	$-3^2.13.23$
16	2.5^6	2.5.19.23	$-2^4.5.7.23$	-5.13	$2^3.3.5.13$	$-2^4.3.13^2$	$3.7^2.29$

TABLE IX. Coefficients of degenerate Γ_7 representations up to $l = 27.5$. Other linear combinations can be found. These are obtained by making some coefficient of the second representation vanish in order to get the simplest values.

l	Denominator	$\lvert 3/2\rangle$ $\lvert 23/2\rangle$ $\lvert 43/2\rangle$	$\lvert -7/2\rangle$ $\lvert -27/2\rangle$ $\lvert -47/2\rangle$	$\lvert 13/2\rangle$ $\lvert 33/2\rangle$ $\lvert 53/2\rangle$	$\lvert -17/2\rangle$ $\lvert -37/2\rangle$
37/2	$2.5^8.139^2$	$2^3.3.5.7.11.17.(107)$	$5.(139)^2$	$2^2.7.13.17^2.23.(41)$	$-2.7.11.23.(109)$
$\beta=1$		$-3.11^3.17^2.23.29$	$2^8.3^2.7.11.23.29.31$	$-2^3.3^6.17.23.29.31$	$-2^6.17.23.29.31.37$
37/2	$2.3.5^6.139^2$	$-2.3.5.13.23.29.31.37$	0	$2^2.3^2.11.17.29.31.37$	$2^3.3^2.13.17.29.31.37$
$\beta=2$		$-2^6.3.7.13.17.31.37$	$-13.17.37.(71)$	$2.7^5.11.13.37$	$-3^2.7.11.13.(149)$
43/2	$2.3.5^{10}.4493$	$2.3^4.5.7^3.13.17.19.37$	$2^2.3.7.13.17.37.41^2$	$3^4.37.(4493)$	$7.29.37.43^2.(47)$
$\beta=1$		$2^5.7.13.29.31.37$	$2^9.3^2.13.17.29.31.37$	$2^2.3.7.11^2.13.17.19.29.31$	$-2^3.7^5.17.19.29.31$
		$2^5.17.19.29.31.41.43$			
43/2	$2.5^7.4493$	$-2^2.5.11^2.29.31.43$	$2.3^3.19.29.31.43$	0	$2^5.13.17.19.31.43$
$\beta=2$		$-17.19.43.(107)$	$7^3.13^2.19.43$	$-2.3.7^4.37.43$	$5^2.13^3.37.43$
		$7.13.23^2.37.41$			
47/2	$2^3.3.5^{11}.17.727$	$2.3^4.13.29.31.37.41.47$	$2^5.7.11.29.31.37.41.47$	$-2.3.11.19.29^2.31.37.41.47$	$3.11.17.19.37.41.47^3$
$\beta=1$		$2^2.3.7^2.13.19.23^2.37.41.47$	$11.13.19.41.47.(521)$	$-3.11.41.47.(3571)$	$2^2.3^7.11.47.(107)$
		$3^2.43.47.(307)$	$2^4.17^2.23.43.(727)$		
47/2	$2^3.3.5^8.17.727$	$-2.3.5^2.11.19.29^3$	$2^9.3.7.13.19.23^2.29$	$2.11^4.13.(71)$	$-13.17.23^2.31.(53)$
$\beta=2$		$2^2.11.31.(2141)$	$3.7^2.23^2.31.37$	$13.19.31.37.(229)$	$2^2.13.19.23^2.31.37.41$
		$3.11.13.19.31.37.41.43$	0		
49/2	$2.3.5^{10}.13.67.149$	$3^3.5^2.23.29.31.37.(61)$	$-2.7.11.29^3.31.37$	$-11.19.37.(4721)$	$-2^2.3.17.19.37.(2797)$
$\beta=1$		$2^2.7^2.19.37.(157)$	$2.3.13^3.(67.149)$	$2^3.3^2.11^3.13^2.37^2.41$	$-2^3.3.11^3.41.43.(79)$
		$-2.23.37^2.41.43.(107)$	$-2^2.5^2.23.41.43.47.(53)$		
49/2	$2.5^{10}.13.67.149$	$-2^2.11^3.17.19.31.41.43$	$-2^5.3.7.17.19.23.31.41.43$	$-2^2.3.17.23.29.41.43.(59)$	$-11.23.29.41.43$
$\beta=2$		$-2^2.3.5.11.17.23.29.41.43$	0	$2.3.13^2.17.19.23.29.37.43$	$2^3.17.19.23.29.37.(89)$
		$2^7.3.11^3.17.19.29.37$	$-3^3.7^2.11.17.19.29.37.47$		
53/2	$2^4.3.5^{12}.13.1327$	$-2.7.23^3.29.31.37.41.43$	$2^{11}.5.7.7.23.31.37.41.43$	$2.3^4.5.37.41.43.(163)$	$-5.7.17.19.37.41.43.(293)$
$\beta=1$		$-2^4.5.7^3.11^2.41.43.(59)$	$-5.13.41.43.(79.193)$	$2^7.5.7.11.13^2.(1327)$	$-5.7.37^2.(20507)$
		$2^3.5.23.47.(16\,693)$	$7^2.23.47.(10259)$	$2.3^2.7^2.13.17.23.47^3.53$	
53/2	$2^4.3.5^{11}.13.1327$	$-2.29.47.53.(4519)$	$2^9.3^4.5.47.53.(127)$	$-2.5.7.23.31.47.53.(131)$	$-5^3.17.19.23.31.47.53$
$\beta=2$		$2^4.5.17^2.23.31.37.47.53$	$5.7.13.19^2.23.31.37.47.53$	0	$-5.23.31.37.41.43.47.53$
		$2^3.5.7.17^2.31.37.41.43.53$	$3^2.7.23^2.31.37.41.43.53$	$-2.7.13.17^3.31.37.41.43$	
55/2	$2.5^{12}.29.401$	$2^2.3^2.31.37.41.(337)$	$-2^2.3.5.7^2.13.37.41.(107)$	$-3.5.13.17.23.37.41.43^2$	$-5.13.17.23.37.41.(181)$
$\beta=1$		$3.5.23.41.(3203)$	$2.3^2.5.23.29^2.(401)$	$2.3^2.5.11.13.23.43.(257)$	$2.5.13^3.43.(1013)$
		$-3.5.7.13.19^2.43.47$	$3^3.7.13.17.43.47.53^2$	$-3.7.17.43.47.53.(151)$	
55/2	$2.3.5^{11}.29.401$	$3.7.13.17.23.31.43.(71)$	$-5.7^3.17^3.23.43$	$2^2.5.7.43.(10\,771)$	$2^2.3.5.7^3.11^2.23^2.43$
$\beta=2$		$2^2.5^3.7.13^3.17.37.43$	0	$2.3.5.7.11.17.37.41^3$	$-2.3.5.7.17.23.37.41.43^2$
		$2^{12}.5.17.23.37.41.47$	$2^8.23.37.41.43^2.47$	$2^2.3^2.13.23.37.41.47.53$	

TABLE X. Coefficients of component $|\hat{0}\ -\rangle$ of degenerate Γ_3 representations up to $l = 27$. See next table.

l	Denominator		[0 − ⟩ [20 − ⟩	[5 − ⟩ [25 − ⟩	[10 − ⟩	[15 − ⟩
13	$5^4.47.127$	($\gamma=1$)	$2^{10}.7.11.13$	17.(73)	$-2^5.11.17.19.23$	
13	$5^5.47.127$	($\gamma=2$)	$2.17^3.19.23$	$-2^5.7.11.13.19.23$	$3^4.7^3.13$	
15	$2^2.5^7.11$	($\gamma=1$)	$-2.3^4.5.7.13.17$	$3^4.5.11.19$	$-2.3.19^3.23$	7.13.19.23.29
15	$2^2.5^6.11$	($\gamma=2$)	2.3.5.17.19.23	3.5.7.11.13.23	2.7.13	3.29.(37)
17	$2^4.3.5^8$	($\gamma=1$)	$3^2.5.19.23.29$	$-5.11.13.17.23.29$	$2^2.3.17.29.(41)$	$-2.11^2.17.31$
17	$2^4.3.5^6$	($\gamma=2$)	$-3^2.5.13.17.19$	$-5.11.19^2$	$2^2.3.13.23$	2.13.23.29.31
18	$5^5.107$	($\gamma=1$)		$-2.3^2.5.29.31$	$-3.7.11.29.31$	$2^2.(107)$
18	107	($\gamma=2$)		-7.11	2.3.5	0
19	$5^7.47.71$	($\gamma=1$)	2.3.5.7.11.17.19.31	$-2.3^2.5.23.31.(43)$	$-11.13^2.23.29.31$	$2.3.7.13^2.17.23.29$
19	$5^8.47.71$	($\gamma=2$)	$2^2.3.5.19.23.29^3$	$-2^4.5.7.11.17.29$	$-2.3^4.7.17.(79)$	$-3.11.31.(727)$
20	$5^8.7.11^2$			$2^4.3^4.29.31.37$	2.11.31.37.(59)	17.37.(563)
$\gamma=1$			$-2^7.13.17.19$			
20	$5^7.7.11^2$			$2.3^2.17.(179)$	$-11.17.29.(61)$	2.29.31.(43)
$\gamma=2$			$2^2.13.19.29.31.37$			
21	$5^{10}.7.23.37$		$2.5.7.11.13^3.23.29.31$	$2^2.3^2.17.19.29.31^3$	$3^2.13.17.19.(337)$	$2^2.3.11^4.13^3.19$
$\gamma=1$			$-2^2.3^2.37.41.(167)$			
21	$5^9.7.23.37$		$-2.3^2.5.7.11.17.19.23$	$3^2.13.(5813)$	$2^4.29.31.(167)$	$-3.17.29.31.(181)$
$\gamma=2$			13.17.19.29.31.37.41			
22	5^{10}			2.7.13.19.29.31	$-3^4.7.17.19$	2.19.37.(53)
$\gamma=1$			$2^7.13.37.41$			
22	5^{10}			$11^2.29.37$	$-2^5.13.17.31.37$	7.13.31
$\gamma=2$			$-3^2.7.19.31.41$			
23	$2^5.5^8.13.67.71$		$2^8.3.5.13.23.37.41$	11.19.37.41.(641)	$-3.7^3.17.19.29.31.37.41$	$2^3.3^2.11.13.19^2.29.31.41$
$\gamma=1$			$-2.3^3.11.29.31.43.(47)$			
23	$2^5.3.5^{10}.11.13.67.73$		$2^6.5.7.11.13.19.23.29.31$	$-3^5.7.29.31.(13\ 109)$	$-11.17.(107.691)$	$2^3.3.7.11^2.13.19^3.37.(71)$
$\gamma=2$			2.7.19.37.41.43.(1759)			
23	$2^4.3.5^9.11.13.71.73$		$-2^4.5.11.13.17.29.31.43.(53)$	$3^3.17.19^3.23.29.31.43$	$-7.11.19.23.43.(1741)$	$-2^3.3.7^2.11^2.13.17.23.37.43$
$\gamma=3$			$2.17.23.37^3.41.(157)$			
24	$2^3.5^9.13$			$-11.13^2.31.37.41$	$3^5.19.37.41$	$2^3.3.11.41.(97)$
$\gamma=1$			$-2.3.43.(163)$			
24	2^35^913			$3.13^2.17^2.19.31$	11.(2389)	$2^3.3^2.19.23^2.37$
$\gamma=2$			$-2.11.19.37.41.43$			
25	$3^4.5^{12}.7^2.37.43.67$		$-2^4.3^2.23.31.37.43^4.(67)$	$2^3.5.11^3.13.29.31.37.(167)$	$3^3.5.7.11^2.13.19.29.37.(881)$	$2^7.3^2.5.11.13^3.29.(1249)$
$\gamma=1$			2.5.13.29.41.43.(41 189)	$-2^8.7^2.13.19^2.23.29.41.43.47$		
25	$3^4.5^{11}.7^2.37.43$		$-2.3^2.11.13.29.37.43^2.(61)$	5.23.37.(95 273)	$-2.3^3.5.7.11.19.23.31.37.(59)$	$3^2.5.23.31.(149.313)$
$\gamma=2$			$-5.11.23.31.41.43.(317)$	$2.7^2.11.31.41.43.47.(223)$		
25	$3.5^8.37.43.67$		0	$-2^3.7.11.19.31.37.41.47$	$-2^6.3.37.41^3.47$	$2.3^4.7.11.19.41.47$
$\gamma=3$			$2^9.7.19.23^2.43.47$	$5^3.7.13^2.19.23.43$		
26	$3.5^8.19.29$			$2.7.11.23.31^2.41$	$-2^2.7^3.19.23.41$	$-2^3.3.23.37.(67)$
$\gamma=1$			2.37.43.(157)	$5.7^2.17.37.43.47$		
26	$3.5^{10}.19.29$			11.37.43.(233)	$2^3.19.37.43.(107)$	$3.7^7.41.43$
$\gamma=2$			7.23.41.(1091)	$-2^3.5.7.17.23.41.47$		
27	$2.5^{13}.7.47.61.491$		$2^7.3^3.7.13.17.23.29.31.(311)$	5.17.(1 164 409)	2.5.19.37.(383.1553)	$-2.3.5.17.37.41.(8209)$
$\gamma=1$			$2^6.5.11^4.17.23.37.41.43.47$	$-3^8.7.11^2.13.23.37.41.43.47$		
27	$5^{12}.7.47.59.491$		$-7.23.29.31.37.41.(883)$	$-2^{11}.3^3.5.13^3.37^3.41$	$-2^6.3.5.13.17.19.41.(769)$	$-2^4.3^4.5.13.(53783)$
$\gamma=2$			$2.3.5.13.23.37^2.43^3.47$	$-2^5.3.7.17.23.43.47.(977)$		
27	$5^{12}.7.47.59.61$		$-2^8.3.7.13.29.37^3.43$	$-2.3^2.5.23.31.37.43.(149)$	$2^2.5.17.19.23.31.43.(853)$	$-3.5.23.29^2.31.41.43.(83)$
$\gamma=3$			$-2^3.5.31.41.47.(521)$	$2^7.7^5.13.17.31.41.47$		

TABLE XI. Coefficients of components $|\hat{2}\rangle$ of degenerate Γ_3 representations up to $l = 27$. The related $|\hat{0}\ -\rangle$ components are given in Table X with the same γ. Some values are obtained by an arbitrary orthogonalization which leads to simpler results than any of those obtained by creating zeros.

l	Denominator	$[\ 2\rangle$ $[12\rangle$ $[22\rangle$	$[\ -3\rangle$ $[-13\rangle$ $[-23\rangle$	$[\ 7\rangle$ $[17\rangle$ $[27\rangle$	$[\ -8\rangle$ $[-18\rangle$
13	$2^2.5^6.47.127$	$-2^4.5.11.(137)$	$5.(47.127)$	$2.3^6.5.7.17.19$	$-2^2.5.17.19.(101)$
$\gamma=1$		$2.11.17.19.23^3$	$3^2.11.13.17.19.23$		
13	$2.5^5.47.127$	$5.7.13.17.19.23$	0	$-2.5.11.13.23^3$	$-5.7^3.11.13.23$
$\gamma=2$		$2.7.13.(113)$	$-2^2.7^3.(79)$		
15	$2^2.5^6.11$	$2.3.5.13$	$2^6.3.5.11^2$	$-3.5.7.13^2.19$	$2.3.5.7.19.23$
$\gamma=1$		$-2^2.13.19.23$	$3.7.13.19.23$		
15	$2^2.5^5.11$	$2.5.7.19.23$	0	$-5.13.23$	$2.5.13.17^2$
$\gamma=2$		$2^2.3.7.23^2$	$-3^4.17^2$		
17	$2^6.3.5^8$	$-2^2.5.17.23.29$	$3.5.7^2.17.23.29$	$2.3^4.5.13.17.29$	$2^4.13.17.19^2.29$
$\gamma=1$		$-2.17.(199)$	$3.17.(281)$	$2^{11}.3^3.11.31$	
17	$2^6.3.5^4$	$2^2.5.13^3$	$3.5^3.13$	$2.3^4.5.23$	$-2^4.5^2.23$
$\gamma=2$		$2.13.23.29$	$3.13.23.29$	0	
18	$2.5^8.107$	$-3^2.5.11.23.29.31$	$-3.5.7.11.23.29.31$	$2^3.3^4.13.29.31$	$-3^4.11.29.31$
$\gamma=1$		$2^{11}.11.31$	$3^3.7^2.11.(37)$	$2^3.3.7.17.(79)$	$2.3.7.17.(61)$
18	$2^2.5^8.107$	$2^2.5.7.23^3$	$-3.5.23.(71)$	$-2.3^4.7.11.13.19^2$	$2^6.7.(151)$
$\gamma=2$		$2.3^2.7.29.(103)$	$3.7.29^3.31$	$2^3.3.11.17.29.31$	$-2^3.3^3.11.17.29.31$
19	$2.5^8.47.71$	$3^4.5.11.13^2.17.31$	$-2^3.3^2.5.13^4.31$	$-2^2.3^4.7.13^2.23.31$	$7.23^3.31$
$\gamma=1$		$-2.3.11^3.13^2.23.29$	$3.7.11.19^2.23.29$	$13^2.17.23.29$	$2^5.7^2.17.23.29.37$
19	$2^2.5^9.47.71$	$2^8.5.7.23.29$	$2.5.7.11.17.23.29^3$	$-2^2.11.17.29.(103)$	$-2^2.3^2.11.17.29.(163)$
$\gamma=2$		$-2^3.3.7.17.31.37^2$	$3.17^3.31.(107)$	$3^2.7.11.31.(601)$	$2.3^2.7.11^3.31.37$
20	$2^2.5^9.7.11^2$	$-2^4.3.5.17.23.29.31.37$	$-2.3.5.17.29.31.37$	$-2^2.7^3.13.29.31.37$	$2^2.3^2.13^2.29.31.37$
$\gamma=1$		$-2^3.11.37.(197)$	$-3.37.(1231)$	$17.(79.97)$	$2.3.7^2.11^4.17.19$
20	$2.5^6.7.11^2$	$3.5.23^3$	$-2^3.3.5.(103)$	$-2^2.5^2.7^3.13.17$	$-3^2.17.(233)$
$\gamma=2$		$2^5.11.17.29.31$	$-3^5.17.29.31$	$-29.31.37$	0
21	$2.5^9.7.23.37$	$3.5.13.29.31.(79)$	$-2^{11}.5.13.19.29.31$	$2.3^6.7.17.19.29.31$	$3^4.17.19^3.31$
$\gamma=1$		$2^3.3.13.17.19.(179)$	$-2^4.3^5.13^3.19$	$2^5.3^2.13.37.(197)$	$2.3.7^2.23^2.37^3$
21	$5^6.7.23.37$	$2^5.3.5.11^2.17.19$	$3^2.5.17.(101)$	$-2^{12}.5^2.7.13$	$2^3.3^2.13.17^2.29$
$\gamma=2$		$-3.29.31.(67)$	$-2.3.17^3.29.31$	$17.19.29.31.37$	0
22	$2^2.5^{12}$	$2.3^3.5.7^3.29.31$	$-2^{13}.3.7.29.31$	$-2^5.13.17.19.31$	$2^4.13^3.17.19.31$
$\gamma=1$		$2.3^2.7.13^2.19$ $-3^2.11.13.37.41.43$	$2^2.19.(827)$	$-2^5.7.13.37.41^2$	$-7.13.37.(139)$
22	2.5^{12}	$-2.3.5.13.19.29.37$	$2.3^5.13.19.29.37$	$-2^3.7.17.23^2.37$	$2^2.7^5.17.37$
$\gamma=2$		$2^3.13.17^2.31.37$ $-7.11.19.31.41.43$	$2^{10}.7.13.31.37$	$2.3^2.19.31.(53)$	$7^2.19.31.37^2$

TABLE XI. (*Continued*).

l	Denominator	$[2\rangle$ $[12\rangle$ $[22\rangle$	$[-3\rangle$ $[-13\rangle$ $[-23\rangle$	$[7\rangle$ $[17\rangle$ $[27\rangle$	$[-8\rangle$ $[-18\rangle$
23	$2^7.5^{11}.13.67.71$	$-2^7.3^2.11.13.37.41.(313)$	$3.7.11^3.37.41.(149)$	$3^5.11.17.19.29.37^3.41$	$2^4.3.11^3.17.19.29.31.37.41$
$\gamma=1$		$-2.3^2.13^5.19.29.31.37.41$	$2.11.13.19.29^3.31.37.41$	$2.3.7.11.29.31.41.(401)$	$3^8.7.11.29.31.(137)$
		$7^2.29.31.43.(109)$	$2^5.13^2.23.29.31.43.(137)$		
23	$2^7.3.5^{11}.11.13.67.73$	$2^7.3.7.13.19.23^2.29.31.43^2$	$19.29.31.(103.3617)$	$7.17.31.(779\,981)$	$2^4.3^2.7.17.(419.601)$
$\gamma=2$		$2.3^7.7.11.13.(10\,321)$	$-2.3.7^3.13.41^2.(59.83)$	$2.19.37.(142\,903)$	$-3.19.37.41.(73.2411)$
		$3^5.7.11.19.37.41.43.(331)$	$2^5.3.7.11.13^2.19.23.37.41.43$		
23	$2^6.3.5^8.11.13.71.73$	$2^7.3.13^3.17.23.29.31.43$	$-7.17.23.29.31.43.(59)$	$19.23.31.43.(3371)$	$-2^4.3^2.19^3.23.43.(79)$
$\gamma=3$		$2.3^3.7^2.11.13.17.19.23.43$	$2.3.13.17.19.23.43.(743)$	$2.7.17.23.37.43.(1013)$	$-3.7.17.23.37.41.43.(73)$
		$-3^3.11.17.23.37.41.(137)$	$2^5.3.11.13^4.17.37.41$		
24	$2^5.5^{10}.13$	$2^5.3^2.29.31.37.41$	$3.11.29.31.37.41$	$-3.5^2.7^2.11.19.37.41$	$2^3.3.11.17.19.37.41$
$\gamma=1$		$-2.3.11^2.13.19.37.41$	$-2.13.19.41.(73)$	$2.3^3.11.17^2.(71)$	$-3^2.5^2.11.(173)$
		$-23.43.(73)$	$2^7.13^2.23.43.47$		
24	$2^5.5^{10}.13$	$2^5.3.5^2.11.19.29.31$	$3^2.19.29.31.(79)$	$7^2.(157)$	$-2^3.3^4.17.(89)$
$\gamma=2$		$-2.3^2.11.13.(149)$	$-2.3^3.11.13.17^2.37$	$-2.19.37.41.(67)$	$3.7^2.17^2.19.37.41$
		$-3.11.19.23.37.41.43$	0		
25	$2^2.3^4.5^{10}.7^2.37.43.67$	$-2.13.23.31.37.(36\,527)$	$-2^3.7.13.31.37.(73.557)$	$2^4.11.13.19.29.37.(4007)$	$2^3.3^3.13.19.29.37.(3671)$
$\gamma=1$		$2.11^2.13.19.29.41^2.(197)$	$-2^2.29.(199.27\,749)$	$2^3.3.7.11.13.29.41^3.(173)$	$3.7.11.13.29.41.43.(3701)$
		$-13.19^2.23.29.41.43.47.(127)$	$2^4.11^2.13.23.29.31^2.41.43.47$		
25	$3^4.5^{11}.7^2.37.43$	$-2^6.11.29.37.(761)$	$7.11.23.29.37.(1949)$	$-2^3.19.23^3.31.37.(107)$	$2^2.3^3.11.19.23.31.37.(307)$
$\gamma=2$		$11.19.23.31.(109.271)$	$2.11.13.17^4.23.31.(59)$	$-3.7.23.31.41.(1543)$	$2.3.7.23.31.41.43.(601)$
		$-2^5.11.31.41^3.43.47$	$2.11.31.41.43.47.(397)$		
25	$2^2.3.5^{11}.37.43.67$	$2^5.7.19.23.29.31.37.41.47$	$-2^3.19.29.31.37.41^3.47$	$-7.11.13^2.37.41.47^3$	$2.3.7.37.41.47.(1693)$
$\gamma=3$		$-2^3.7.17^2.41.47.(251)$	$-2^2.7^5.13.19.41.47$	$-2.3.11^3.17^2.19.47.(83)$	$2^4.3.11.19.43.47.(181)$
		$-2^4.7.11^2.19.23.43.(103)$	$-7.19.23.43.(83.103)$		
26	$2^2.3.5^{11}.19.29$	$2^4.7.17^2.29.31.41$	$-2^5.3^3.7.31.41.(61)$	$-11^2.19^2.23.41.47^2$	$-2.3^2.17.19.23.41.(103)$
$\gamma=1$		$-2^5.3.7.11^2.23.37.41$	$2^2.3^2.13.23.29^2.37.41$	$2.7.11.23.37.43.(109)$	$2^3.3^4.7.19^2.23.37.43$
		$-2^2.23^2.37.43.47$	$3^4.7^2.37.43.47$		
26	$2.3.5^{11}.19.29$	$-2^2.23^3.29.31.37.43$	$2.3^5.23.31.37.43$	$-7.19^2.37.43.(59)$	$2.3^6.7.17.19.37.43$
$\gamma=2$		$-2.3.43.(4327)$	$-2^4.3^4.7.13.29^2.43$	$-2^{15}.11.13^2.41$	$2^5.3^4.7^2.19^2.41$
		$2^2.7.11^2.13^2.23.41.47$	$-3^2.7.23.41.47.(109)$		
27	$2.5^{13}.7.47.61.491$	$-2^3.11^2.13^4.17.23.29^2.31$	$-2^2.3.5.11^2.17.23.31.(5581)$	$5.7^3.47^2.(61.491)$	$2^8.5.11^2.(79\,631)$
$\gamma=1$		$2^8.3.5.13.17.37.(8699)$	$2^7.5.11^2.13.17.37.(1993)$	$2.5.17.37.41.43.(103.197)$	$-2^6.3^2.5.17.37.41.43.(83)$
		$2.3^2.5.7.17.23.37^3.41.43.47$	$-2^2.7.17.23.37.41.43.47.(401)$	$3^7.7.13.23.37.41.43.47.53$	
27	$5^{12}.7.47.59.491$	$-3.13.23.31.37.41.(3583)$	$-2.3^2.5.13^3.23.31.37.41$	0	$-2^3.3.5.13.17.37.41.(571)$
$\gamma=2$		$2.3^2.5.11^2.23^2.41.(59)$	$2^2.3.5.41.(257.1831)$	$2^4.3^7.5.11^2.13.23^2.43$	$-2.3.5^3.13.43.(25\,583)$
		$-2^2.3.5.7.13.23.43.47.(1637)$	$2^5.3.7.13.23.43.47.(743)$	$2^3.7^3.17.19^2.23.43.47.53$	
27	$2^2.5^{12}.7.47.59.61$	$2^2.37.43.(83.859)$	$2.3.5.37.43.(39\,659)$	0	$2^5.5.11^2.17.23^3.31.37.43$
$\gamma=3$		$-2^3.3^3.5.11^2.13.23.31.43.(59)$	$5.13.23.31.43.(10\,177)$	$-3^2.5.23.29^2.31^3.41$	$2.3^2.5^3.23.31.41.(1783)$
		$2^2.5.7.19^2.31.37^2.41.47$	$2.7.31.41.47.(4457)$	$2.3.7.13.17.31.41.47.53.(89)$	

of two components of nondegenerate Γ_3 representations.

With Tables V–VIII all representations can be obtained for integer values or half-integer values of j up to 27 except for the Γ_7 representations from $l = 18.5$, the Γ_4 representations from $l = 18$, and the Γ_3 representations from $l = 13$. Tables IX–XI give some degenerate Γ_7 and Γ_3 representations which enable any representation for $j \leqslant 27$ to be obtained.

Detailed indications concerning the use of such tables have been given in the previous article. With the Γ_1 representation for $l = 6$, one can obtain the Γ_6 representations for $j = \frac{11}{2}$ and $\frac{13}{2}$, the Γ_2 representations for $j = 5$, 6, and 7, the Γ_8 representations for $j = \frac{9}{2}, \frac{11}{2}, \frac{13}{2}$, and $\frac{15}{2}$, the Γ_5 representations for $j = 4$, 5, 6, 7, and 8, and the Γ_9 representations for $j = \frac{7}{2}, \frac{9}{2}, \frac{11}{2}, \frac{13}{2}, \frac{15}{2}$, and $\frac{17}{2}$, using $3jm$ coefficients of SU(2) an angular momentum of which is smaller than or equal to $\frac{5}{2}$. Tables of analytic formulas for Clebsch–Gordan coefficients with an angular momentum up to 5 are in Ref. 11. A very concise table of $3jm$ coefficients with an angular momentum up to 2 can be found in Appendix C of Ref. 12; it is not difficult to extend it to higher values.

In his book, Butler[8] chooses to define the second Γ_9 representation for $j = \frac{15}{2}$ from the product of the Γ_5 representation for $j = 7$ by $|s = \frac{1}{2}\rangle$: in the product there are the Γ_8 and the Γ_9 representations, which can be obtained from the Γ_1 representation for $l = 6$, which is the parent of this Γ_5 representation. Similarly, he chooses the second Γ_5 representation for $j = 8$ in the product of the Γ_8 representation for $j = \frac{15}{2}$ and $|s = \frac{1}{2}\rangle$: the parent is also the Γ_1 representation with $l = 6$. The parent of the first Γ_9 representation for $j = \frac{15}{2}$ and the first Γ_5 representation for $j = 8$ is the Γ_1 representation for $l = 10$. However, the Γ_9 representations for $j = \frac{15}{2}$ of the tables are, in fact, Γ_9'' representations, the parents of which are the Γ_3 representations for $l = 7$ and $l = 8$, respectively.

[1] R. Conte, J. Raynal, and E. Soulié, J. Math. Phys. **25**, 1176 (1984).
[2] J. Raynal, J. Math. Phys. **25**, 1187 (1984).
[3] J. Raynal and R. Conte, J. Math. Phys. **25**, 3375 (1984).
[4] R. Conte and J. Raynal, J. Math. Phys. **26**, 2413 (1985).
[5] A. G. Mc Lellan, J. Chem. Phys. **34**, 1350 (1961).
[6] T. Molien, Sitzungber. K. Preuss. Akad. Wiss. **2**, 1152 (1897).
[7] J. Griffiths, *The Theory of Transition Metal Ions* (Cambridge U. P., Cambridge, 1971), p. 386.
[8] P. H. Butler, *Point Group Symmetry Applications: Methods and Tables* (Plenum, New York, 1981), p. 80.
[9] M. Kibler, J. Mol. Spectrosc. **26**, 111 (1968); Int. J. Quantum Chem. **3**, 795 (1969).
[10] N. V. Cohan, Proc. Cambridge Philos. Soc. **54**, 28 (1958).
[11] D. A. Varshalovich, A. N. Moscaliev, and V. K. Khersonskii, *Quantum Theory of Angular Momenta* (Nauka, Leningrad, 1975), p. 228 (in Russian).
[12] A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1959), p. 911.

Algebra and physics of the unitary multiplicity-free representations of $\overline{SL}(4,\mathbb{R})$

Dj. Šijački
Institute of Physics, P.O. Box 57, Belgrade, Yugoslavia
Y. Ne'eman[a),b)]
Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel

(Received 13 March 1985; accepted for publication 2 May 1985)

The systematics of the multiplicity-free unitary irreducible representations of SL(4,ℝ) are restudied, and an amended list is presented. An automorphism essential to the physical application for particles and fields in Minkowski space is described.

I. PHYSICAL APPLICATIONS

Present knowledge about the unitary irreducible representations (unirreps) of the SL(n,ℝ) groups is incomplete. In particular, very little is known about the unirreps of the double-covering groups $\overline{SL}(n,\mathbb{R})$. The cases $n = 2,3,4$ are important in physics.[1] SL(2,ℝ) and its multiple coverings appear in numerous problems such as classical and quantized strings, projective transformations, integration over random surfaces, etc. The case $n = 3$ has been utilized to classify the excitations of deformed nuclei and of hadronic states lying along Regge trajectories. SL(4,ℝ) plays a role in the strong coupling theory and in various dynamical spectrum generating algebras. A particularly useful application for any n relates to the representations of diffeomorphisms on n-dimensional manifolds.[2] These are realized nonlinearly over the linear representations of the SL(n,ℝ) subgroups.[3] Diffeomorphisms appear in the theory of gravity, in hydrodynamics and magnetohydrodynamics, etc. In recent years, with the realization of the existence of a double covering of the diffeomorphisms,[4] to be represented nonlinearly over $\overline{SL}(4,\mathbb{R})$ in the case of space-time, a program was launched, aimed at the construction of $\overline{SL}(4,\mathbb{R})$ or even "world" spinors as spinorial "manifields" and their wave or field equations. These would appear as a superposition of matter fields, representing, for instance, phenomenological hadrons with their system of excitations, in special relativity and in the corresponding transition to three possible versions of general relativity with an active local invariance group of the tetrad frames ("G-structures"): (a) the tetrad formulation of Einstein's theory,[5] as in the case of finite spinors ("anholonomic" application of a local Lorentz invariance), (b) the Einstein–Cartan picture,[6] in which the spinor fields in addition supply a spin source term to Cartan's (algebraic) torsion equation (nonprogagating local torsion), and (c) the metric affine picture,[7] where the anholonomic invariance is further extended to $\overline{SL}(4,\mathbb{R})$, with the manifields supplying a shear source term[8] to the (algebraic) nonmetricity equation (nonpropagating local nonmetricity).

Alternatively, dropping the anholonomic treatment altogether, the manifield can be used in two approaches as a holonomic world spinor: (d) Einstein's Riemannian general relativity, and (e) a tentative affine theory,[9] with propagating torsion, curvature, and nonmetricity, but where only torsion or curvature is not confined.

Cases (a)–(c) have recently been resolved through appropriate wave equations.[10] A tentative gravitational Lagrangian has been proposed[11] for case (e) and a holonomic equation has been constructed.[12] Case (d) has also been recently resolved.[13] All of these require knowledge of the unitary irreducible representations of $\overline{SL}(4,\mathbb{R})$, or at least of the multiplicity-free unirreps [in which any representation of the maximal compact $\overline{SO}(4)$ subgroup will appear at most once in the reduction over that subgroup]. Some such representations (the "ladder" class) were constructed in connection with dynamical groups.[14] The first comprehensive study of the entire system of representations was published by Kihlberg.[15] This has been followed by further results due to Šijački[16] and Borisov.[17] The present authors published a supposedly comprehensive catalog of all multiplicity-free unirreps.[9] Unfortunately, some of the representations listed fail the test of fulfilling the algebraic commutation relations,[18] as was pointed out by Friedman and Sorkin,[19] who published what was purported to be a corrected list. We find, however, that in as much as Ref. 9 was incorrect through overlisting, Ref. 19 erred through underlisting. Considering the importance of the issue, we have now surveyed these same systematics once again, with more insight and some hindsight. Hopefully, this article will thus supply a "final" catalog of the multiplicity-free unirreps of $\overline{SL}(4,\mathbb{R})$. As for the rigorous mathematical results, we point out that the unitary duals of GL(3,ℝ) and GL(4,ℝ) have been determined by Speh.[20]

In Sec. II we outline the relations between the $\overline{SL}(4,\mathbb{R})$, SL(4,ℝ), and SO(3,3) groups and their maximal compact subgroups, we give the relevant commutation relations, and state the procedure for the construction of all unirreps. In Sec. III we find the two quotient groups of $\overline{SO}(4)$ from which all multiplicity-free unirreps are to be obtained, and list the corresponding group generator matrix elements. In Secs. IV and V we discuss the irreducibility and unitarity properties of the $\overline{SL}(4,\mathbb{R})$ unirreps. In Sec. VI we exhibit a

[a)] Wolfson Chair Extraordinary in Theoretical Physics.
[b)] Also at the University of Texas, Austin, Texas 78712.

 0022-2488/85/102457-08$02.50

list of all $\overline{\mathrm{SL}}(4,\mathbb{R})$ multiplicity-free unirreps, and make a connection and comments to previous work. Finally, in Sec. VII, we present a deunitarizing automorphism and make use of it to lay out a basis for the $\overline{\mathrm{SL}}(4,\mathbb{R})$ field structure.

II. $\overline{\mathrm{SL}}(4,\mathbb{R})$ GROUP STRUCTURE AND REPRESENTATIONS

The $\overline{\mathrm{SL}}(4,\mathbb{R})$ group is a 15-parameter noncompact Lie group. The space of the group parameters is simply connected. The maximal compact subgroup of this group is $\overline{\mathrm{SO}}(4)$, the double covering group of the SO(4) group. There is a four-element center of $\overline{\mathrm{SL}}(4,\mathbb{R})$, which is isomorphic to $Z_2 \otimes Z_2$. The factor group of $\overline{\mathrm{SL}}(4,\mathbb{R})$ with respect to a two-element (diagonal) subgroup Z_2^d of $Z_2 \otimes Z_2$ is isomorphic to SL(4,$\mathbb{R}$), i.e.,

$$\overline{\mathrm{SL}}(4,\mathbb{R})/Z_2^d \simeq \mathrm{SL}(4,\mathbb{R}),$$

while the factor group of $\overline{\mathrm{SL}}(4,\mathbb{R})$ with respect to the whole center $Z_2 \otimes Z_2$ is isomorphic to SO(3,3), i.e.,

$$\overline{\mathrm{SL}}(4,\mathbb{R})/[Z_2 \otimes Z_2] \simeq \mathrm{SO}(3,3).$$

These relations are summarized by the following diagram of exact sequences:

$$\begin{array}{ccccccccc}
 & & & & 1 & & 1 & & \\
 & & & & \downarrow & & \downarrow & & \\
1 & \to & Z_2^d & \to & Z_2 \otimes Z_2 & \to & Z_2 & \to & 1 \\
 & & & & \downarrow & & \downarrow & & \\
1 & \to & Z_2^d & \to & \overline{\mathrm{SL}}(4,\mathbb{R}) & \to & \mathrm{SL}(4,\mathbb{R}) & \to & 1 \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & \mathrm{SO}(3,3) & & \mathrm{SO}(3,3) & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 1 & & 1 & &
\end{array}$$

The maximal compact subgroups of the groups $\overline{\mathrm{SL}}(4,\mathbb{R})$, SL(4,$\mathbb{R}$), and SO(3,3) are the groups $\overline{\mathrm{SO}}(4) \simeq \mathrm{SU}(2) \otimes \mathrm{SU}(2)$, $\mathrm{SO}(4) \simeq [\mathrm{SU}(2) \otimes \mathrm{SU}(2)]/Z_2^d$, and $\mathrm{SO}(3) \otimes \mathrm{SO}(3)$, respectively. The relations between the maximal compact subgroups are given by the above diagram, with each group substituted by its maximal compact subgroup. If j_1, j_2 are the Casimir labels of the $\mathrm{SU}(2) \otimes \mathrm{SU}(2)$ group, then in an arbitrary representation $Z_2 \otimes Z_2$ is represented by $\{1,(-)^{2j_1}\} \otimes \{1,(-)^{2j_2}\}$ and Z_2^d is accordingly represented by $\{1,(-)^{2j_1} = (-)^{2j_2}\}$.

Let Q_{ab}, $a,b = 0,1,2,3$, be the $\overline{\mathrm{SL}}(4,\mathbb{R})$ generators. The $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations read

$$[Q_{ab}, Q_{cd}] = i g_{bc} Q_{ad} - i g_{ad} Q_{cb}, \tag{2.1}$$

where for the structure constants g_{ab} one can take the invariant metric tensors: either $\delta_{ab} = (+1,+1,+1,+1)$ with respect to the $\overline{\mathrm{SO}}(4)$ subgroup or $\eta_{ab} = (+1,-1,-1,-1)$ with respect to the Lorentz subgroup $\overline{\mathrm{SO}}(1,3)$ of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ group. The metric tensor g_{ab} is $\overline{\mathrm{SL}}(4,\mathbb{R})$ covariant. The antisymmetric part (when $g_{ab} = \eta_{ab}$) of Q_{ab}, i.e., $M_{ab} = Q_{[ab]}$, form the six Lorentz generators, while the remaining nine symmetric operators, i.e., $T_{ab} = Q_{(ab)}$, generate the relativistic (four-dimensional) shear transformations. The $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations are now given by the following relations:

$$\begin{aligned}
[M_{ab}, M_{cd}] &= -i(\eta_{ac} M_{bd} - \eta_{ad} M_{bc} \\
&\quad - \eta_{bc} M_{ad} + \eta_{bd} M_{ac}), \\
[M_{ab}, T_{cd}] &= -i(\eta_{ac} T_{bd} + \eta_{ad} T_{bc} \\
&\quad - \eta_{bc} T_{ad} - \eta_{bd} T_{ac}), \\
[T_{ab}, T_{cd}] &= -i(\eta_{ac} M_{bd} + \eta_{ad} M_{bc} \\
&\quad + \eta_{bc} M_{ad} + \eta_{bd} M_{ac}),
\end{aligned} \tag{2.2}$$

We will conduct the study of the multiplicity-free unitary irreducible representations of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ group in the basis of its maximal compact subgroup $\overline{\mathrm{SO}}(4)$. In this way one has on one hand an advantage of carrying out a rather straightforward calculation, and on the other hand, of applying immediately the most general mathematical theorems on the completeness of the results, which refer to the case when the unirreps of a noncompact group are analyzed in the basis of its maximal compact subgroup. The $\overline{\mathrm{SO}}(4) \simeq \mathrm{SU}(2) \otimes \mathrm{SU}(2)$ subgroup is generated by

$$J_i^{(1)} = \tfrac{1}{4} \epsilon_{ijk} M_{jk} + \tfrac{1}{2} T_{0i}, \quad J_i^{(2)} = \tfrac{1}{4} \epsilon_{ijk} M_{jk} - \tfrac{1}{2} T_{0i}, \tag{2.3}$$

where $i,j,k = 1,2,3$. The remaining nine (noncompact) operators transform with respect to $\mathrm{SU}(2) \otimes \mathrm{SU}(2)$ as the components of the (1,1) irreducible tensor operator Z. We will write them in the spherical basis as $Z_{\alpha,\beta}$, $\alpha,\beta = 0, \pm 1$. The minimal set of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations now reads as follows:

$$\begin{aligned}
&[J_0^{(p)}, J_\pm^{(q)}] = \delta_{pq} J_\pm^{(p)}, \quad p,q = 1,2, \\
&[J_+^{(p)}, J_-^{(q)}] = 2\delta_{pq} J_0^{(p)}, \\
&[J_0^{(1)}, Z_{\alpha,\beta}] = \alpha Z_{\alpha,\beta}, \\
&[J_0^{(2)}, Z_{\alpha,\beta}] = \beta Z_{\alpha,\beta}, \\
&[J_\pm^{(1)}, Z_{\alpha,\beta}] = (2 - \alpha(\alpha \pm 1))^{1/2} Z_{\alpha \pm 1,\beta}, \\
&[J_\pm^{(2)}, Z_{\alpha,\beta}] = (2 - \beta(\beta \pm 1))^{1/2} Z_{\alpha,\beta \pm 1}, \\
&[Z_{+1,+1}, Z_{-1,-1}] = -(J_0^{(1)} + J_0^{(2)}).
\end{aligned} \tag{2.4}$$

The remaining commutation relations can be obtained by making use of the Jacobi identity.

All (multiplicity-free) unirreps of a noncompact group can be constructed explicitly by the following three-step procedure, which is based on the work of Harish-Chandra.[21]

(1) One determines the matrix elements of the group generators in the basis of all homogeneous Hilbert spaces over the maximal compact subgroup and its quotient groups.

(2) One determines all sublattices of the maximal compact subgroup labels that are invariant under the action of the noncompact operators. Each invariant lattice determines the basis of the representation invariant Hilbert space.

(3) One determines the constraints on the representation labels by imposing a condition of Hermiticity on the generators in each Hilbert space corresponding to the above invariant sublattices, starting with the most general unitary, positive definite Hilbert space scalar product.

III. MATRIX ELEMENTS OF THE $\overline{\mathrm{SL}}(4,\mathbb{R})$ GENERATORS

Let the group elements k of the maximal compact subgroup SU(2) ⊗ SU(2) be parametrized by two sets of Euler angles, i.e.,

$$k(\alpha_1,\beta_1,\gamma_1,\alpha_1.\beta_2,\gamma_2)=k_1(\alpha_1,\beta_1,\gamma_1)k_2(\alpha_2,\beta_2,\gamma_2).$$

The first possible homogeneous vector spaces are those defined by the whole SU(2) ⊗ SU(2) group. The complete set of Wigner's D function provides a basis, i.e.,

$$\{((2j_1+1)(2j_2+1))^{1/2}D^{j_1}_{n_1m_1}(\alpha_1,\beta_1,\gamma_1)D^{j_2}_{n_2m_2}(\alpha_2,\beta_2,\gamma_2)\},$$

where (j_1,j_2) are the SU(2) ⊗ SU(2) Casimir labels. The labels (n_1,n_2) determine an additional multiplicity of the (j_1,j_2) values, besides the assumed $(2j_1+1)(2j_2+1)$ multiplicity corresponding to the (m_1,m_2), $|m_1|\leqslant j_1$, $|m_2|\leqslant j_2$. Since we are interested in the multiplicity-free representations, we ought to confine ourselves in this case to the subspace of this homogeneous space with basis vectors

$$\{((2j_1+1)(2j_2+1))^{1/2}D^{j_1}_{0m_1}(\beta_1,\gamma_1)D^{j_2}_{0m_2}(\beta_2,\gamma_2)\}.$$

The second possible homogeneous space is defined over the quotient [SU(2)/U(1)] ⊗ [SU(2)/U(1)], with the corresponding basis given by the vectors

$$\{((2j_1+1)(2j_2+1))^{1/2}D^{j_1}_{0m_1}(\beta_1,\gamma_1)D^{j_2}_{0m_2}(\beta_2,\gamma_2)\}.$$

Thus, we recover the multiplicity-free Hilbert subspace of the first possibility. The third possible homogeneous vector space is defined over the quotient group [SU(2) ⊗ SU(2)]/SU(2), where the elements of the "denominator" SU(2) group are parametrized by the Euler angles $(\alpha,\beta,\gamma)=(\alpha_1,\beta_1-\beta_2,\alpha_2)$. In this case, the Casimir labels j_1 and j_2 are equal mutually, i.e., $j=j_1=j_2$. The Hilbert space basis is given by the Wigner functions, $\{(2j+1)^{1/2}\times D^{j}_{m_1m_2}(\gamma_1,\beta,\gamma_2)|\beta=\beta_1+\beta_2\}$. Hereby we have listed all possibilities (with nontrivial maximal compact subgroup labels), and therefore, in order to determine all multiplicity-free unirreps of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ group, one has to look for solutions in just two homogeneous vector spaces given by quotient groups of the maximal compact subgroup SU(2) ⊗ SU(2), with corresponding basis vectors as follows:

$$[\mathrm{SU}(2)/\mathrm{U}(1)]\otimes[\mathrm{SU}(2)/\mathrm{U}(1)]\to\left\{\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle\right\}=\{((2j_1+1)(2j_2+1))^{1/2}\times D^{j_1}_{0m_1}(\beta_1,\gamma_1)D^{j_2}_{0m_2}(\beta_2,\gamma_2)\},\tag{3.1}$$

and

$$[\mathrm{SU}(2)\otimes\mathrm{SU}(2)]/\mathrm{SU}(2)\to\left\{\left|\begin{matrix}j\\ m_1 & m_2\end{matrix}\right\rangle\right\}=\{2(j+1)^{1/2}D^{j}_{m_1m_2}(\gamma_1,\beta,\gamma_2)\}.\tag{3.2}$$

The SU(2) ⊗ SU(2) group generator matrix elements are well known, while the matrix elements of the noncompact generators were determined in Ref. 9, by making use of the decontraction formula. They agree with the ones obtained in a general analysis of $\overline{\mathrm{SL}}(4,\mathbb{R})$ unirreps.[22]

In the basis (3.1) we find

$$J^{(1)}_0\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle=m_1\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle,$$

$$J^{(2)}_0\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle=m_2\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle,$$

$$J^{(1)}_\pm\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle=(j_1(j_1+1)-m_1(m_1\pm1))^{1/2}\left|\begin{matrix}j_1 & j_2\\ m_1\pm1 & m_2\end{matrix}\right\rangle,$$

$$J^{(2)}_\pm\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle=(j_2(j_2+1)-m_2(m_2\pm1))^{1/2}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\pm1\end{matrix}\right\rangle,$$

and

$$\begin{aligned}\left\langle\begin{matrix}j_1' & j_2'\\ m_1' & m_2'\end{matrix}\right|Z_{\alpha,\beta}\left|\begin{matrix}j_1 & j_2\\ m_1 & m_2\end{matrix}\right\rangle&=(-)^{j_1'-m_1'}(-)^{j_2'-m_2'}\\&\times\begin{pmatrix}j_1' & 1 & j_1\\ -m_1' & \alpha & m_1\end{pmatrix}\begin{pmatrix}j_2' & 1 & j_2\\ -m_2' & \beta & m_2\end{pmatrix}\\&\times\langle j_1'j_2'||Z||j_1j_2\rangle,\end{aligned}\tag{3.3}$$

$$\begin{aligned}\langle j_1'j_2'||Z||j_1j_2\rangle&=-i(-)^{j_1'+j_2'}((2j_1'+1)(2j_2'+1)\\&\times(2j_1+1)(2j_2+1))^{1/2}\\&\times(e_1+ie_2-\tfrac{1}{2}[j_1'(j_1'+1)\\&-j_1(j_1+1)+j_2'(j_2'+1)-j_2(j_2+1)])\\&\times\begin{pmatrix}j_1' & 1 & j_1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}j_2' & 1 & j_2\\ 0 & 0 & 0\end{pmatrix},\end{aligned}$$

where $e=e_1+ie_2$, $e_1,e_2\in\mathbb{R}$ is the $\overline{\mathrm{SL}}(4,\mathbb{R})$ representation label. The 3-j symbol

$$\begin{pmatrix}j' & 1 & j\\ 0 & 0 & 0\end{pmatrix},$$

with half-integer entries is to be evaluated by taking the corresponding expression for integer entries and continuing it to the half-integer ones.

In the basis (3.2) we find

$$J^{(1)}_0\left|\begin{matrix}j\\ m_1 & m_2\end{matrix}\right\rangle=m_1\left|\begin{matrix}j\\ m_1 & m_2\end{matrix}\right\rangle,$$

$$J^{(2)}_0\left|\begin{matrix}j\\ m_1 & m_2\end{matrix}\right\rangle=m_2\left|\begin{matrix}j\\ m_1 & m_2\end{matrix}\right\rangle,$$

$$J^{(1)}_\pm\left|\begin{matrix}j\\ m_1 & m_2\end{matrix}\right\rangle=(j(j+1)-m_1(m_1\pm1))^{1/2}\left|\begin{matrix}j\\ m_1\pm1 & m_2\end{matrix}\right\rangle,$$

$$J^{(2)}_\pm\left|\begin{matrix}j\\ m_1 & m_2\end{matrix}\right\rangle=(j(j+1)-m_2(m_2\pm1))^{1/2}\left|\begin{matrix}j\\ m_1 & m_2\pm1\end{matrix}\right\rangle,$$

and

$$\left\langle \begin{matrix} j' \\ m_1' \quad m_2' \end{matrix} \middle| Z_{\alpha,\beta} \middle| \begin{matrix} j \\ m_1 \quad m_2 \end{matrix} \right\rangle$$
$$= (-)^{j'-m_1'}(-)^{j'-m_2'} \begin{pmatrix} j_1' & 1 & j_1 \\ -m_1' & \alpha & m_1 \end{pmatrix}$$
$$\times \begin{pmatrix} j_2' & 1 & j_2 \\ -m_2' & \beta & m_2 \end{pmatrix} \langle j' || Z || j \rangle, \tag{3.4}$$
$$\langle j' || Z || j \rangle = -i((2j'+1)(2j+1))^{1/2}$$
$$\times (e_1 + ie_2 - \tfrac{1}{2}[j'(j'+1) - j(j+1)]),$$

where $e = e_1 + ie_2$, $e_1, e_2 \in \mathbb{R}$ is the $\overline{\mathrm{SL}}(4,R)$ representation label.

IV. INVARIANT LATTICES

We will treat separately the above two cases. In the first case, (3.1) and (3.3), one has *a priori* a general lattice of all (j_1, j_2) points

$$\{(j_1, j_2) | j_1, j_2 = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots\}.$$

In this case, owing to the 3-j symbols, the reduced matrix elements of the noncompact operators vanish if $j_1' = j_1$ and/or $j_2' = j_2$. Thus, the most general lattice of the (j_1, j_2) points splits into eight sublattices. Since only the four possibilities $(j_1', j_2') = (j_1 \pm 1, j_2 \pm 1)$ occur, the (j_1, j_2) content of each sublattice is determined by $j_1 + j_2 \pmod 2$, $j_1 - j_2 \pmod 2$, and the "minimal" (j_1, j_2) value. The eight such lattices we label explicitly as follows (cf. Fig. 1):

$$L_1 = L(0,0), \quad L_2 = L(\tfrac{1}{2},\tfrac{1}{2}), \quad L_3 = L(0,1) = L(1,0),$$
$$L_4 = L(\tfrac{1}{2},\tfrac{3}{2}) = L(\tfrac{3}{2},\tfrac{1}{2}), \quad L_5 = L(\tfrac{1}{2},0), \quad L_6 = L(0,\tfrac{1}{2}), \tag{4.1}$$
$$L_7 = L(0,\tfrac{3}{2}), \quad L_8 = L(\tfrac{3}{2},0).$$

In order to determine which of these sublattices are invariant under the action of the noncompact generators, we will make use of the following explicit forms of the reduced matrix elements:

$$\langle j_1 - 1\, j_2 - 1 || Z || j_1 j_2 \rangle$$
$$= -i(-)^{2(j_1+j_2)}(e_1 + ie_2 + j_1 + j_2)(j_1 j_2)^{1/2},$$
$$\langle j_1 + 1\, j_2 - 1 || Z || j_1 j_2 \rangle$$
$$= -i(-)^{2j_2-1}(e_1 + ie_2 - j_1 + j_2 - 1)((j_1+1) j_2)^{1/2}, \tag{4.2}$$

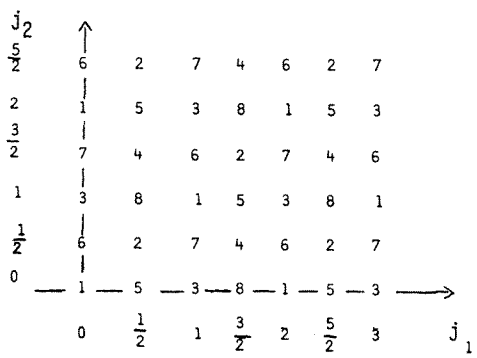


FIG. 1. L_i, $i = 1,2,\ldots,8$, sublattices [Eq. (4.1)].

$$\langle j_1 - 1\, j_2 + 1 || Z || j_1 j_2 \rangle$$
$$= -i(-)^{2j_1-1}(e_1 + ie_2 + j_1 - j_2 - 1)(j_1(j_2+1))^{1/2}.$$

It is obvious now that, owing to the $(j_1)^{1/2}$ and/or $(j_2)^{1/2}$ factors in (4.2), the lattices $L(0,0)$ and $L(0,1)$ are invariant, i.e., the reduced matrix elements $\langle j_1 - 1\, j_2 - 1 || Z || j_1 j_2 \rangle$, $\langle j_1 + 1\, j_2 - 1 || Z || j_1 j_2 \rangle$, and $\langle j_1 - 1\, j_2 + 1 || Z || j_1 + j_2 \rangle$ vanish at the edge (j_1, j_2) points $(j_1, 0)$ and $(0, j_2)$. Thus we find two invariant lattices

$$L(0,0), L(0,1) = L(1,0), \tag{4.3}$$

with the $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations satisfied for every $e_1, e_2 \in \mathbb{R}$. Whenever $e_2 \neq 0$ these two lattices are irreducible.

In the case of the lattices $L(\frac{1}{2},0)$ and $L(\frac{3}{2},0)$, the reduced matrix elements $\langle j_1 - 1\, j_2 - 1 || Z || j_1 j_2 \rangle$ and $\langle j_1 + 1\, j_2 - 1 || Z || j_1 + j_2 \rangle$ vanish for the edge points $(j_1, 0)$ but the $\langle j_1 - 1\, j_2 + 1 || Z || j_1 j_2 \rangle$ do not vanish for the points $(0, j_2)$. Thus, these lattices are not invariant, and the algebra commutation relations are not satisfied for arbitrary $e_1 e_2$ values. However, one can constrain the representation labels e_1, e_2 and find invariant sublattices. If we take $e_1 = 1 - j_{10}$, $j_{10} = \frac{1}{2}, 1, \frac{3}{2}, \ldots, e_2 = 0$, then $\langle j_1 - 1\, j_2 + 1 || Z || j_1 j_2 \rangle$ vanish for all points with $j_1 - j_2 = j_{10}$, and we thus find the triangularlike infinite irreducible sublattices

$$L(0,0; j_1 - j_2 \geqslant j_{10}), \quad j_{10} = 2,4,6,\ldots,$$
$$L(\tfrac{1}{2},0; j_1 - j_2 \geqslant j_{10}), \quad j_{10} = \tfrac{1}{2}, \tfrac{5}{2}, \tfrac{9}{2}, \ldots,$$
$$L(1,0; j_1 - j_2 \geqslant j_{10}), \quad j_{10} = 1,3,5,\ldots,$$
$$L(\tfrac{3}{2},0; j_1 - j_2 \geqslant j_{10}), \quad j_{10} = \tfrac{3}{2}, \tfrac{7}{2}, \tfrac{11}{2}, \ldots.$$

In a common notation these lattices read

$$L(j_0, 0; j_1 - j_2 \geqslant j_0),$$
$$j_0 = \tfrac{1}{2}, 1, \tfrac{3}{2}, \quad j_1 + j_2 \equiv j_0 \pmod 2. \tag{4.4}$$

To satisfy the commutation relations of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ group for these sublattices it is necessary that $e_1 = 1 - j_{10}, e_2 = 0$.

In the case of the lattices $L(0,\frac{1}{2})$ and $L(0,\frac{3}{2})$, the reduced matrix elements $\langle j_1 - 1\, j_2 - 1 || Z || j_1 j_2 \rangle$ and $\langle j_1 - 1\, j_2 + 1 || Z || j_1 j_2 \rangle$ vanish for the edge points $(0, j_2)$, but the $\langle j_1 + 1\, j_2 - 1 || Z || j_1 j_2 \rangle$ elements do not vanish for the points $(j_1, 0)$. These lattices are not invariant, and the algebra commutation relations are not satisfied. In this case one can again constrain the e_1, e_2 labels and find invariant sublattices. If we take $e_1 = 1 - j_{20}$, $j_{20} = \frac{1}{2}, 1, \frac{3}{2}, \ldots, e_2 = 0$, then $\langle j_1 + 1\, j_2 - 1 || Z || j_1 j_2 \rangle$ vanish for all points with $j_2 - j_1 = j_{20}$, and we thus find another set of triangularlike infinite irreducible sublattices

$$L(0,0; j_2 - j_1 \geqslant j_{20}), \quad j_{20} = 2,4,6,\ldots,$$
$$L(0,\tfrac{1}{2}; j_2 - j_1 \geqslant j_{20}), \quad j_{20} = \tfrac{1}{2}, \tfrac{5}{2}, \tfrac{9}{2}, \ldots,$$
$$L(0,1; j_2 - j_1 \geqslant j_{20}), \quad j_{20} = 1,3,5,\ldots,$$
$$L(0,\tfrac{3}{2}; j_2 - j_1 \geqslant j_{20}), \quad j_{20} = \tfrac{3}{2}, \tfrac{7}{2}, \tfrac{11}{2}, \ldots.$$

In a common notation these lattices read

$$L(0, j_0; j_2 - j_1 \geqslant j_0),$$
$$j_0 = \tfrac{1}{2}, 1, \tfrac{3}{2}, \quad j_1 + j_2 \equiv j_0 \pmod 2. \tag{4.5}$$

The $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations are satisfied for the la-

bels corresponding to the sublattice points provided that $e_1 = 1 - j_{20}, e_2 = 0$.

Finally, if we take $e_1 = 1, e_2 = 0$, then both $\langle j_1 - 1\, j_2 + 1 | \, |Z| \, | j_1 j_2 \rangle$ and $\langle j_1 + 1\, j_2 - 1 | \, |Z| \, | j_1 j_2 \rangle$ vanish for all points on the line $j_1 = j_2$, and $\langle j_1 - 1\, j_2 - 1 | \, |Z| \, | j_1 j_2 \rangle$ vanishes for $j_1 = j_2 = 0$. Therefore, we find an additional invariant sublattice $L(0,0; j_1 = j_2 = j), j = 0,1,2,\ldots$. The corresponding $\overline{\mathrm{SL}}(4,\mathbb{R})$ representations are a special case of the representations corresponding to the lattices of (4.7).

It is rather straightforward to check in the case of the $L(0,0)$ and $L(1,0)$ lattices that the sublattices of points $j_1 - j_2 \geqslant -1, -2, -3, \ldots$ or $j_2 - j_1 \geqslant -1, -2, -3, \ldots$ split under the action of the noncompact operators into a lattice of the form (4.3) or (4.4) plus an additional sublattice, and that for the latter the positive definiteness of the corrresponding Hilbert space scalar product [determined by (5.3) and (5.6)] is not satisfied. There are, therefore, no more irreducible invariant sublattices correponding to (3.1) and (3.3).

In the second case, (3.2) and (3.4), one *a priori* has a half-line-like lattice of points $\{(j_1, j_2) = (j, j) \mid j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\}$. The explicit form of the reduced matrix elements (3.4) is given by

$$\begin{aligned}
\langle j - 1 | \, |Z| \, | j \rangle &= -i((2j+1)(2j-1))^{1/2}(e_1 + ie_2 + j), \\
\langle j | \, |Z| \, | j \rangle &= -i(2j+1)(e_1 + ie_2), \\
\langle j + 1 | \, |Z| \, | j \rangle &= -i((2j+3)(2j+1))^{1/2}(e_1 + ie_2 - j - 1).
\end{aligned} \tag{4.6}$$

One can see immediately that $\langle j - 1 | \, |Z| \, | j \rangle$ vanishes when $j = \frac{1}{2}$, for every e_1, e_2 and that one has an invariant lattice of points $L(\frac{1}{2}, \frac{1}{2}; j_1 = j_2 = j), j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$. The same matrix element vanishes for $j = 0$, provided $e_1 = e_2 = 0$. However, owing to the existence of a nontrivial reduced matrix element $\langle j | \, |Z| \, | j \rangle$ in this case, one can explicitly verify that the $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations are satisfied for $e_1 = 0$ and an arbitrary value of e_2. Thus in the second case we find two irreducible invariant lattices for $e_2 \in \mathbb{R}$,

$$\begin{aligned}
L(0) &= \{(j, j) \mid j = 0, 2, 4, \ldots\}, \\
L(\tfrac{1}{2}) &= \{(j, j) \mid j = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots\}.
\end{aligned} \tag{4.7}$$

V. UNITARITY

The next question we want to discuss is that of the unitarity of the multiplicity-free representations of $\overline{\mathrm{SL}}(4,\mathbb{R})$, or in other words, the Hermiticity of the corresponding generators. Since $\overline{\mathrm{SL}}(4,\mathbb{R})$ is a noncompact group, its unitary representations are necessarily infinite-dimensional. Unitarity is a matter which depends on the Hilbert space one is working in, i.e., it depends on the corresponding scalar product. In order to obtain all multiplicity-free unirreps of $\overline{\mathrm{SL}}(4,\mathbb{R})$, we start with the most general scalar product of any two functions f and g,

$$(f, g) = \int dk'\, dk\, f^*(k')\kappa(k', k)g(k), \tag{5.1}$$

where $\kappa(k', k)$ is a scalar product kernel, $k, k' \in \mathrm{SU}(2) \otimes \mathrm{SU}(2)$, and dk is an invariant $\mathrm{SU}(2) \otimes \mathrm{SU}(2)$ measure. We have shown in Ref. 9 that for the most general multiplicity-free $\overline{\mathrm{SL}}(4,\mathbb{R})$ representation, the noncompact operator matrix elements take on the following form:

$$\begin{aligned}
&\begin{pmatrix} j_1' & j_2' \\ m_1' & m_2' \end{pmatrix} Z_{\alpha,\beta} \begin{pmatrix} j_1 & j_2 \\ m_1 & m_2 \end{pmatrix} \\
&\quad = \kappa(j_1, j_2) \begin{pmatrix} j_1' & j_2' \\ m_1' & m_2' \end{pmatrix} Z_{\alpha,\beta} \begin{pmatrix} j_1 & j_2 \\ m_1 & m_2 \end{pmatrix},
\end{aligned} \tag{5.2}$$

where $\kappa(j_1, j_2)$ are the matrix elements of the kernel. The positive definiteness of the scalar product, i.e., $(f, f) > 0$ for every f yields

$$\kappa(j_1, j_2) \geqslant 0, \tag{5.3}$$

and the Hermiticity of the scalar product, i.e., $(f, g) = (g, f)^*$ implies

$$\kappa(j_1, j_2) = \kappa^*(j_1, j_2). \tag{5.4}$$

The Hermiticity of the noncompact operators reads in the spherical basis as follows:

$$Z^{+}_{\alpha,\beta} = (-)^{\alpha-\beta} Z_{-\alpha,-\beta}. \tag{5.5}$$

Making use of this condition, and of (3.3) and (5.4), we arrive at

$$\begin{aligned}
&(e_1 + ie_2 - \tfrac{1}{2}[\, j_1'(j_1' + 1) - j_1(j_1 + 1) \\
&\quad + j_2'(j_2' + 1) - j_2(j_2 + 1)\,])\kappa(j_1', j_2') \\
&\quad = (-e_1 + ie_2 - \tfrac{1}{2}[\, j_1'(j_1' + 1) - j_1(j_1 + 1) \\
&\qquad + j_2'(j_2' + 1) - j_2(j_2 + 1)\,])\kappa(j_1, j_2).
\end{aligned} \tag{5.6}$$

This equation provides us with two cases, i.e., $e_1 = 0, e_2 \in \mathbb{R}$ and $e_1 \neq 0, e_2 = 0$.

It is at this stage that we inspect the unitarity of the representations, as well as the positive definiteness of the scalar product, for each of the irreducible sublattices found in the above, (4.3)–(4.5) and (4.7).

The $\overline{\mathrm{SL}}(4,\mathbb{R})$ representations corresponding to the invariant lattices $L(0,0)$ and $L(1,0)$ of (4.3), are already unitary with a trivial kernel for the scalar product, $\kappa(j_1, j_2) = 1$ for every j_1, j_2, provided $e_1 = 0$, $e_2 \in \mathbb{R}$. These unirreps of $\overline{\mathrm{SL}}(4,\mathbb{R})$ form the principal series, which we denote by $D^{\mathrm{pr}}(0,0; e_2)$ and $D^{\mathrm{pr}}(1.0; e_2)$. In the case $e_1 \neq 0$, $e_2 = 0$, from (5.3) and (5.6), we find that there is a solution for e_1 if

$$|e_1| < 1 - |j_1 - j_2| \quad \text{and} \quad |e_1| < 2 + j_1 + j_2, \tag{5.7}$$

for every (j_1, j_2) point of a given lattice. We find by inspection that $0 < |e_1| < 1$ for the lattice $L(0,0)$, and that there are no solutions for the lattice $L(1,0)$. These unirreps form the supplementary series, and we denote them by $D^{\mathrm{supp}}(0,0; e_1)$. The matrix elements of the kernel are now

$$\kappa(j_1, j_2) = \frac{\Gamma(j_1 + j_2 + e_1 + 1)\Gamma(1 - e_1)\Gamma(|j_1 - j_2| + e_1 + 2)\Gamma(2 - e_1)}{\Gamma(j_1 + j_2 - e_1 + 1)\Gamma(1 + e_1)\Gamma(|j_1 - j_2| - e_1 + 2)\Gamma(2 + e_1)} \kappa(0,0). \tag{5.8}$$

For the $\overline{\mathrm{SL}}(4,\mathbb{R})$ representations corresponding to the irreducible lattices of (4.4) and (4.5) to be unitary, a nontrivial kernel is required for the scalar product. The unitarity condition (5.6) can then be satisfied for any $e_1 = 1 - j_{10}$, $j_{10} = \frac{1}{2},1,\frac{3}{2},\ldots$, or $e_1 = 1 - j_{20}$, $j_{20} = \frac{1}{2},1,\frac{3}{2},\ldots$, and $e_2 = 0$, and that the positive definiteness condition is also satisfied by the scalar product. The corresponding unirreps form the discrete series of multiplicity-free unirreps of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ group. We denote them by $D^{\mathrm{disc}}(1-e_1,0)$, and by $D^{\mathrm{disc}}(0,1-e_1)$, $e_1 = \frac{1}{2},0,-\frac{1}{2},-1,\ldots$, and they correspond, respectively, to the irreducible lattices of (4.4) and (4.5). For the discrete series, (5.6) yields

$$\kappa(j_1,j_2) = \frac{\Gamma(j_1+j_2+e_1+1)\Gamma(|j_1-j_2|+e_1+2)}{\Gamma(j_1+j_2-e_1+1)\Gamma(|j_1-j_2|-e_1+2)} \times \kappa(\min(j_1),\min(j_2)), \tag{5.9}$$

where $\kappa(\min(j_1),\min(j_2))$ is either $\kappa(1-e_1,0)$ or $\kappa(0,1-e_1)$.

The $\overline{\mathrm{SL}}(4,\mathbb{R})$ representations corresponding to the irreducible lattices $L(0)$ and $L(\frac{1}{2})$ of (4.7) are, as we have already stated, unitary for $e_1 = 0$ and an arbitrary $e_2 \in \mathbb{R}$. This result follows from an explicit verification of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations. These representations form the ladder unirreps, and we denote them by $D^{\mathrm{ladd}}(0,e_2)$ and $D^{\mathrm{ladd}}(\frac{1}{2};e_2)$.

The irreducibility of the Hilbert spaces in which we have defined the multiplicity-free unirreps of $\overline{\mathrm{SL}}(4,\mathbb{R})$ is guaranteed by construction—none of them possesses an invariant subspace under the action of the group generators.

The second-order Casimir operator for $\overline{\mathrm{SL}}(4,\mathbb{R})$ is defined by

$$C_2 = Q_{ab}Q^{ba} = -4 + \tfrac{1}{4}(e_1 + ie_2)^2. \tag{5.10}$$

For the principal and the ladder series it is given by

$$C_2 = -4 - \tfrac{1}{4}e_2^2, \tag{5.11}$$

while for the supplementary and the discrete series it takes, respectively, the values

$$C_2 = -4 + \tfrac{1}{4}e_1^2 \tag{5.12}$$

and

$$C_2 = -4 + \tfrac{1}{4}e_1^2 \to -4 + \tfrac{1}{4}(j_0 - 1)^2, \tag{5.13}$$

where $j_0 = \frac{1}{2},1,\frac{3}{2},2,\ldots$ is either j_{10} or j_{20}.

VI. SUMMARY OF $\overline{\mathrm{SL}}(4,\mathbb{R})$ MULTIPLICITY-FREE UNIRREPS

We have parametrized, in this work, the unitary irreducible representations of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ group in terms of the parameter $e = e_1 + ie_2$. The representations are defined in Hilbert spaces which are symmetric homogeneous spaces over certain quotient groups K' of the maximal compact subgroup $K = \mathrm{SU}(2) \otimes \mathrm{SU}(2)$, i.e., in the spaces $L^2(K')$ of square-integrable functions over K' with respect to the invariant measure over K'. We have considered the most general scalar product (5.1) of the Hilbert space elements with, in general, a nontrivial kernel κ. The K' representation eigenvector labels, which define a basis of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ representation Hilbert space, are constrained to belong to certain irreducible lattices L. Therefore, we denote the unirrep Hilbert spaces by $H(K',\kappa,L)$.

There are, besides the trivial representation, four series of multiplicity-free unirreps of the $\overline{\mathrm{SL}}(4,\mathbb{R})$ group.

Principal series: $D^{\mathrm{pr}}(0,0;e_2)$ and $D^{\mathrm{pr}}(1,0;e_2)$, $e_1 = 0$, $e_2 \in \mathbb{R}$. They are defined in the Hilbert spaces $H(K'_1,\kappa,L)$, where $K'_1 = [\mathrm{SU}(2)/\mathrm{U}(1)] \otimes [\mathrm{SU}(2)/\mathrm{U}(1)]$, $\kappa(j_1,j_2) = 1$, $\forall j_1,j_2$, and the irreducible lattices are, respectively, $L(0,0)$ and $L(1,0)$ [cf. (4.3)]. The generator matrix elements are given by (3.3), and the Casimir invariant is given by (5.11).

Supplementary series: $D^{\mathrm{supp}}(0,0;e_1)$, $0 < |e_1| < 1$, $e_2 = 0$. They are defined in the Hilbert spaces $H(K'_1,\kappa,L)$, where $K'_1 = [\mathrm{SU}(2)/\mathrm{U}(1)] \otimes [\mathrm{SU}(2)/\mathrm{U}(1)]$, $\kappa(j_1,j_2)$ is nontrivial and given by (5.8), and the irreducible lattice is $L(0,0)$ [cf. (4.3)]. The generator matrix elements are given by (3.3) and (5.2), and the Casimir invariant is given by (5.12).

Discrete series: $D^{\mathrm{disc}}(1-e_1,0)$ and $D^{\mathrm{disc}}(0,1-e_1)$, $e_1 = 1 - j_0, j_0 = \frac{1}{2},1,\frac{3}{2}, e_2 = 0$. They are defined in the Hilbert spaces $H(K'_1,\kappa,L)$, where $K'_1 = [\mathrm{SU}(2)/\mathrm{U}(1)] \otimes [\mathrm{SU}(2)/\mathrm{U}(1)]$, $\kappa(j_1,j_2)$ is nontrivial and given by (5.9), and the irreducible lattices are, respectively, $L(j_0,0;j_1-j_2 \geqslant j_0)$ and $L(0,j_0;j_2-j_1 \geqslant j_0)$ [cf. (4.4) and (4.5)]. The generator matrix elements are given by (3.3) and (5.2), and the Casimir invariant is given by (5.13).

Ladder series: $D^{\mathrm{ladd}}(0;e_2)$ and $D^{\mathrm{ladd}}(\frac{1}{2};e_2)$, $e_1 = 0$, $e_2 \in \mathbb{R}$. They are defined in the Hilbert spaces $H(K'_2,\kappa,L)$, where $K'_2 = [\mathrm{SU}(2) \otimes \mathrm{SU}(2)]/\mathrm{SU}(2)$, $\kappa(j_1,j_2) = \kappa(j,j) = 1$, $\forall j$, and the irreducible lattices are, respectively, $L(0)$ and $L(\frac{1}{2})$ [cf. (4.7)]. The generator matrix elements are given by (3.4), and the Casimir invariant is given by (5.11).

Let us comment briefly on the previous work on $\overline{\mathrm{SL}}(4,\mathbb{R})$ multiplicity-free unirreps. The ladder series, with $e_2 = 0$, were obtained by Dothan and Ne'eman,[14] and in this work e_2 was constrained by the algebraic structure of the physical model they considered. The general ladder series, $e_2 \neq 0$, were obtained by Mukunda,[23] by means of an analytic continuation in the SU(4) labels, and by Šijački,[16] who solved the commutation relations explicitly. Kihlberg[15] failed to check the $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations in the multiplicity-free case, when continuing the $\overline{\mathrm{SO}}(4)$ labels to the half-integer values; in addition he did not consider the homogeneous space over the K'_2 group and thus obtained only the $e_2 = 0$ ladder unirreps. Friedman and Sorkin[19] realized the importance of checking the commutation relations for the continued values (or equivalently checking whether the relevant sublattices are invariant), and made an attempt to find all multiplicity-free unirreps. However, they did not actually solve their relevant equations (A.6)–(A.17): instead, they made use of our result[9,18] according to which only $\langle j_1 \pm 1\, j_2 \pm 1 | \, |Z| \, j_1 j_2\rangle$ and $\langle j_1 \pm 1\, j_2 \mp 1 | \, |Z| \, | j_1 j_2\rangle$ matrix elements are nonzero; unfortunately, this result only holds in the Hilbert space over the K'_1 group. The $\langle jj | \, |Z| \, | jj\rangle$ matrix elements are indeed nonzero (and proportional to e_2) for the Hilbert space over the K'_2 group. When embedding a Hilbert space over K'_2 into a Hilbert space over K'_1, in order to prevent nonvanishing $\langle j \pm 1\, j \mp 1 | \, |Z| \, | jj\rangle$ matrix elements, $e_2 = 0$ is required, and thus only the D^{ladd}, $e_2 = 0$ unirreps survive. The supplementary unirreps were altogether missed in their work as well.

The $\overline{\mathrm{SL}}(4,\mathbb{R})$ multiplicity-free representation labels used in previous work are given in terms of the labels of this work by $R = e_2$ (Ref. 23), $\eta = e_2$ (Ref. 16), $(a_1,a_2) = (e_2, -e_1)$ (Ref. 15), $(p_1,p_2) = (-e_1, -e_2)$ (Refs. 9 and 18), and $k = -4(e_1 + ie_2)^2$ (Ref. 19).

VII. THE DEUNITARIZING AUTOMORPHISM $\mathscr{A}$

The $\overline{\mathrm{SL}}(4,\mathbb{R})$ generators M_{ab}, T_{ab}, $a,b = 0,1,2,3$ of (2.2) can be rearranged according to the following set: $J_i = \frac{1}{2}\epsilon_{ijk}M_{jk}$, $N_i = T_{0i}$, $K_i = M_{0i}$, T_{ij}, T_{00}, where $i,j = 1,2,3$. The $\overline{\mathrm{SL}}(4,\mathbb{R})$ commutation relations now read

$$\begin{aligned}
&[J_i,J_j] = i\epsilon_{ijk}J_k,\\
&[J_i,N_j] = i\epsilon_{ijk}N_k,\\
&[J_i,K_i] = i\epsilon_{ijk}K_k,\\
&[N_i,N_j] = i\epsilon_{ijk}J_k,\\
&[K_i,K_j] = -i\epsilon_{ijk}J_k,\\
&[J_i,T_{jk}] = i\epsilon_{ijl}T_{lk} + i\epsilon_{ikl}T_{jl},\\
&[J_i,T_{00}] = 0,\\
&[T_{ij},T_{kl}] = -i(\delta_{ik}\epsilon_{jlm} + \delta_{il}\epsilon_{jkm} + \delta_{jk}\epsilon_{ilm} + \delta_{jl}\epsilon_{ikm})J_m,\\
&[K_i,N_j] = -i(T_{ij} + \delta_{ij}T_{00}),\\
&[K_i,T_{jk}] = -i(\delta_{ij}N_k + \delta_{ik}N_j),\\
&[N_i,T_{jk}] = -i(\delta_{ij}K_k + \delta_{ik}K_j),\\
&[K_i,T_{00}] = -2iN_i,\\
&[N_i,T_{00}] = -2iK_i,\\
&[T_{ij},T_{00}] = 0.
\end{aligned} \tag{7.1}$$

The compact operators are J_i and N_i, while the remaining ones K_i, T_{ii} and T_{00} are noncompact. Note the following subgroups:

$$\begin{aligned}
&\overline{\mathrm{SO}}(4) &&: J_i,N_i,\\
&\overline{\mathrm{SO}}(1,3)\simeq\mathrm{SL}(2,C) &&: J_i,K_i,\\
&\overline{\mathrm{SL}}(3,\mathbb{R}) &&: J_i,T_{ij}.
\end{aligned} \tag{7.2}$$

The commutation relations (6.1) are invariant under the automorphism

$$\begin{aligned}
&J_i \to J_i,\\
&N_i \to iK_i,\\
\mathscr{A}:\quad &K_i \to iN_i,\\
&T_{ij} \to T_{ij},\\
&T_{00} \to T_{00}.
\end{aligned} \tag{7.3}$$

As a result of this automorphism, the vector spaces carrying unirreps of $\overline{\mathrm{SL}}(4,\mathbb{R})$ can also realize the action of $\overline{\mathrm{SL}}(4,\mathbb{R})_{\mathscr{A}}$, except that some of the latter group's matrices will not be unitary, having been multiplied by $\sqrt{-1}$ in (7.3). This is essential for most physical applications. Indeed, ordinary tensor fields carry finite nonunitary representations of $\mathrm{SL}(4,\mathbb{R})$, and Dirac or Bargmann–Wigner spinor fields carry finite nonunitary representations of $\mathrm{SL}(2,\mathbb{C}) \subset \overline{\mathrm{SL}}(4,\mathbb{R})$. In both cases, the physical generators of Lorentz transformations—the boosts—are entirely orbital. This can be seen in the following way.

The Noether theorem determines the following structure for the total angular momentum:

$$M_{ab} = \int_\sigma \{(x_a\Theta_b{}^\mu - x_b\Theta_a{}^\mu) + \mathscr{S}_{ab}{}^\mu\}d\sigma_\mu + \text{h.c.}, \tag{7.4}$$

where h.c. denotes the Hermitian conjugate expression, σ_μ is a spacelike hyperplane,

$$\Theta_a{}^\mu = e_a{}^\mu\mathscr{L} - \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}\partial_a\phi \tag{7.5}$$

is the canonical energy-momentum tensor, $\mathscr{L}$ is the Lagrangian density, ϕ the field, $S_{ab}{}^\mu$ is the intrinsic spin tensor density

$$\mathscr{S}_{ab}{}^\mu = \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}S_{ab}\,\phi, \tag{7.6}$$

and S_{ab} is the matrix representation of M_{ab} on the ϕ vector space. These matrices are unitary for the S_{ij}, which belong to the compact subgroup $\overline{\mathrm{SO}}(3)$; they represent the J_{ij} in Eq. (2.3),

$$J_{ij} = \epsilon_{ijk}(J^{(1)}_k + J^{(2)}_k). \tag{7.7}$$

The S_{0i}, on the other hand, represent the noncompact special Lorentz transformation generators K_i. In any finite-dimensional representation of $\overline{\mathrm{SL}}(4,\mathbb{R})$ or $\mathrm{SL}(2,\mathbb{C})$, they are given by anti-Hermitian matrices. For example, in the four-dimensional defining representation, they are given by the real symmetric matrices $E_{0i} + E_{i0}$, where

$$(E_{mn})^A_B = \delta^A_m\delta_{Bn}$$

(A,B are the row and column indices).

The first bracket in the expression (7.4) for M_{ab} is the orbital angular momentum, the second is the spin. As a result of the addition of $(S_{ab})^+$, the M_{ij} indeed contain both orbital and spin angular momentum, but the intrinsic spin piece cancels in M_{0i}. The physical boost for all known physical fields is entirely orbital and contributes to the kinetic energy only.

In unitary representations of $\overline{\mathrm{SL}}(4,\mathbb{R})$, the boost possesses a nonvanishing intrinsic piece and raises the mass or potential energy, connecting the particle to a higher excited state. To avoid this unphysical result, we identify the physical generators instead in $\overline{\mathrm{SL}}(4,\mathbb{R})_{\mathscr{A}}$. Here $M^{\mathscr{A}}_{0i}$ is given by the finite non-Hermitian matrices of (iT_{0i}). It is this deunitarizing automorphism which allowed the authors in Refs. 9 and 12 to claim that the $\overline{\mathrm{SO}}(4)$ compact subgroup matrices can be used for the Lorentz $\overline{\mathrm{SO}}(1,3)$. In Ref. 10, both possibilities were investigated, the unitary $\overline{\mathrm{SL}}(4,\mathbb{R})$ case representing hypothetical particles obeying a Majorana-like equation. Note that the $\sqrt{-1}$ in the automorphism can also be absorbed in the space-time manifold, with $x^4 = ix^0$. This is the "Pauli metric" in which the metric is indeed Euclidean and thus $\overline{\mathrm{SO}}(1,3) \to \overline{\mathrm{SO}}(4)$.

In the general affine[9] approach to particle physics and gravity, the fundamental symmetry is that of the $\overline{\mathrm{GA}}(4,\mathbb{R})$ group. In studying the unirreps of this group in the space of quantum states we have found that for hadrons, $\overline{\mathrm{SL}}(3,\mathbb{R})$ (see Refs. 18 and 24) is the relevant "little group," from which

one induces the $\overline{GA}(4,\mathbb{R})$ unirreps. This stability group has to be represented unitarily, since its states form a basis of the quantum mechanical Hilbert space. In the $\overline{GA}(4,\mathbb{R})$ representations on fields, the $\overline{SL}(4,\mathbb{R})$ homogeneous subgroup is actively realized in the space of the field components, and its representations thus define the general affine fields. The two pictures have to merge for the stability subgroup, so that the $\overline{SL}(3,\mathbb{R}) \subset \overline{SL}(4,\mathbb{R})$, when represented on fields, has to be unitary. This $\overline{SL}(3,\mathbb{R})$ is generated by the J_i and T_{ij}, and we observe that these operators are indeed unaffected by the deunitarizing automorphism (7.3).

ACKNOWLEDGMENTS

One of the authors (Dj.Š.) wishes to thank the Wolfson Chair Extraordinary in Theoretical Physics of Tel Aviv University for its hospitality and financial support. The other author (Y.N.) was supported in part by the U.S./Israel Binational Science Foundation.

[1]Y. Ne'eman, Found. Phys. **13**, 467 (1983).
[2]Y. Ne'eman, "Representation of the group of diffeomorphisms in the physics of particles, gravity and fluids," to be published in Ann. NY Acad. Sci. (Proceedings of the 1984 Tel Aviv International Conference on Collective Phenomena).
[3]A. B. Borisov and V. I. Ogievetsky, Teor. Mat. Fiz. **21**, 329 (1974).
[4]Y. Ne'eman, Ann. Inst. H. Poincaŕe, **28**, 369 (1978).
[5]See, for example, S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972), Chap. XII, Sec. 5.
[6]See, for example, F. W. Hehl, P. von der Heyde, G. D. Kerlick, and J. M. Nester, Rev. Mod. Phys. **48**, 393 (1976).
[7]F. W. Hehl, G. D. Kerlick, and P. von der Heyde, Phys. Lett. B **63**, 446 (1976).
[8]F. W. Hehl, E. A. Lord, and Y. Ne'eman, Phys. Lett. B **71**, 432 (1977) and Phys. Rev. D **17**, 418 (1978).
[9]Y. Ne'eman and Dj. Šijački, Ann. Phys. (NY) **120**, 292 (1979).
[10]A. Cant and Y. Ne'eman, "Spinorial infinite equations fitting metric-affine gravity," Tel Aviv University report TAUP N156-84 (to appear in J. Math. Phys.).
[11]Dj. Šijački, Phys. Lett. B **109**, 435 (1982).
[12]J. Mickelsson, Commun. Math. Phys. **88**, 551 (1983).
[13]Y. Ne'eman, "World spinors in Riemannian gravity," Tel Aviv University report TAUP N157-85.
[14]Y. Dothan and Y. Ne'eman, in *Resonant Particles*, edited by B. A. Munir (Ohio U., Athens, Ohio, 1965), p. 17; reprinted in *Symmetry Groups in Nuclear and Particle Physics*, edited by F. J. Dyson (Benjamin, New York, 1966), p. 287,
[15]A. Kihlberg, Ark. Fys. **32**, 241 (1966).
[16]Dj. Šijački, Ph.D thesis, Duke University, 1974; Ann. Israel Phys. Soc. **3**, 35 (1980).
[17]A. B. Borisov, Rep. Math. Phys. **13**, 141 (1978).
[18]Y. Ne'eman and Dj. Šijački, Proc.Nat. Acad. Sci. USA **76**, 561 (1979), and, more especially, **77**, 1761 (1980).
[19]J. L. Friedman and R. D. Sorkin, J. Math. Phys. **21**, 1269 (1980).
[20]B. Speh, Mat. Ann. **258**, 113 (1981).
[21]Harish-Chandra, Proc. Nat. Acad. Sci. USA **37**, 170, 362, 366, 691 (1951).
[22]Dj. Šijački, "The continuous unitary irreducible representations of $\overline{SL}(4,\mathbb{R})$," Institute of Physics, Belgrade, preprint.
[23]N. Mukunda, in *Non-Compact Groups in Particles Physics*, edited by Y. Chow (Benjamin, New York, 1966).
[24]Dj. Šijački, in *Frontiers in Particle Physics '83* (World Scientific, Singapore, 1984).

Character constraints on duality

John J. Sullivan
Physics Department, University of New Orleans, New Orleans, Louisiana 70148

(Received 27 February 1985; accepted for publication 31 May 1985)

Duality identifies recoupling coefficients and isoscalar factors of the unitary groups with matrix elements of double coset representatives of the symmetric groups. When the double coset factorization is with respect to the same subgroup chain the matrix is that of the ordinary irreducible representation in a basis symmetry adapted to the subgroup chain. The invariant character of the symmetric group element then acts as a constraint on the magnitude, phase, and multiplicity resolution for diagonal elements. An algorithm for determining the content of this constraint is given and is shown to be consistent with a phase convention proposed previously.

I. DUALITY

In Ref. 1 (hereafter referred to as I) we have extensively analyzed the duality which exists between tensor representations of the general linear groups GlnC and the ordinary irreducible representations (irreps) of the symmetric groups S_L using double coset (DC) decompositions $S_{iL} \backslash S_L / S_{L_j}$. Duality identifies 9-$j$ recoupling coefficients and isoscalar factors of Un with double coset matrix elements (DCME) and weighted double coset matrix elements (WDCME) in S_L. The DCME are elements of the transformation matrix between equivalent bases symmetry-adapted to the two subgroup chains. When the subgroups are identical ($_iL = L_i$) the DCME form the irreducible representation matrix of the group element chosen to represent the DC. In these cases the invariant character of the element in S_L provides an additional constraint on the magnitudes and phases of diagonal DCME. In I we argue duality may be applied directly to the unimodular SUn subgroups if appropriate conventions for phase and multiplicity resolutions are observed. For the phase convention proposed in I character theory also implies constraints on the resolution of multiplicities involved in the DCME. This work analyzes these constraints and gives an algorithm distinguishing diagonal and off-diagonal resolutions of multiplicities for outer products of the type $(\lambda{:}\lambda_1{}^2)$.

In the next section we introduce notation by restating the problem in terms used in I. The character as evaluated by Nakayama's formula is related to its expression in terms of DCME and an algorithm for distinguishing the multiplicities is given. In the final section two sum formulas for diagonal DCME, which derive from the identifications of duality, are noted.

II. DCME EXPRESSION FOR THE CHARACTER

The DC for the decomposition $S_{iL} \backslash S_L / S_{L_j}$ are enumerated by sets $_iL_j$ of non-negative integers, which in the array

$$\begin{bmatrix} L & L_j \\ {}_iL & {}_iL_j \end{bmatrix}$$

sum to the left column and to the top row. Because of backcoupling relations there is no loss in generality in limiting the range of indices i and j to 2 nor in setting $_2L_2 = 0$. The DC representative element is chosen as that permutation that cyclically permutes the ordered sets $(_2L_1 + {}_1L_2)$ sequentially $_1L_2$ times. The DCME are labeled by arrays

$$\begin{matrix} \begin{bmatrix} \lambda & \lambda_j \\ {}_i\lambda & {}_i\lambda_j \end{bmatrix} \begin{matrix} r \\ {}_ir \end{matrix} , \\ r' \quad r_j \end{matrix}$$

where the λ's label irreps of the respective subgroups. The irreps in any row or column must couple as outer products and the indices r label the multiplicities $|\lambda{:}\lambda_j|$ of these outer products. The DCME are orthogonal on the sets $(r'{}_i\lambda_i r | r\lambda_j r_j)$, with λ and $_i\lambda_j$ acting as fixed parameters. For $_iL = L_i$ and setting $_2L_2 = 0$ we consider DCME of the form

$$\begin{matrix} \begin{bmatrix} \lambda & \lambda_1 & \lambda_2 \\ {}_1\lambda & {}_1\lambda_1 & \lambda_2 \\ {}_2\lambda & {}_2\lambda & 0 \end{bmatrix} \begin{matrix} r \\ s \end{matrix} \\ r' \quad s' \end{matrix}$$

representing a DC element with cycle structure $(2^{L_2}, 1^{{}_1L_1})$. The trace of this element gives the character as

$$\lambda(2^{L_2}, 1^{{}_1L_1}) = \sum_{\substack{\lambda_j \overline{\lambda}_1 \\ r\ s}} |{}_1\lambda_1|\ |\lambda_2| \begin{matrix} \begin{bmatrix} \bar{\lambda} & \lambda_1 & \lambda_2 \\ \lambda_1 & {}_1\lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_2 & 0 \end{bmatrix} \begin{matrix} r \\ s \end{matrix} \\ r \quad s \end{matrix} . \tag{1}$$

The backcoupling relations give the equivalent sums

$$\lambda(2^{L_2}, 1^{{}_1L_1}) = \sum_{\epsilon \lambda_2 \lambda_1 r} |{}_1\lambda_1|\ |\lambda_2|\ |\lambda{:}\epsilon_1\lambda_1| \begin{matrix} \begin{bmatrix} \epsilon & \lambda_2 & \lambda_2 \\ \lambda_2 & 0 & \lambda_2 \\ \lambda_2 & \lambda_2 & 0 \end{bmatrix} r \\ r \end{matrix}$$

$$= \sum_{\epsilon_1 \bar{\lambda}_1} |{}_1\lambda_1|\ |\lambda{:}\epsilon_1\lambda_1|\ \epsilon(2^{L_2}) .$$

DCME of the form

$$\begin{matrix} \begin{bmatrix} \lambda & \lambda_1 & \lambda_1 \\ \lambda_1 & 0 & \lambda_1 \\ \lambda_1 & \lambda_1 & 0 \end{bmatrix} r \\ r' \end{matrix}$$

are transposition phases orthogonal and symmetric in the multiplicities $(r'|r)$. In I we have argued transposition phase matrices may be put in the form of a permutation matrix (one unit $+1$ in any row or column) times a phase $\phi(\lambda{:}\lambda_1\lambda_2)$ independent of the multiplicity and specified by a convention

 0022-2488/85/102465-04$02.50

based on the Littlewood–Richardson rules for decomposing an outer product into its irreducible constituents. That is, under transposition the multiplicities either map to themselves and thus contribute to the trace or they are off-diagonal in pairs and do not contribute to the trace. Diagonalization of the transposition matrix gives ± 1 eigenvalues for each multiplicity pairing and corresponds to using the wreath product $S_{2L}/S_2 V(S_L)^2$ in the subgroup chain. The paired multiplicities separate into symmetric and antisymmetric species. This determines, along with the phase ϕ, the resolution of the symmetrized multiplicities $|\lambda{:}\lambda_1 \otimes \sigma|$, where $\sigma = [2]$ or $[1^2]$. To show that the character constraint is not trivial we cite two examples: $[3,2,1](2^3) = 0$ requires $|[3,2,1]{:}[2,1] \otimes [2]| = 1 = |[3,2,1]{:}[2,1] \otimes [1^2]|$, while $[4^2,2^2](2^6) = +80$ requires $|[4^2,2^2]{:}[3,2,1] \otimes [2]| = 2$ and $|[4^2,2^2]{:}[3,2,1] \otimes [1^2]| = 0$.

It is therefore reasonable to ask what are the conditions which must be satisfied for an irrep with $|\lambda{:}\lambda_1{}^2| \neq 0$ to contribute to the sum for the character and as a corollary require the multiplicity to be resolved in diagonal or off-diagonal form.

III. CHARACTER VIA NAKAYAMA'S FORMULA

The evaluation of the ordinary irreducible characters of the symmetric group is a widely studied subject and several iterative procedures are well known. We proceed using Nakayama's formula as applied to an element containing a q-cycle b times. James and Kerber[2] give an excellent presentation of this subject and it is the relevant portions of their Sec. 2.7 we give here. At the close of this section we relate this procedure to the DC decompositions introduced above.

In a manner similar to Eq. (1), Nakayama's formula expands the irreducible characters of S_L in terms of subtraces of a subgroup sequence adapted to the element being considered. For an element with b q-cycles the Nakayama formula becomes

$$\lambda(q^b \pi) = \operatorname{sgn}\sigma f^{\lambda}(q)\hat{\lambda}(\pi),$$

where

$$\lambda \vdash L \text{ and } \hat{\lambda} \vdash (L - qb). \tag{2}$$

The element π is a permutation acting on the set $(L - qb)$ fixed by q^b. The phase $\operatorname{sgn}\sigma$ and the dimension $f^{\lambda}(q)$ are discussed below. An irrep λ has a unique q-core, q-quotient, and q-weight w such that if $b > w$ the character vanishes and if $b = w$, $\hat{\lambda}$ is the q-core and $f^{\lambda}(q)$ is the dimension of the representation induced by the q-quotient. These entities are evaluated as follows.

Expand λ in its row determinatal form $\lambda = |[l_i - i + j]| \equiv |[\beta_{ij}]|$ so that $\beta_{ij} > \beta_{i+1,j}$ and $\beta_{ij} < B_{ij+1}$. A column sequence β_{ij}, $1 \leqslant i \leqslant j$, with $\beta_{j+1,j} < 0$ is called a β-sequence for λ. The Nakayama formula involves the sequential stripping of q-rims from the λ frame. Each q-rim removal corresponds to the reduction of a suitable β_k in a β-sequence to $\beta_k - q$. Examination of the determinantal form shows that if and only if $0 \leqslant \beta_k - q \neq \beta_i$ for all $i \neq k$ will the resulting reduced β-sequence (perhaps reordered) be itself a β-sequence giving a nonvanishing contribution to the character. These conditions are conveniently recorded in graphic form on a q-columned abacus as introduced by James. The columns 0 to $q-1$ classify the β-elements by their value modulo q. A β-sequence is recorded by placing a bead in each position β_k when the positions are numbered in natural order starting with zero in the upper left-hand corner and proceeding left to right and downward. Nonvanishing q-rim removals then correspond to any possible move of one bead upward one position in a fixed column. Moving all beads to the uppermost positions gives the β-sequence determing the q-core $\hat{\lambda}$. The number of unit moves in each column is the column q-weight $w^{(i)}$ with the q-weight being the total number of moves $w = \Sigma w^{(i)}$. The initial bead sequence in a fixed column determines a β-sequence for an irrep $\lambda^{(i)} \vdash w^{(i)}$. The direct product of these irreps $\lambda^{(i)}$ forms the q-quotient of λ with induced dimension

$$f^{\lambda}(q) = \binom{w}{w^{(0)} w^{(1)} \cdots w^{(q-1)}} \bigotimes_i |\lambda^{(i)}|. \tag{3}$$

The phase is determined by numbering the beads in natural ordering in their initial position and comparing this to the beads numbered in natural order in the q-core positions. The parity of the premutation σ necessary to bring these two orderings to coincidence determines the phase $\operatorname{sgn}\sigma$. Before proceeding we illustrate the above for $q = 3$ and

$$\lambda = [4^2,2^2] = \begin{vmatrix} [4] & [5] & [6] & [7] \\ [3] & [4] & [5] & [6] \\ [0] & [1] & [2] & [3] \\ \varnothing & [0] & [1] & [2] \end{vmatrix}$$

with β-sequence (7,6,3,2) graphed as

$$\begin{array}{cccc} & \cdot & \cdot & 1 \\ & 2 & \cdot & \cdot \\ & 3 & 4 & \\ \hline w^{(i)} & 2 & 2 & 0 \\ \lambda^{(i)} & [1^2] & [2] & [0] \end{array}$$

with three-core

$$\begin{matrix} 2 & 4 & 1 \\ 3 & \cdot & \cdot \end{matrix} = [0] \quad \text{and} \quad \sigma = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

so that $[4^2,2^2](3^4) = -\binom{4}{2} = -6$. Thus for $b > w$ the character must vanish and Eq. (2) is established for $b = w$. To extend this result for $b < w$ and relate it to the DC decomposition let us examine more closely the procedure for $q = 2$.

Each bead move corresponds to removing a segment [2] or $[1^2]$ from the preceding frame λ. That is the skew diagrams $\lambda/[2]$ and $\lambda/[1^2]$ contain common constituents in which separate (unsymmetrized) modes are removed from distinct rows and columns. Because of the sign change these constituents cancel and do not contribute to the Nakayama formula. The abacus structure counts only those constituents of $\lambda/[2]$ that correspond to the removal of symmetrized nodes and analogously antisymmetrized nodes in $\lambda/[1^2]$. In the abacus structure a move in which the immediately preceding position in the other row is occupied by a bead requires a reordering of the bead numbering to be a β-sequence. The transposition of rows in the determinantal form results in a (-1) phase factor and the move corresponds to removing an antisymmetrized segment $[1^2]$ from the λ-

frame. A move requiring no reordering to obtain a β-sequence corresponds to removing the symmetrized segment [2] from the λ-frame. The permutation σ gives the cummulative number of transpositions in ordering necessary to reach the two-core numbered in natural order. The irreps $\lambda^{(i)}$ determine the symmetry of the sequences of moves within a given column by which the two-core positions can be reached. *The decomposition of the outer product induced by the two-quotient then determines the symmetries of the sequences of all moves within all columns by which the two core can be reached.* It follows for $L_2 = w$, the necessary and sufficient condition for the diagonal element

$$\begin{matrix}\begin{bmatrix}\lambda & \lambda_1 & \lambda_2\\ \lambda_1 & {}_1\lambda_1 & \lambda_2\\ \lambda_2 & \lambda_2 & 0\end{bmatrix} & \begin{matrix}r\\ s\end{matrix}\\ \begin{matrix}r & s\end{matrix} & \end{matrix}$$

to be nonzero is that λ_2 be contained in the induction from the two-quotient of λ and ${}_1\lambda_1$ be the two-core of λ. For fixed ${}_i\lambda_j$ these elements contribute to the character in proportion to the multiplicity $|\lambda_2{:}\lambda^{(0)}\lambda^{(1)}|$, where $\lambda^{(0)}\lambda^{(1)}$ is the two-quotient of λ. Illustrative examples are given in Table I.

Extension to the case $b < w$ proceeds in the same manner except one does not continue to reduce to the two-core but must sum over all intermediate diagrams related to irreps μ within $w - b$ moves of the two-core. Each intermediate diagram itself has a two-quotient $\mu^{(0)}\mu^{(1)}$ such that λ_2 must occur in the induction $(\lambda^{(0)}/\mu^{(0)})(\lambda^{(1)}/\mu^{(1)})$, which dimension determines $f^{\mu}(q)$ in the sum

$$\lambda(q^b\pi) = \sum_{\mu} \operatorname{sgn}\sigma_{\mu} f^{\mu}(q)\mu(\pi). \tag{4}$$

The phase $\operatorname{sgn}\sigma_{\mu}$ is that necessary to change the bead numbering in the pattern corresponding to λ to the natural order for the pattern corresponding to μ. We illustrate by calculating the character [4,3,1] $(2^k, 1^{8-2k})$ via Eq. (4) and list the results in Table II. Extension to $q > 2$ proceeds in the same manner.

IV. DIAGONAL DCME AND DUALITY

Because of the identifications due to duality other sums involving the diagonal DCME can be obtained. The identification of the DCME as a recoupling coefficient of Un requires

$$\sum_{\substack{\lambda\lambda_i\\ rr_j}} \langle\lambda\rangle_n \begin{matrix}\begin{bmatrix}\lambda & \lambda_1 & \lambda_2\\ \lambda_1 & {}_1\lambda_1 & {}_1\lambda_2\\ \lambda_2 & {}_1\lambda_2 & {}_2\lambda_2\end{bmatrix} & \begin{matrix}r\\ r_1\\ r_2\end{matrix}\\ \begin{matrix}r & r_1 & r_2\end{matrix} & \end{matrix} = \langle {}_1\lambda_1\rangle_n \langle {}_2\lambda_2\rangle_n \langle {}_1\lambda_2\rangle_n\,, \tag{5}$$

where $\langle\lambda\rangle_n$ is the dimension of the irrep λ in Un. Given the irreps ${}_1\lambda_1$, ${}_2\lambda_2$, and ${}_1\lambda_2$ the dimension n may be chosen arbitrarily within the range ${}_1n_1 + {}_2n_2 + 2{}_1n_2 \geqslant n \geqslant \operatorname{Max}({}_1n_1, {}_2n_2, {}_1n_2)$, where ${}_in_j$ is the number of rows in the frame ${}_i\lambda_j$.

Similarly the DCME of the decomposition $S_{iL} \setminus S_L / S_{L_j}$ can be identified with DCME of the decomposition $U_i n \setminus Un / Un_j$ for suitably chosen ${}_in$ and n_j. A DC representative specified by

$$\begin{bmatrix} n & n_1 & n_2\\ n_1 & {}_1n_1 & {}_1n_2\\ n_2 & {}_1n_2 & {}_2n_2\end{bmatrix}$$

simply interchanges the distinct dimensions ${}_1n_2 \leftrightarrow {}_1n_2$ so in the defining irrep it has the simple block matrix form

$$\begin{bmatrix} E_{{}_1n_1} & 0 & 0 & 0\\ 0 & 0 & E_{{}_1n_2} & 0\\ 0 & E_{{}_1n_2} & 0 & 0\\ 0 & 0 & 0 & E_{{}_2n_2}\end{bmatrix},$$

where E_n is the unit matrix of dimension n. The character of this element in Un is given by

$$\begin{aligned}&\{\lambda\}(1^{n-{}_1n_2}, -1^{{}_1n_2})\\ &\quad = \sum_{\substack{\lambda_j\lambda_j\\ r\ r_j}} \langle {}_1\lambda_1\rangle_n \langle {}_2\lambda_2\rangle_n \langle {}_1\lambda_2\rangle_n\\ &\qquad \times \begin{matrix}\begin{bmatrix}\lambda & \lambda_1 & \lambda_2\\ \lambda_1 & {}_1\lambda_1 & {}_1\lambda_2\\ \lambda_2 & {}_1\lambda_2 & {}_2\lambda_2\end{bmatrix} & \begin{matrix}r\\ r_1\\ r_2\end{matrix}\\ \begin{matrix}r & r_1 & r_2\end{matrix} & \end{matrix}.\end{aligned} \tag{6}$$

It is convenient to use the Frobenius–Schur formula relating characters in Gln and S_L in the form $\Sigma_{\lambda}\lambda(\mu)\{\lambda\}(\epsilon) = S_{\mu}(\epsilon)$.

TABLE I. Evaluation of two-cores, two-quotients, and constraints on multiplicity.

Irrep $\lambda \vdash 2L$	Trace $\lambda(2^L)$	Two-core $\hat{\lambda}$	Two-quotient $\lambda^{(0)}\lambda^{(1)}$	Multiplicity resolution σ assumes values [2] and $[1^2]$
[3,2,1]	-0	[3,2,1]	[0][0]	$\lvert\lambda{:}[2,1]\otimes\sigma\rvert = 1$
$[4,2,1^2]$	-6	[0]	$[1^2][2]$	$\lvert\lambda{:}[3,1]\otimes[1^2]\rvert = 1 = \lvert\lambda{:}[2,1^2]\otimes[1^2]\rvert$
$[3^2,2]$	-6	[0]	$[1^2][2]$	$\lvert\lambda{:}[3,1]\otimes[1^2\rvert = 1 = \lvert\lambda{:}[2,1^2]\otimes[1^2]\rvert$
[4,3,1]	-2	[0]	$[2^2][0]$	$\lvert\lambda{:}[2^2]\otimes[1^2]\rvert = 1 = \lvert\lambda{:}[3,1]\otimes\sigma\rvert$
[5,3,2]	-10	[0]	$[1^2][3]$	$\lvert\lambda{:}[4,1]\otimes[1^2]\rvert = 1 = \lvert\lambda{:}[3.1^2]\otimes[1^2]\rvert = \lvert\lambda{:}[3,2]\otimes\sigma\rvert$
$[4,3,2^2,1]$	-5	[0]	$[2^3][0]$	$\lvert\lambda{:}[2^3]\otimes[1^2]\rvert = 1 = \lvert\lambda{:}[3,1^3]\otimes\sigma\rvert = \lvert\lambda{:}[2^2,1^2]\otimes\sigma\rvert$, $\lvert\lambda{:}[3,2,1]\otimes\sigma\rvert = 2$
$[4^2,2^2]$	$+80$	[0]	[2,1][2,1]	$\lvert\lambda{:}\lambda_i\otimes[2]\rvert = 1\{\lambda_i\lvert[4,2],[3,1^3],[2^3],[3^2],[4,1^2],[2^2,1^2]\}$, $\lvert\lambda{:}[3,2,1]\otimes[2]\rvert = 2$
[6,5,3,2]	-140	[0]	$[1^2][3^2]$	$\lvert\lambda{:}[4^2]\otimes[1^2]\rvert = 1 = \lvert\lambda{:}[3^2]\otimes[1^2]\rvert$, $\lvert\lambda{:}[4,3,1]\otimes[2]\rvert = 2$, $\lvert\lambda{:}[4,3,1]\otimes[1^2]\rvert = 3$
[4,2,1] for element	$+1$ $(2^3,1)$	[1]	[3][0]	$\begin{bmatrix}[4,2,1] & [3,1] & [3]\\ [3,1] & [1] & [3]\\ [3] & [3] & [0]\end{bmatrix} = +1\sum\begin{bmatrix}[4,2,1] & \lambda_1 & [2,1]\\ \lambda_1 & [1] & [2,1]\\ [2,1] & [2,1] & [0]\end{bmatrix} = 0$

TABLE II. Characters of $[4,3,1]\,(2^k,1^{8-2k})$.

k	μ	$\mu^{(0)}$	$\lambda^{(0)}/\mu^{(0)a}$	Character
4	[0]	[0]	$[2^2]$	-2
3	[2]	[1]	[2,1]	-2
2	[4]	[2]	[2]	-1
	$[2,1^2]$	$[1^2]$	$[1^2]$	$+3$ $\}+2$
1	$[4,1^2]$	[2,1]	[1]	$+10$

[a] $(\lambda^{(1)}/\mu^{(1)}) = [0]$ for all entries

The power sum $S_\mu(\epsilon)$ for $\mu \vdash L$ with cycle structure $\mu = \otimes_k (k)^{m_k}$ for an element with eigenvalues $(\epsilon) = (\epsilon_1,\epsilon_2,\ldots,\epsilon_n)$ in the defining irrep is defined as $S_\mu(\epsilon) = \otimes_k S_k(\epsilon)^{m_k}$ with $S_k(\epsilon) = \Sigma \epsilon_i{}^k$. For our purposes we may choose ${}_2n_2 = 0 = {}_2L_2$. Substituting the character formulas in terms of DCME we obtain the sum relation

$$\sum |{}_1\lambda'_1|\,|\lambda'_2| \begin{bmatrix} \lambda & \lambda'_1 & \lambda'_2 & r' \\ \lambda'_1 & {}_1\lambda'_1 & \lambda'_2 & s' \\ \lambda'_2 & \lambda'_2 & 0 & \\ r' & s' & & \end{bmatrix} \times \langle {}_1\lambda_1 \rangle_{{}_1n_1} \langle \lambda_2 \rangle_{n_2} \begin{bmatrix} \lambda & \lambda_1 & \lambda_2 & r \\ \lambda_1 & {}_1\lambda_1 & \lambda_2 & s \\ \lambda_2 & \lambda_2 & 0 & \\ r & s & & \end{bmatrix} = n^{L_2}\,{}_1n_1{}^{L_1}. \tag{7}$$

We also wish to point out the character of the DC representative

$$\begin{bmatrix} n & n_1 & n_2 \\ n_1 & {}_1n_1 & n_2 \\ n_2 & n_2 & 0 \end{bmatrix}$$

in Un can be evaluated by imbedding Un/S_n in the natural way[3] such that the defining irrep subduces as $\{1\}_n\downarrow = [n] + [n-1,1]$. Taking symmetrized powers on both sides we have $\{\lambda\}_n\downarrow = ([n] + [n-1,1])\odot\lambda = \Sigma_k [n-1,1]\odot\lambda/[k]$, where $\odot$ indicates the inner plethysm (symmetrized inner product in S_n) and $\lambda/[k]$ ranges over all skew diagrams resulting from all possible one-rowed deletions. Since the DC representative is also an element of S_n one can evaluate the character of the class $(2^{n_2},1^{n_1})$ over the representations subduced in S_n by $\{\lambda\}\downarrow$. This has limited utility, however, because the inner plethysms are difficult to evaluate except for some special cases.

[1] J. J. Sullivan, J. Math Phys. **24**, 424 (1983).

[2] G. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Vol. 16, *Encyclopedia of Mathematics and its Applications* (Addison–Wesley, Reading, MA, 1981).

[3] P. H. Butler and R. C. King, J. Math. Phys. **14**, 1176 (1973).

Nonlinear evolution equations associated with a Riemann–Hilbert scattering problem

A. Degasperis,[a] P. M. Santini,[a] and M. J. Ablowitz
Department of Mathematics and Computer Science, Clarkson University, Potsdam, New York 13676

(Received 7 January 1985; accepted for publication 29 March 1985)

In an earlier paper nonlinear evolution equations associated with a Riemann–Hilbert scattering problem, which reduces, in an appropriate limit, to the Zakharov–Shabat–AKNS scattering problem, were considered. Here we discuss certain necessary constraints associated with the scattering problem and their impact upon the associated evolution equations. Moreover, the direct linearization of the nonlinear evolution equations and an algorithm to construct an N-soliton solution are given.

I. INTRODUCTION

The inverse spectral (or scattering) transform (IST) method is a well-established technique to solve and investigate certain nonlinear partial differential equations of evolution type, a number of which are physically relevant.[1]

Attention has been recently given to the intermediate long wave (ILW) equation[2–7] because it brings into the field some novelty; that is, it is an integrodifferential, rather than purely differential, nonlinear equation, that is, integrable via a spectral problem based on a differential Riemann–Hilbert (RH) boundary value problem rather than an ordinary differential equation. Moreover, the ILW equation depends on a parameter which we call η, in such a way as to coincide, as η vanishes, with the Korteweg–de Vries (KdV) equation,[8] and, as η goes to infinity, with the Benjamin–Ono equation.[9–11]

In analogy with the well-known connection between the Korteweg–de Vries equation and the modified Korteweg–de Vries equation, a modified ILW equation (whose $\eta\to 0$ limit is the modified KdV equation) has also been introduced and investigated.[12,13]

Further progress in this direction has been made by extending[14] the class of intermediate-type long-wave equations, and by introducing[15,16] an intermediate version of the Kadomtsev–Petviashvili equation[17] (whose $\eta\to 0$ limit is of course the Kadomtsev–Petviashvili equation).

More recently,[18] a class of matrix nonlinear integral evolution equations was generated through the following 2×2 matrix spectral problem:

$$\psi^-(x,z) = G(x,z)\psi^+(x,z), \quad x\in\mathbb{R}, \tag{1a}$$

$$G(x,z)\equiv I + z\sigma_3 + U(x), \tag{1b}$$

where I is the identity matrix, $\sigma_3 = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$, z plays the role of spectral parameter, and $U(x)$ is a (complex) z-independent potential function.

Given the matrix function $U(x)$, (1) defines a homogeneous RH boundary value problem on a strip of the complex x plane. The matrices $\psi^{\pm}(x,z)$ are the boundary values of a function $\Psi(x,z)$ holomorphic in the horizontal strip between $\operatorname{Im} x = 0$ and $\operatorname{Im} x = \eta$:

$$\psi^+(x,z)\equiv\lim_{y\downarrow 0}\Psi(x+iy,z), \quad x\in\mathbb{R}, \tag{2a}$$

$$\psi^-(x,z)\equiv\lim_{y\uparrow \eta}\Psi(x+iy,z), \quad x\in\mathbb{R}. \tag{2b}$$

It turns out that $\Psi(x,z)$ can be written as

$$\Psi(x,z) = \exp[-i\zeta(z)x]\left(\frac{1}{2i\eta}\int_{-\infty}^{\infty}\coth\left[\frac{\pi}{\eta}(x'-x)\right]\times h(x',z)dx' + \text{const}\right), \quad 0<\operatorname{Im} x<\eta, \tag{3}$$

where

$\exp[\eta\zeta(z)] = I + z\sigma_3$ and $h(x,z)$, defined by

$$h(x,z) = -(I+z\sigma_3)^{-1}\exp[i\zeta(z)x]U(x)\psi^+(x,z) \tag{4}$$

is Hölder continuous on $x\in\mathbb{R}$ and satisfies the condition $|\int_{-\infty}^{\infty} h(x,z)dx| < \infty$. Moreover, formula (3) implies the following periodicity condition:

$$\psi^-(x,z) = (E\psi^+)(x,z), \tag{5a}$$

where $E = \exp(i\eta\,\partial_x)$ is the formal shift operator

$$(Ef)(x) = f(x+i\eta). \tag{5b}$$

It was shown in Ref. 18 that the linear problem (1) and the associated class of evolution equations reduce, in the limit $\eta\to 0$, to the generalized Zakharov–Shabat–AKNS scattering problem[19] and to the associated class of nonlinear evolution equations.[19,20] Moreover, for the class of nonlinear equations associated with (1), an infinite family of conservation laws was derived and only elementary properties of the spectral problem were essential for that derivation. In fact, the emphasis in Ref. 18 was mainly on the novel nonlinear evolution equations, such as an intermediate version of the nonlinear Schrödinger equation, and on their associated Lax pair.

In this paper we present new results concerning the RH boundary value problem (1) and the class of evolution equations associated with it.

II. THE BASIS CONSTRAINTS

In the theory of matrix RH problems[21] of the type (1) an important role is played by the determinant of $G(x,z)$. In our case

$$\det G(x,z) = 1 - z^2 - z\,\operatorname{tr}(\sigma_3 U(x)) + \operatorname{tr} U(x) + \det U(x). \tag{6}$$

All the results of this paper are derived when the potential matrix $U(x)$ is subjected to the following two scalar constraints:

$$\operatorname{tr}(\sigma_3 U(x)) = 0, \tag{7a}$$

[a] Permanent address: Dipartimento di Fisica, Universita di Roma I, 00185 Roma, Italy.

 0022-2488/85/102469-04$02.50

$$\operatorname{tr} U(x) + \det U(x) = 0, \tag{7b}$$

or equivalently

$$U(x) = \sqrt{1 + Q^2(x)} - 1 + Q(x), \tag{8}$$

where $Q(x)$ is the off-diagonal part of $U(x)$. In this case the determinant of $G(x,z)$ takes the particularly simple form

$$\det G(x,z) = 1 - z^2, \tag{9}$$

independent of x with the following important consequences.

(i) The matrix $G(x,z)$ is invertible for every $x \in \mathbb{R}$; this is a necessary condition for the solvability of (1).

(ii) The total index κ of the matrix RH problem (1) is zero, since

$$\kappa = (2\pi)^{-1}[\arg(\det G(x,z))]_{-\infty}^{+\infty}, \tag{10}$$

where $[\theta(x)]_{-\infty}^{\infty} = \theta(\infty) - \theta(-\infty)$.

Then an important theorem due to Gohberg and Krein[22] shows that "generically" the two partial indices κ_1, κ_2 ($\kappa = \kappa_1 + \kappa_2$) are both zero. This fact guarantees the existence and uniqueness of a bounded fundamental matrix $\Psi(x,z)$ associated with (1).

III. THE REDUCED CLASS OF EVOLUTION EQUATIONS

The existence and uniqueness of bounded solutions of (1) can be used in the construction of the IST method for the class of evolution equations introduced in Ref. 18, if and only if the constraints (7) are compatible with the evolution equations themselves. It will be shown in the following that this is indeed the case. Hence the constraints (7), introduced as requirements for the solvability of (1), are in fact a reduction of the class of equations introduced in Ref. 18 to the following class of matrix nonlinear evolution equations:

$$Q_t = \sigma_3 \gamma(\mathscr{L}) Q, \tag{11}$$

where

$$\mathscr{L}F \equiv i\sigma_3(\sqrt{1+Q^2}\,\mathscr{D}F + \tfrac{1}{2}Q\mathscr{D}^{-1}([Q,F]/\sqrt{1+Q^2})), \tag{12}$$

F is off-diagonal, $\gamma(y)$ is an arbitrary polynomial in y, and

$$(\mathscr{D}f)(x) \equiv \frac{1}{\eta}\int_{-\infty}^{\infty} dy \left\{\sinh\left[\frac{\pi}{\eta}(y-x)\right]\right\}^{-1} f(y), \tag{13a}$$

$$(\mathscr{D}^{-1}f)(x) \equiv -\frac{1}{\eta}\int_{-\infty}^{\infty} dy \coth\left[\frac{\pi}{\eta}(y-x)\right] f(y). \tag{13b}$$

In order to show that (8) is a reduction for the class of evolution equations associated with (1), one has to show that the set of the matrices U satisfying (8) is closed with respect to the elementary deformations $\delta U^{(n)}$ such that $\delta\psi^{\pm} = B^{\pm}\psi^{\pm}$. Namely, one has to prove that if $U(x)$ satisfies (8) then

$$\delta U^{(n)} = (2\sqrt{1+Q^2})^{-1}\{Q,\delta Q^{(n)}\} + \delta Q^{(n)}. \tag{14}$$

In Ref. 18 it was shown that the elementary deformations $\delta U^{(n)}$, such that $\delta\psi^{\pm} = B^{\pm}\psi^{\pm}$, are given by

$$\delta U^{(n)} = b_n\{((E-1)L^n\sigma_3)(1+U) + [L^n\sigma_3, U]\}, \tag{15}$$

with

$$B^{+}(x,z) = b_n \sum_{j=0}^{n} L^{n-j}\sigma_3 z^j, \quad b_n \text{ arbitrary constants}, \tag{16}$$

where $L^n\sigma_3$, defined in Ref. 18, is written here in the following more convenient form:

$$L^n\sigma_3 \equiv -\tfrac{1}{2}(i\mathscr{D}^{-1}+1)F_n - \tfrac{1}{2}(i\mathscr{D}-1)G_n, \tag{17}$$

in terms of the diagonal and off-diagonal matrices $F_n(x)$ and $G_n(x)$, respectively, which are constructed through the following recursion relations:

$$F_{n+1} = -\sigma_3((1+U_c)F_n + \tfrac{1}{2}\{\mathscr{D}G_n, Q\} + \tfrac{1}{2}[G_n, Q]), \tag{18a}$$

$$G_{n+1} = \sigma_3(\tfrac{1}{2}\{F_n, Q\} - (i/2)[\mathscr{D}^{-1}F_n, Q] + (i/2)\{\mathscr{D}G_n, (1+U_c)\} + \tfrac{1}{2}[G_n, U_c]), \tag{18b}$$

$$F_1 = 0, \quad G_1 = 2Q, \tag{18c}$$

where $[\,,]$ and $\{\,,\}$ are the usual commutator and anticommutator between matrices and U_c is the diagonal part of U. The class of evolution equations is obtained by replacing $\delta U^{(n)}/b_n$ by U_t.

Using Eqs. (15), (17), and (8), one can show that

$$\delta U^{(n)} - \{(2\sqrt{1+Q^2})^{-1}Q, \delta Q^{(n)}\} - \delta Q^{(n)} = -\frac{Q^2(\operatorname{tr} F_n)}{2\sqrt{1+Q^2}} + \sqrt{1+Q^2} \times\left(F_n - \frac{1}{2\sqrt{1+Q^2}}[Q, G_n]\right). \tag{19}$$

Moreover, if (8) holds, one can prove by induction that the recursion equations (18) decouple in the following way:

$$F_n = (2\sqrt{1+Q^2})^{-1}[Q, G_n], \tag{20a}$$

$$G_{n+1} = \mathscr{L}G_n, \quad G_1 = 2Q. \tag{20b}$$

Then, from (19) and (20a), one immediately gets

$$\delta U^{(n)} = (2\sqrt{1+Q^2})^{-1}\{Q, \delta Q^{(n)}\} + \delta Q^{(n)}. \tag{21}$$

From (15), (17), and (20) one finally obtains the evolution equations (11) and from (16), (17), and (20) one gets the corresponding time evolution of function ψ,

$$\psi_t^{+} = \frac{\alpha_n}{2}\left(\sum_{j=0}^{n-1} z^j L^{n-j}\sigma_3 + z^n\sigma_3\right)\psi^{+}, \tag{22}$$

where

$$L^j\sigma_3 = -\tfrac{1}{2}(i\mathscr{D}^{-1}+1)((\sqrt{1+Q^2})^{-1}[Q, \mathscr{L}^{j-1}Q]) - (i\mathscr{D}-1)\mathscr{L}^{j-1}Q \tag{23}$$

and the polynomial $\gamma(y)$, introduced in (11), is taken to be $\gamma(y) \equiv \alpha_n y^n$.

The first three equations of the class (11) are (see Ref. 18) the following:

(i) an intermediate wave equation

$$\gamma(y) = -icy, \quad U = \begin{pmatrix} u-1 & \rho v \\ v & u-1 \end{pmatrix}, \quad \rho \in \mathbb{R}, \tag{24a}$$

$$v_t = c\sqrt{1+\rho v^2}\,\mathscr{D}v, \quad u = \sqrt{1+\rho v^2};$$

(ii) an intermediate nonlinear Schrödinger equation

$$\gamma(y) = icy^2, \quad U = \begin{pmatrix} u-1 & i\rho\psi^* \\ i\psi & u-1 \end{pmatrix}, \quad \rho \in \mathbb{R}, \tag{24b}$$

$$i\psi_t + c[\sqrt{1-\rho|\psi|^2}\,\mathscr{D}(\sqrt{1-\rho|\psi|^2}\,\mathscr{D}\psi)$$

$$-\rho\psi\mathscr{D}^{-1}\mathrm{Re}(\psi^*\mathscr{D}\psi)] = 0,$$

$$u = \sqrt{1-\rho|\psi|^2};$$

(iii) an intermediate modified KdV equation

$$\gamma(y) = icy^3, \quad U = \begin{pmatrix} u-1 & \rho v \\ v & u-1 \end{pmatrix}, \quad \rho\in\mathbb{R},$$

$$v_t = c\sqrt{1+\rho v^2}\,\mathscr{D}(\sqrt{1+\rho v^2}\,\mathscr{D}(\sqrt{1+\rho v^2}\,\mathscr{D}v)\mathscr{D}^{-1}v\mathscr{D}v),$$

$$u = \sqrt{1+\rho v^2}. \tag{24c}$$

Taking the $\eta\to 0$ limit of Eqs. (24a)–(24c) one obtains the linear wave equation, the nonlinear Schrödinger equation, and the modified KdV equation, respectively.

The limit $\eta\to\infty$ can be immediately performed,[18] replacing $\mathscr{D}$ and $\mathscr{D}^{-1}$ by H and $-H$, respectively, where

$$(Hf)(x) \equiv \frac{1}{i\pi}\int_{-\infty}^{\infty} dy\,(y-x)^{-1}f(y) \tag{25}$$

is the Hilbert transform. We conclude this section by noticing that Eq. (24a) can be written in the following simple and suggestive form:

$$\mathscr{D}^{-1}\theta_t = c\sin\theta, \quad \theta = \theta(x,t), \tag{26}$$

where $v(x,t) = i\sin\theta(x,t)$. In the limit $\eta\to\infty$ Eqs. (26) become

$$H\theta_t = -c\sin\theta, \tag{27}$$

which we refer to as the sine-Hilbert equation, in analogy with the sine-Gordon equation $\theta_{xt}\sin\theta$.

IV. THE DIRECT LINEARIZATION

Postponing to a separate paper the presentation of the IST method for the solution of the Cauchy problem associated with Eq. (11), we now present the direct linearization (DL)[23,24] for the class (11).

The DL is an algebraic approach based on the existence of a linear integral equation which provides a large class of solutions of the evolution equations (11).

Proposition: Let $\mu^{\pm}(x,t,z)$ be the solutions of the integral equations

$$\mu^{\pm}(x,t,z) + \int_l \mu^{\pm}(x,t,z')R^{\pm}(x,t,z')\frac{d\lambda(z')}{z'-z} = I, \tag{28}$$

where l and $d\lambda(z)$ are an arbitrary contour and measure

$$R^{\pm}(x,t,z) \equiv \psi_0^{\pm}(x,t,z)A(z)(\psi_0^{\pm}(x,t,z))^{-1},$$

$$A(z) \text{ arbitrary}, \tag{29}$$

and $(\psi_0^{\pm}(x,t,z),\ U(x,t))$ is a given solution of (1) and (22) (where, of course, ψ and U are replaced by ψ_0 and U_0). Assuming that the homogeneous version of (28) has only the trivial solution, then the matrices $\psi^{\pm}$, defined through

$$\psi^{\pm}(x,t,z) \equiv \mu^{\pm}(x,t,z)\psi_0^{\pm}(x,t,z), \tag{30}$$

solve Eq. (1) if the potential $U(x,t)$ is given by

$$U(x,t) = U_0(x,t) + \int_l [\mu^-(x,t,z)R^-(x,t,z)\sigma_3 - \sigma_3\mu^+(x,t,z)R^+(x,t,z)]d\lambda(z). \tag{31}$$

The proof is direct and as in the spirit of the method as it was introduced in Ref. 23; the constructive procedure used to obtain Eqs. (29) and (31) is illustrated in detail in Ref. 24.

V. N-SOLITON SOLUTION

The N-soliton solution for the class (11) can be obtained by setting $U_0(x,t) = 0$ (and then $\psi_0^+(x,t,z) = \exp[-i\zeta(z)x + (\alpha_n/2)z^n\sigma_3 t]$),

$$A_{lm}(1-\delta_{lm})\theta((-1)^l\mathrm{Im}\,z) \tag{32a}$$

[$\theta(x)$ is the usual step function], and

$$d\lambda(z) \equiv \begin{cases} \sum_{j=1}^{N} c_j\,\delta(z-z_j), & \mathrm{Im}\,z_j>0, \quad \mathrm{Im}\,z>0, \\ -\sum_{j=1}^{N} \bar{c}_j\,\delta(z-\bar{z}_j), & \mathrm{Im}\,\bar{z}_j<0, \quad \mathrm{Im}\,z<0. \end{cases} \tag{32b}$$

In this case Eq. (28) reduces to a $2N$th-order algebraic system; in particular, if $N=1$ we have the following one-soliton solution:

$$u_{11}(x,t) = u_{22}(x,t) = (\bar{z}_1 - z_1)\sinh[\eta(k_1-\bar{k}_1)]/d(x,t), \tag{33a}$$

$$u_{12}(x,t) = \bar{c}_1 e^{\eta\bar{k}_1 - \gamma}1[\cosh(\eta\bar{k}_1)e^{-\phi_+(k_1)} + \cosh(\eta k_1)e^{-\phi_-(\bar{k}_1)}]/d(x,t), \tag{33b}$$

$$u_{21}(x,t) = c_1 e^{-\eta k_1 - \gamma}1[\cosh(\eta\bar{k}_1)e^{\phi_+(k_1)} + \cosh(\eta k_1)e^{\phi_-(\bar{k}_1)}]/d(x,t), \tag{33c}$$

where

$$\bar{z}_1 \equiv z(\bar{k}_1) = \tan(\eta\bar{k}_1), \quad z_1 \equiv z(k_1) = \tan(\eta k_1), \tag{34a}$$

$$d(x,t) \equiv \cosh[\eta(k_1-\bar{k}_1)] + \cosh[\phi_+(k_1) - \phi_-(\bar{k}_1)], \tag{34b}$$

$$\phi_{\pm}(k) \equiv 2ikx - \alpha_n(z(k))^n t \pm \gamma_1, \tag{34c}$$

$$-c_1\bar{c}_1/(z_1-\bar{z}_1)^2 \equiv e^{2\gamma_1 + \eta(k_1-\bar{k}_1)}, \tag{34d}$$

and $k_I \equiv \frac{1}{2}\,\mathrm{Im}(k_1-\bar{k}_1)$ and $k_R \equiv \frac{1}{2}\,\mathrm{Re}(k_1-\bar{k}_1)$ satisfy the inequality

$$k_I^2 + k_R^2 - (\pi/2\eta)k_I < 0. \tag{35}$$

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant No. DMS-8202117, the Office of Naval Research under Grant No. N00014-76-C-0867, and the Air Force Office of Scientific Research under Grant No. AFOSR-84-0005.

[1] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981); F. Calogero and A. Degasperis, *Spectral Transform and Solitons I* (North-Holland, Amsterdam, 1982); R. K. Dodd, J. C. Eilbeck, J. D. Gibbon, and H. C. Morris, *Solitons and Nonlinear Waves* (Academic, New York, 1983).

[2] R. I. Joseph, J. Phys. A **10**, L225 (1977).

[3] T. Kubota, D. R. S. Ko, and D. Dobbs, J. Hydronaut. **12**, 157 (1978).

[4] R. I. Joseph and R. Egri, J. Phys. A **11**, L97 (1978).

[5] H. H. Chen and Y. C. Lee, Phys. Rev. Lett. **43**, 264 (1979).

[6] J. Satsuma, M. J. Ablowitz, and Y. Kodama, Phys. Lett. A **73**, 283 (1979).

[7] Y. Kodama, M. J. Ablowitz, and J. Satsuma, J. Math. Phys. **23**, 564 (1982).

[8] C. S. Gardner, J. M. Green, M. D. Kruskal, and R. M. Miura, Phys. Rev. Lett. **19**, 1095 (1967).

[9] T. B. Benjamin, J. Fluid Mech. **29**, 559 (1967).

[10]H. Ono, J. Phys. Soc. Jpn. **39**, 1082 (1975).
[11]A. S. Fokas and M. J. Ablowitz, Stud. Appl. Math. **68**, 1 (1983).
[12]J. Satsuma, T. R. Taha, and M. J. Ablowitz, J. Math. Phys. **25**, 900 (1984).
[13]J. Gibbons and B. Kupershmidt, Phys. Lett. A **79**, 31 (1980).
[14]D. R. Lebedev and A. O. Radul, "Generalized internal long wave equations: construction, Hamiltonian structure, and conservation laws," preprint ITEP-102 (1982) (to appear in Commun. Math. Phys.).
[15]V. B. Matveev and M. A. Sall, Dokl. Akad. Nauk. SSSR **261**, 533 (1981).
[16]A. I. Bobenko, V. B. Matveev, and M. A. Sall, Dokl. Akad. Nauk. SSSR **265**, 1357 (1982) [Sov. Phys. Dokl. **27**, 610 (1982)].
[17]B. B. Kadomtsev and V. I. Petviashvili, Dokl. Akad. Nauk. SSSR **192**, 753 (1970) [Sov. Phys. Dokl. **15**, 539 (1970)].
[18]A. Degasperis and P. M. Santini, Phys. Lett. A **98**, 240 (1983).
[19]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Stud. Appl. Math. **53**, 249 (1974).
[20]F. Calogero and A. Degasperis, Nuovo Cimento B **32**, 201 (1976).
[21]N. I. Muskhelishvili, *Singular Integral Equations* (Gordon and Breach, New York, 1967).
[22]I. Gohberg and M. G. Krein, Usp. Mat. Nauk. **13**, 2 (1958).
[23]A. S. Fokas and M. J. Ablowitz, Phys. Rev. Lett. **47**, 1096 (1981).
[24]P. M. Santini, M. J. Ablowitz, and A. S. Fokas, J. Math. Phys. **24**, 892 (1984).

Transforms associated to square integrable group representations. I. General results

A. Grossmann, J. Morlet,[a] and T. Paul[b]
Centre de Physique Théorique,[c] Section II, Centre National de la Recherche Scientifique, Luminy, Case 907, F-13288 Marseille Cedex 9, France

(Received 20 December 1984; accepted for publication 12 April 1985)

Let G be a locally compact group, which need not be unimodular. Let $x \to U(x)$ $(x \in G)$ be an irreducible unitary representation of G in a Hilbert space $\mathscr{H}(U)$. Assume that U is square integrable, i.e., that there exists in $\mathscr{H}(U)$ at least one nonzero vector g such that $\int |(U(x)g,g)|^2\, dx < \infty$. We give here a reasonably self-contained analysis of the correspondence associating to every vector $f \in \mathscr{H}(U)$ the function $(U(x)g,f)$ on G, discussing its isometry, characterization of the range, inversion, and simplest interpolation properties. This correspondence underlies many properties of generalized coherent states.

I. INTRODUCTION

This paper is the first of a series concerned with applications of various families of "generalized coherent states" to quantum mechanics, wave propagation, and signal analysis.

Many properties of the classical (canonical) coherent states[1,2] are closely tied to the Weyl–Heisenberg group. In particular, the fundamental formula

$$\mathbf{1} = \int |z\rangle d^2 z \langle z| \tag{1.1}$$

is a way of writing the orthogonality relations[3,4] for the irreducible representation of that group.

Aslaksen and Klauder[5] have considered the analogous states for the two-parameter group of shifts and dilations and found that the "fiducial vector" ("analyzing wavelet" in our terminology) cannot be arbitrary, in contrast to the Weyl–Heisenberg case.

The same two-parameter group appeared in Ref. 6 in the study of decomposition of signals into "wavelets of constant shape"; the restriction on the analyzing wavelet was there called an "admissibility condition."

Another (equivalent) representation of the same group, together with an appropriate choice of the analyzing wavelet, has given rise to a realization of quantum mechanics on a Hilbert space of function analytic on a half-plane.[7,8] Other groups were used to define coherent states: SU(2) for spin coherent states,[9] SU(1,1) in Ref. 10, and a general definition was proposed by Perelomov.[11]

In this paper we shall be concerned with "coherent states" associated with certain representations of arbitrary (in particular not necessarily unimodular) locally compact groups.

Let G be a locally compact group, U a continuous irreducible representation of G in a Hilbert space $\mathscr{H}(U)$, and g a vector in $\mathscr{H}(U)$.

We consider the family of vectors

$$|x\rangle = U(x)g \quad (x \in G) \tag{1.2}$$

in $\mathscr{H}(U)$. This family depends on the choice of g.

Since U is irreducible, the linear span of the vectors $|x>$ is dense in $\mathscr{H}(U)$.

One can then ask the question whether there exists a (suitably normalized) invariant measure $d\mu(x)$ on G, such that

$$\int |x\rangle d\mu(x) \langle x| = \mathbf{1}, \tag{1.3}$$

where $|x\rangle\langle x|$ is defined by $|x\rangle\langle x|f> = (U(x)g,f)U(x)g$, and $\mathbf{1}$ is the identity operator in $\mathscr{H}(U)$.

The answer, in general, is no; this can be seen by taking $G = \mathbb{R}$ (additive), and by considering the one-dimensional irreducible representation space $\mathbb{C}$, with $U(x)$ the operator of multiplication by e^{ix}.

However, if the representation U is "square integrable" in a sense that will be defined below, then there exists a dense set of vectors in $\mathscr{H}(U)$, which give rise to (1.3). If g is such a vector (called "admissible") then the correspondence $f \to \psi$, with

$$\psi(x) = \langle x|f\rangle = (U(x)g,f) \quad (x \in G, \quad f \in \mathscr{H}(U)), \tag{1.4}$$

can be shown to be a multiple of an isometry between $\mathscr{H}(U)$ and $L^2(G,d\mu(x))$ and so (1.3) holds. The range of this transform is a closed subspace of $L^2(G,d\mu(x))$, and can be characterized by a reproducing kernel. If the group is unimodular, the set of admissible vectors is the whole space $\mathscr{H}(U)$, but this is not the case if the group is not unimodular (e.g., the affine group).

The purpose of this first paper is to give general results about transformations defined by (1.4). All the results derived here can be found in the mathematical literature and are part of the study of orthogonality relations for generalized square integrable representations. See, in particular, Refs. 12 and 13. We write them here in a form that is covenient for the applications we have in mind, using tools familiar to mathematical physicists (e.g., we give another proof of orthogonality relations with the help of quadratic forms).

The second paper of the series will be devoted to the particular case of the "$ax + b$" group, which has given rise to applications in applied mathematics.[6]

Further papers will be concerned with discrete versions of (1.2) and (1.3), with analyticity properties, special cases, and applications.

[a] Permanent address: ELF Aquitaine Company, ORIC Lab. 370 Bis Av. Napoleon Bonaparte, Rueil-Malmaison, France.
[b] Allocataire D.G.R.S.T.
[c] Laboratoire propre.

 0022-2488/85/102473-07$02.50

II. SQUARE INTEGRABLE REPRESENTATIONS; ADMISSIBLE VECTORS

In this section we present basic notions that will be used in the construction of transforms defined in Sec. IV.

A. Notations

G will denote a locally compact group, with identity e, and $x, y, \ldots$ elements of G. It is well known that there exists in such a group a left-invariant and a right-invariant (Haar) measure.[3,4] The left-invariant measure, with a fixed normalization, will be written dx. So

$$d(yx) = dx \quad (y \in G). \tag{2.1}$$

The right-invariant measure will be denoted by $d_R x$:

$$d_R(xy) = d_R(x) \quad (y \in G). \tag{2.2}$$

One has

$$d_R x = \Delta^{-1}(x)dx, \quad d(xy) = \Delta(y)dx, \tag{2.3}$$

where $\Delta(x)$ (the modular function) is a positive-valued character

$$\Delta(e) = 1, \quad \Delta(x) > 0, \quad \Delta(xy) = \Delta(x)\Delta(y). \tag{2.4}$$

If $\Delta(x) \equiv 1$, the group G is said to be unimodular. Notice a collision of terminologies: the group $GL(n,\mathbb{R})$ of nonsingular $n \times n$ matrices is unimodular, even though it contains nonunimodular matrices.

The inhomogeneous group $IGL(n,\mathbb{R})$, a semidirect product of translations in $\mathbb{R}^n$ and of $GL(n;\mathbb{R})$, is not unimodular.

When one deals with semidirect products, it is convenient to use the right-invariant measure $d_R x$, since it is the product of the right-invariant measures of the factors (see Ref. 14, p. 210).

We have

$$d(x^{-1}) = \Delta^{-1}(x)dx = d_R x. \tag{2.5}$$

If $x \to \Phi(x)$ is any function on the group, define

$$\check{\Phi}: \check{\Phi}(x) = \Phi(x^{-1}). \tag{2.6}$$

Notice that $\check{\check{\Phi}} = \Phi$.

Then

(Φ is left integrable) $\Leftrightarrow$ ($\check{\Phi}$ is right integrable)

and the corresponding integrals are equal, since

$$\int \check{\Phi}(x)d_R x = \int \Phi(x^{-1})d(x^{-1}) = \int \Phi(x')dx'.$$

We shall need the following statement: Let $\Phi(x)$ be a complex-valued function on the group G, such that

$$\bar{\Phi}(x) = \Phi(x^{-1}) \quad (x \in G) \tag{2.7}$$

[here $\bar{\Phi}(x)$ is the complex conjugate of $\Phi(x)$]. If, for some p, with $1 \leqslant p < +\infty$, we have $\Phi \in L^p(G,dx)$ and if Φ satisfies (2.7), then also $\bar{\Phi} \in L^p(G,dx)$.

Here $\Phi \in L^p(G,dx)$ means

$$\int |\Phi(x)|^p \, dx < \infty,$$

the integral being taken over G.

B. Left and right regular representation

Left regular: If $\Phi \in L^2(G,dx)$ and $a \in G$, we define $\lambda(a)\Phi \in L^2(G,dx)$ by

$$(\lambda(a)\Phi)(x) = \Phi(a^{-1}x) \quad (x \in G). \tag{2.8}$$

Right regular: If $\Psi \in L^2(G,d_R x)$ and $a \in G$, we define $\rho(a)\Psi \in L^2(G,d_R x)$ by

$$(\rho(a)\Psi)(x) = \Psi(xa) \quad (x \in G). \tag{2.9}$$

The two representations λ and ρ act unitarily—in different spaces in general, namely $L^2(G,dx)$ and $L^2(G,d_R x)$.

C. Definition of square integrable representations and of admissible vectors

Let $x \to U(x)$ be a strongly continuous unitary representation of the locally compact group G in a complex Hilbert space $\mathscr{H}(U)$.

A vector $g \in \mathscr{H}(U)$ is said to be *admissible* if

$$\int |(U(x)g,g)|^2 \, dx < +\infty. \tag{2.10}$$

In (2.10) the left-invariant measure dx can be replaced by the right-invariant measure $d_R x$, since

$$(U(x)g,g) = (g,U(x^{-1})g);$$

so

$$\int |(U(x)g,g)|^2 \, dx = \int |(g,U(x^{-1})g)|^2 \, dx = \int |(U(x^{-1})g,g)|^2 \, dx = \int |(U(x)g,g)|^2 \, d_R x.$$

Definition (2.1): U will be called square integrable if (i) U is irreducible and, (ii) there exists in $\mathscr{H}(U)$ at least one nonzero admissible vector.

Remarks: (a) Any representation unitarily equivalent to a square-integrable representation is also square-integrable.

(b) If G is compact, any irreducible representation U of G is square integrable. We shall see below examples of square-integrable representations of groups that are not compact.

(c) As an example of an irreducible representation that is *not* square-integrable, consider the one-dimensional representation $x \to e^{iax}$ of $\mathbb{R}$.

(d) If G is unimodular and if U is a square-integrable representation of G, then every vector in $\mathscr{H}(U)$ is admissible. (See, e.g., Ref. 3.) We shall see that the situation is different if G is not unimodular.

III. ORTHOGONALITY RELATIONS

A. Historical comments

Orthogonality relations were derived by Schur (beginning of this century) for finite groups, Weyl (in the 1920's) for compact groups, and Bargmann and Godement for (square-integrable representations of) unimodular groups in the 1950's; nonunimodular groups have been investigated more recently.[12,13]

Orthogonality relations in the unimodular case are expressed by the following equality [3,4]:

$$\int_G \overline{(U(x)g_1,f_1)}(U(x)g_2,f_2)dx = \lambda\,(g_2,g_1)(f_1,f_2)$$

for every g_1, g_2, f_1, f_2 in $\mathscr{H}(U)$, where λ depends only on the square-integrable representation U.

B. Statement of orthogonality relations

We have the following[12,13] theorem.

Theorem 3.1: Let U be a square integrable representation of G, acting on the Hilbert space $\mathscr{H}(U)$. Then there exists in $\mathscr{H}(U)$ a unique self-adjoint positive operator C such that the following hold.

(i) The set of admissible vectors coincides with the domain of C.

(ii) Let g_1 and g_2 be any two admissible vectors. Let f_1 and f_2 be any two vectors in $\mathscr{H}(U)$. Then

$$\int_G \overline{(U(x)g_1,f_1)}(U(x)g_2,f_2)dx = (Cg_2,Cg_1)(f_1,f_2). \tag{3.1}$$

(iii)If the group G is unimodular, then C is a multiple of the identity.

The proof in the general case uses an extension of the Schur lemma and is given in the Appendix.

C. A special case

If $g_1 = g_2 = f_1 = f_2$, then (3.1) gives that for any admissible vector g, one has

$$(Cg,Cg) = \frac{1}{\|g\|^2}\int |(U(x)g,g)|^2\,dx, \tag{3.2}$$

where $\|\cdot\|$ denotes the norm in $\mathscr{H}(U)$.

If $g_1 = g_2 = g$, one has

$$\int \overline{(U(x)g,f_1)}(U(x)g,f_2)dx = \frac{\int |(U(x)g,g)|^2\,dx}{\|g\|^2}(f_1,f_2). \tag{3.3}$$

IV. L_g TRANSFORM AND R_g TRANSFORM

In this section, we define the transforms described in the Introduction and prove their isometry.

A. Definitions

Let U be a square-integrable representation of G acting on $\mathscr{H}(U)$, and let g be a nonzero admissible vector [see (2.10)].

Associate to g the positive number

$$c_g = \frac{1}{\|g\|^2}\int |(U(x)g,g)|^2\,dx, \tag{4.1}$$

which, by Sec. II A, is also

$$c_g = \frac{1}{\|g\|^2}\int |(U(x)g,g)|^2\,d_R x.$$

Notice that, by the results of Sec. II A one has

$$c_{U(x_0)g} = \Delta\,(x_0)^{-1}c_g, \tag{4.2}$$

where $\Delta\,(x_0)$ is the modular function.

For any $f\in\mathscr{H}(U)$ consider the complex-valued functions $L_g f$ and $R_g f$ on G, defined by

$$(L_g f)(x) = (1/\sqrt{c_g})(U(x)g,f) \quad (x\in G), \tag{4.3}$$

$$(R_g f)(x) = (1/\sqrt{c_g})(g,U(x)f) \quad (x\in G). \tag{4.4}$$

One has

$$(R_g f)(x) = (L_g f)(x^{-1}). \tag{4.5}$$

The function $(L_g f)(x)$ will be called the L_g *transform* of f. It depends on the representation U and on the choice of the admissible vector g. For reasons which will become clear, we shall sometimes call g the *analyzing wavelet*.

Similarly, $(R_g f)(x)$ will be called the R_g transform of f.

Remark: By (4.2) and (4.3) we have

$$(L_{U(x_0)g}f)(x) = \Delta\,(x_0)^{1/2}(L_g f)(xx_0). \tag{4.6}$$

By the same argument we have also

$$(R_{U(x_0)g}f)(x) = \Delta\,(x_0)^{1/2}(R_g f)(x_0^{-1}x). \tag{4.7}$$

B. Continuity and boundedness

By the continuity of $U(x)$ and continuity of scalar product, the functions $(L_g f)(x)$ and $(R_g f)(x)$ are continuous on G.

By the Schwarz inequality, $L_g f$ and $R_g f$ are bounded on G:

$$|(L_g f)(x)| \leqslant (1/\sqrt{c_g})\|f\|\,\|g\|, \tag{4.8}$$

$$|(R_g f)(x)| \leqslant (1/\sqrt{c_g})\|f\|\,\|g\|, \tag{4.9}$$

for every $x\in G$.

C. Intertwining (covariance)

By the definition of L_g, we have, for every $f\in\mathscr{H}(U)$, $a\in G$, $x\in G$,

$$(L_g f)(a^{-1}x) = (U(a^{-1}x)g,f) = (U(x)g,U(a)f),$$

which can be written as

$$\lambda\,(a)L_g = L_g U(a). \tag{4.10}$$

Similarly we have

$$(R_g f)(xa) = (g,U(x)U(a)f),$$

giving

$$\rho(a)R_g = R_g U(a). \tag{4.11}$$

D. Isometry of L_g and of R_g

We have the following proposition.

Proposition 4.1: (i) The correspondence $f\to L_g f$ is isometric from $\mathscr{H}(U)$ into $L^2(G,dx)$; that is, for every $f_1\in\mathscr{H}(U)$, $f_2\in\mathscr{H}(U)$, we have

$$\int \overline{(L_g f_1)(x)}(L_g f_2)(x)dx = (f_1,f_2). \tag{4.12}$$

(ii) The correspondence $f\to R_g f$ is isometric from $\mathscr{H}(U)$ into $L^2(G,d_R x)$: for every $f_1\in\mathscr{H}(U)$, $f_2\in\mathscr{H}(U)$, we have

$$\int \overline{(R_g f_1)(x)}\,(R_g f_2)(x)d_R x = (f_1,f_2). \tag{4.13}$$

Proof: This proposition is a corollary of Theorem 3.1; by (3.3), we have

$$(f_1,f_2)=\frac{\|g\|^2}{\int|(U(x')g,g)|^2\,dx'}\int\overline{(U(x)g,f_1)}\,(U(x)g,f_2)dx$$
$$=\frac{1}{c_g}\int\overline{(U(x)g,f_1)}\,(U(x)g,f_2)dx$$
$$=\int\overline{(L_gf_1)(x)}\,(L_gf_2)(x)dx.$$

The change of variable $x\to x^{-1}$, $d(x^{-1})=d_R(x)$ gives

$$(f_1,f_2)=\frac{1}{c_g}\int\overline{(U(x^{-1})g,f_1)}\,(U(x^{-1})g,f_2)d(x^{-1})$$
$$=\frac{1}{c_g}\int\overline{(g,U(x)f_1)}(g,U(x)f_2)d_R(x)$$
$$=\int\overline{(R_gf_1)(x)}\,(R_gf_2)(x)d_R(x),$$

and the proposition is proved.

Remark: Since L_g is isometric from $\mathscr{H}(U)$ into $L^2(G,dx)$, L_g is unitary from $\mathscr{H}(U)$ to $L_g\mathscr{H}(U)\subset L^2(G,dx)$. By (4.10) we see that $L_g\mathscr{H}(U)$ is invariant under the left regular representation. So L_g is an intertwining operator between U and the restriction of the left regular representation of G.

This construction allows us to consider the representation U as a subrepresentation of the left regular representation of G. The following section will give a characterization of the range of L_g.

The same remark is also valid for R_g.

V. CHARACTERIZATION OF THE RANGES OF R_g and L_g

Let g be an admissible vector for U. Consider, on G, the complex-valued function $p_g(x)$, defined by

$$p_g(x)=(1/c_g)(U(x)g,g)=(1/\sqrt{c_g})(L_gg)(x)$$
$$=(1/\sqrt{c_g})(R_gg)(x^{-1}). \tag{5.1}$$

The function p_g satisfies $\bar{p}_g(x)=p_g(x^{-1})$ and belongs, by Sec. II A, to $L^2(G;dx)\cap L^2(G,d_Rx)$.

Proposition 5.1: (i) Let Φ belong to $L^2(G,dx)$. Then Φ belongs to $L_g\mathscr{H}(U)\subset L^2(G,dx)$ if and only if the equation

$$\Phi(x)=\int p_g(y^{-1}x)\Phi(y)dy \tag{5.2}$$

holds for every $x\in G$.

(ii) Let Ψ belong to $L^2(G,d_Rx)$. Then Ψ belongs to $R_g\mathscr{H}(U)\subset L^2(G,d_Rx)$ if and only if the equation

$$\Psi(x)=\int p_g(yx^{-1})\Psi(y)d_R(y) \tag{5.3}$$

holds for every $x\in G$.

Proof: Notice first that the integrals (5.2) and (5.3) converge for every x, since the integrand is the product of two square-integrable functions.

(i) Suppose Φ belongs to $L_g\mathscr{H}(U)$. This means that $\Phi(y)=(L_gf)(y)$ for some $f\in\mathscr{H}(U)$. By the definition of p_g [(5.1)] and of L_g [(4.3)], we have

$$\int p_g(y^{-1}x)\Phi(y)dy=\frac{1}{\sqrt{c_g}}\int(L_gg)(y^{-1}x)\Phi(y)dy$$
$$=\frac{1}{\sqrt{c_g}}\int\overline{L_gg(x^{-1}y)}\Phi(y)dy$$
$$=\frac{1}{\sqrt{c_g}}\int\overline{(\lambda(x)L_g)(y)}\Phi(y)dy.$$

From (4.10) we obtain

$$\int p_g(y^{-1}x)\Phi(y)dy=\frac{1}{\sqrt{c_g}}\int\overline{L_g(U(x)g)(y)}\Phi(y)dy$$
$$=(1/\sqrt{c_g})(U(x)g,f)$$
$$=(L_gf)(x)=\Phi(x)$$

by isometry and definition of L_g.

An analogous proof holds for R_g.

We must now prove the converse part of the proposition. To do this we must have an explicit expression of the inverse of L_g.

Lemma 5.2: For every $\Phi\in L^2(G,dx)$ the expression

$$\varphi=\frac{1}{\sqrt{c_g}}\int_G\Phi(x)U(x)g\,dx \tag{5.4}$$

defines a vector in $\mathscr{H}(U)$.

Proof: The integral (5.4) is weakly convergent. Indeed, for any $\psi\in\mathscr{H}(U)$, the function $x\to(\psi,U(x)g)$ is in $L^2(G,dx)$ [it is up to a constant $\overline{(L_g\Psi)(x)}$]. Since $\Phi(x)\in L^2(G,dx)$, the integral

$$\frac{1}{\sqrt{c_g}}\int(\psi,U(x)g)\Phi(x)dx=\int\overline{(L_g\psi)(x)}\Phi(x)dx$$

exists. Furthermore, by the Schwarz inequality in $L^2(G,dx)$ and isometry of L_g we have

$$\left|\frac{1}{\sqrt{c_g}}\int(\Psi,U(x)g)\Phi(x)dx\right|\leqslant\|\Phi\|_{L^2(G,dx)}\|\psi\|_{\mathscr{H}(U)}.$$

Then, by the Riesz theorem, the integral $(1/\sqrt{c_g})\times\int_G\Phi(x)U(x)g\,dx$ defines a vector in $\mathscr{H}(U)$.

Lemma 5.3: If Φ satisfies (5.2) then

$$L_g\left(\frac{1}{\sqrt{c_g}}\int\Phi(x)U(x)g\,dx\right)(y)=\Phi(y). \tag{5.5}$$

Proof: The computation of the left-hand side of (5.5) gives

$$\frac{1}{\sqrt{c_g}\sqrt{c_g}}\int\Phi(x)(U(y)g,U(x)g)dx$$
$$=\frac{1}{c_g}\int\Phi(x)(U(x^{-1}y)g,g)dx$$
$$=\int\Phi(x)p_g(x^{-1}y)dx=\Phi(y)\ \text{by hypothesis.}$$

Lemma 5.3 implies that if Φ satisfies (5.2), then Φ belongs to $L_g\mathscr{H}(U)$.

The same proof holds for R_g.

During the proof of the proposition we have proved the following.

Proposition 5.4: The inverses of L_g and R_g, on their respective domains $L_g\mathscr{H}(U)$ and $R_g\mathscr{H}(U)$ are given by

$$L_g^{-1}\Phi = \frac{1}{\sqrt{c_g}}\int \Phi(x)U(x)g\,dx, \tag{5.6}$$

$$R_g^{-1}\psi = \frac{1}{\sqrt{c_g}}\int \psi(x)U(x^{-1})g\,d_R x, \tag{5.7}$$

where the integrals are taken in the weak sense.

Remarks: (1) The formulas (5.6) and (5.7) express the fact that, on the range of an isometric operator, the adjoint coincides with the inverse.

(2) The condition (5.2) can be rephrased by saying that the space $L_g\mathscr{H}(U)\subset L^2(G,dx)$ is a Hilbert space with reproducing kernel[15,16] (functional Hilbert space in the terminology of Ref. 16). The evaluation functional in $L_g\mathscr{H}(U)$ is

$$\Phi\to\Phi(x) = (e_x,\Phi)_{L^2(G,dx)}, \tag{5.8}$$

with

$$e_x(y) = \bar{p}_g(y^{-1}x) = p_g(x^{-1}y). \tag{5.9}$$

Similarly, (5.3) says that $R_g\mathscr{H}(U)\subset L^2(G,d_R x)$ is a Hilbert space with reproducing kernel in which the evaluation functional is

$$\Psi\to\Psi(x) = (h_x,\Psi)_{L^2(G,d_R x)}, \tag{5.10}$$

with

$$h_x(y) = \bar{p}_g(yx^{-1}) = p_g(xy^{-1}). \tag{5.11}$$

(3) In terms of the dyadic notation used in the Introduction [(1.2)], the isometry property (4.12) can be written as

$$\int \frac{dx}{c_g}|x><x| = \mathbf{1}, \tag{5.12}$$

where the integral is taken in the weak sense.

Indeed, from (5.12) we get

$$\int \frac{<f_1|x>}{\sqrt{c_g}}\frac{<x|f_2>}{\sqrt{c_g}}\,dx = (f_1,f_2),$$

which is (4.12).

The reproducing property (5.2) can be also deduced from (5.12), since (5.12) implies

$$\int \frac{<y|x>}{c_g}\frac{<x|f>}{\sqrt{c_g}}\,dx = \frac{<y|f>}{\sqrt{c_g}},$$

which is (5.2).

The same remark holds for R_g.

VI. COVARIANT INTERPOLATION

A. Interpolation for the left transform

Proposition: Let $x_1,\ldots,x_n$ be n points in G. For $1\leqslant i,j\leqslant n$ consider the number

$$M_{ij} = (U(x_j)g,U(x_i)g) = (U(x_i^{-1}x_j)g,g)$$
$$= c_g p_g(x_i^{-1}x_j). \tag{6.1}$$

Let M be the $n\times n$ positive definite Hermitian matrix with entries M_{ij}. Assume that

$$\det M \neq 0. \tag{6.2}$$

For $1\leqslant j\leqslant n$, define a function $\Phi_j(x)$ on G, by

$$\Phi_j(x) = (U(x_j^{-1}x)g,g) = c_g p_g(x_j^{-1}x). \tag{6.3}$$

Let $\zeta_1,\ldots,\zeta_n$ be any n complex numbers.

Define on G the function $\Phi(x) = \Phi(x_1,\ldots,x_n;\zeta_1,\ldots,\zeta_n;x)$ as

$$\Phi(x) = -\frac{1}{\det M}\times\det\begin{vmatrix} 0 & \Phi_1(x) & \cdots & \Phi_n(x)\\ \zeta_1 & M_{11} & \cdots & M_{1n}\\ \vdots & & & \vdots\\ \zeta_n & M_{n1} & \cdots & M_{nn}\end{vmatrix}.$$

Then we have the following.

(i) $\Phi(x)$ belongs to the range of L_g,

$\Phi\in L_g\mathscr{H}(U)$.

(ii) $\Phi(x)$ satisfies

$$\Phi(x_i) = \zeta_i \quad (i = 1,\ldots,n),$$

i.e., is a solution of the interpolation problem.

(iii) $\Phi(x)$ is of minimal norm, in the following sense: If Φ^{other} is any other function on G satisfying (i) and (ii), then

$$\|\Phi^{\text{other}}\| \geqslant \|\Phi\|,$$

the norm being taken in $L^2(G;dx)$.

(iv) The interpolation procedure is invariant under left multiplication in G, in the following sense: Let a be any element of G. Then

$$M(ax_1,\ldots,ax_n) = M(x_1,\ldots,x_n) \tag{6.4}$$

and

$$\Phi(ax_1,\ldots,ax_n;\zeta_1,\ldots,\zeta_n;ax) = \Phi(x_1,\ldots,x_n;\zeta_1,\ldots,\zeta_n;x), \tag{6.5}$$

so that the left-displaced interpolation problem

$$\Phi_1(ax_i) = \zeta_i \tag{6.6}$$

is solved by the function

$$\Phi_1(x) = \Phi(a^{-1}x). \tag{6.7}$$

B. Interpolation for the right transform

Proposition: Let $x_1,\ldots,x_n$ be n points in G and $\zeta_1,\ldots,\zeta_n$ n complex numbers. Consider the $n\times n$ Hermitian matrix $N = N(x_1,\ldots,x_n)$,

$$N_{ij} = c_g p_g(x_i x_j^{-1}) = (U(x_j^{-1})g,U(x_i^{-1})g).$$

Assume that N_{ij} is invertible, and define on G the function $\Psi(x) = \Psi(x_1,\ldots,x_n;\zeta_1,\ldots,\zeta_n;x)$ as

$$\psi(x) = -\frac{1}{\det N}\times\det\begin{vmatrix} 0 & \Psi_1(x) & \cdots & \Psi_n(x)\\ \zeta_1 & & & \\ \vdots & & N & \\ \zeta_n & & & \end{vmatrix}, \tag{6.8}$$

where $\Psi_j(x) = (U(xx_j^{-1})g,g)$ $(j = 1,\ldots,n)$.

Then, (i) Ψ belongs to $R_g\mathscr{H}(U)\subset L^2(G,d_R x)$; (ii) Ψ takes the prescribed values $\zeta_1,\ldots,\zeta_n$ at the prescribed points $x_1,\ldots,x_n$,

$$\Psi(x_j) = \zeta_j \quad (j = 1,\ldots,n); \tag{6.9}$$

(iii) Ψ is of minimal norm, subject to (i) and (ii); and (iv) for any $a\in G$, the function

$$\Psi_1(x) = \Psi(xa^{-1}) \tag{6.10}$$

takes the values $\zeta_1,...,\zeta_n$ at the points $x_1a,...,x_na$,

$$\Psi_1(x_ja) = \zeta_j \quad (j = 1,...,n). \tag{6.11}$$

An alternative way of writing the interpolating functions $\Phi(x)$ and $\Psi(x)$ is

$$\Phi(x) = \sum_{i=1}^{n}\sum_{j=1}^{n}\zeta_i(M^{-1})_{ij}p_g(x_j^{-1}x), \tag{6.12}$$

$$\Psi(x) = \sum_{i=1}^{n}\sum_{j=1}^{n}\zeta_i(N^{-1})_{ij}p_g(xx_j^{-1}), \tag{6.13}$$

where M^{-1} is the matrix inverse of M, and N^{-1} is the matrix inverse of N.

The interpolations are typical of Hilbert spaces with reproducing kernels and the proof can be adapted from Meschkowsky.[15]

APPENDIX: PROOF OF THEOREM 3.1

In the proof of Theorem 3.1, we shall use the following extension of Schur's lemma.

Proposition A.1: Suppose that (i) G is a group; (ii) U is a unitary irreducible representation of G in a Hilbert space $\mathscr{H}$; (iii) π is a unitary (not necessarily irreducible) representation of G in a Hilbert space $\mathscr{H}'$; (iv) T is a closed operator from $\mathscr{H}$ to $\mathscr{H}'$ with domain $\mathscr{D} \subset \mathscr{H}$ dense in $\mathscr{H}$ and stable under U; and (v) $TU(x) = \pi(x)T$ on $\mathscr{D}$ for every x in G; then T is a multiple of an isometry, and $\mathscr{D} = \mathscr{H}$.

Proof: Let us denote by $(\cdot,\cdot)$ and $(\cdot,\cdot)'$ the scalar products in $\mathscr{H}$ and $\mathscr{H}'$, and by $\|\cdot\|$ and $\|\cdot\|'$ the associated norms.

Consider on $\mathscr{D}$ the scalar product

$$(g,f)_T = (g,f) + (Tg,Tf)'$$

and the associated norm

$$\|g\|_T^2 = \|g\|^2 + \|Tg\|'^2.$$

Then $\mathscr{D}$, equipped with the scalar product $(\cdot,\cdot)_T$, is a Hilbert space which we call $\mathscr{D}_T$.

Since

$$\frac{(\|Tg\|')^2}{\|g\|_T^2} = \frac{(\|Tg\|')^2}{(\|Tg\|')^2 + \|g\|^2} \leqslant 1,$$

T is bounded from $\mathscr{D}_T$ to $\mathscr{H}'$.

Moreover $U(x)$ is unitary in $\mathscr{D}_T$

$$\begin{aligned}\|U(x)g\|_T^2 &= \|U(x)g\|^2 + (\|TU(x)g\|')^2 \\ &= \|g\|^2 + (\|\pi(x)Tg\|')^2 \\ &= \|g\|^2 + (\|Tg\|')^2 = \|g\|_T^2,\end{aligned}$$

for every x in G and g in $\mathscr{D}$, and $U(x)|_{\mathscr{D}}$ is surjective, since for every $g \in \mathscr{D}$, $g = U(x)U(x^{-1})g$, and $\mathscr{D}$ is stable under $U(x^{-1})$.

By the usual Schur's lemma[17] we have that T is a multiple of an isometry from $\mathscr{D}_T$ to $\mathscr{H}'$, that is

$$(\|Tg\|')^2 = \lambda\|g\|_T^2,$$

for every g in $\mathscr{D}$, which gives

$$(\|Tg\|')^2 = \lambda\|g\|^2 + \lambda(\|Tg\|')^2.$$

From this equality we see that $\lambda \neq 1$ and that

$$(\|Tg\|')^2 = [\lambda/(1-\lambda)]\|g\|^2,$$

so T is a multiple of an isometry from $\mathscr{D}$ to $\mathscr{H}'$ and consequently extends to a multiple of an isometry from $\mathscr{H}$ to $\mathscr{H}'$. Since T was assumed closed, one has $\mathscr{D} = \mathscr{H}$, and the proposition is proved.

If we take $\mathscr{H}' = \mathscr{H}$ and $\pi = U$, the same argument together with the classical lemma of Schur shows that T is a multiple of the identity.

In order to prove Theorem 3.1, we first compute the integral (3.1).

For g admissible, consider the following operator T_g from $\mathscr{H}(U)$ to $L^2(G,dx)$: the domain $\mathscr{D}$ of T_g is the set of vectors f in $\mathscr{H}(U)$ such that

$$\int |(U(x)g,f)|^2 dx < +\infty;$$

for f in $\mathscr{D}$, T_gf is defined by

$$(T_gf)(x) = (U(x)g,f).$$

For f in $\mathscr{D}$ and y in G we have

$$\begin{aligned}\int |(U(x)g,U(y)f)|^2\,dx &= \int |(U(y^{-1}x)g,f)|^2\,dx \\ &= \int |(U(x)g,f)|^2\,dx < +\infty,\end{aligned}$$

by the left invariance of dx. So $U(y)f$ belongs to $\mathscr{D}$ for every y in G, i.e., $\mathscr{D}$ is stable under U and

$$T_gU(y) = L(y)T_g \quad \text{on } \mathscr{D}.$$

We see that $\mathscr{D}$ contains the linear span of the set of vectors $U(x)g$, $x \in G$. This linear span is dense in $\mathscr{H}(U)$ by irreducibility of U, so $\mathscr{D}$ is dense in $\mathscr{H}(U)$.

We prove now that T_g is closed: Take a sequence $\{f_n\}$, with $f_n \in \mathscr{D}$ for every n, converging to f in $\mathscr{H}(U)$, and such that T_gf_n converges in $L^2(G,dx)$ to $\varphi \in L^2(G,dx)$. Then T_gf_n converges to φ weakly in $L^2(G,dx)$ and the sequence of $L^2(G,dx)$ norms $\|T_gf_n\|$ is bounded.

By the continuity of the scalar product in $\mathscr{H}(U)$, the sequence of numbers $(U(x)g,f_n)$ converges to $(U(x)g,f)$ for every $x \in G$. Then we have (see Ref. 18, p. 207) that

$$(U(x)g,f) = \varphi(x)$$

so

$$\int |(U(x)g,f)|^2\,dx = \int |\varphi(x)|^2 dx < +\infty,$$

which implies that f belongs to $\mathscr{D}$ and $T_gf = \varphi$. So T_g is closed.

By the extended Schur's lemma (A.1), T_g is a multiple of an isometry: so $\mathscr{D} = \mathscr{H}(U)$ and T_g is bounded.

Now take g_1 and g_2 admissible; then T_{g1} and T_{g2} are bounded and $T_{g_1}^* T_{g_2}$ is a bounded operator in $\mathscr{H}(U)$. Since

$$L(x)T_{g_i} = T_{g_i}U(x) \quad (i = 1,2),$$

for every x in G, we have

$$U(x)T_{g_1}^* T_{g_2} = T_{g_1}^* T_{g_2} U(x),$$

for every x in G.

So, by Schur's lemma, $T_{g_1}^* T_{g_2}$ is a multiple of the identity:

$$T_{g_i}^* T_{g_2} = C_{g_1g_2}\mathbf{1}.$$

This means that, for every f_1, f_2 in $\mathscr{H}$,

$$\int \overline{(U(x)g_1, f_1)}(U(x)g_2, f_2)dx = (T_{g_1} f_1, T_{g_2} f_2)_{L^2(G,dx)}$$
$$= (f_1, T^*_{g_1} T_{g_2} f_2)_{\mathscr{H}(U)}$$
$$= C_{g_1 g_2}(f_1, f_2). \tag{A1}$$

Let us now consider the number $C_{g_1 g_2}$ defined by

$$C_{g_1 g_2} = \frac{\int \overline{(U(x)g_1 f)}(U(x)g_2 f)dx}{\|f\|^2}. \tag{A2}$$

Here, $C_{g_1 g_2}$ is, by (A1), independent of $f \neq 0$.

Let us denote by $\mathscr{A}$ the set of admissible vectors.

The correspondence q: $\mathscr{A} \times \mathscr{A} \to \mathbb{C}$ defined by $q(g_1, g_2) = C_{g_1 g_2}$ for g_1, g_2 in $\mathscr{A}$ is by (A2) a positive, symmetric quadratic form with form domain $\mathscr{A}$. Moreover q is closed: indeed consider on $\mathscr{A}$ the norm $\| \cdot \|_q$ defined by

$$\|g\|_q^2 = \|g\|^2 + q(g,g).$$

Take a $\| \cdot \|_q$-Cauchy sequence of vectors g_n in $\mathscr{A}$. This implies that (i) $\{g_n\}$ is a Cauchy sequence with respect to the $\mathscr{H}(U)$ norm, [so $\{g_n\}$ converges to g in $\mathscr{H}(U)$], and (ii) that

$$\lim_{n,m \to \infty} q(g_n - g_m, g_n - g_m) = 0,$$

which implies that the sequence of functions φ_n of $L^2(G,dx)$ defined by $\varphi_n(x) = (U(x)g_n, f)$ is a Cauchy sequence in $L^2(G,dx)$ and so converges strongly to $\varphi \in L^2(G,dx)$. Consequently φ_n converges weakly to φ and the sequence of norms $\|\varphi_n\|$ is bounded.

Moreover the sequence $(U(x)g_n, f)$ converges for each x in G to $(U(x)g, f)$. Then (see Ref. 18, p. 207), $(U(x)g, f) = \varphi(x)$ which means that, for every f in $\mathscr{H}$, $\int |(U(x)g, f)|^2 dx < +\infty$. In particular $\int |(U(x)g, g)|^2 dx < +\infty$ and so g is admissible. Furthermore

$$\lim_{n \to \infty} \|g_n - g\|_q^2 = \lim_{n \to \infty} \|g_n - g\|^2 + \lim_{n \to \infty} q(g_n - g, g_n - g)$$
$$= 0 + \lim_{n \to \infty} \frac{\|\varphi_n - \varphi\|^2_{L^2(G,dx)}}{\|f\|^2} = 0,$$

which shows that g_n converges in $\| \cdot \|_q$ norm to $g \in \mathscr{A}$.

So $\mathscr{A}$ is complete and then,[19] q is closed. Being as q is a densely defined closed symmetric positive form, by the second representation theorem,[20] there exists a unique positive operator C with domain $\mathscr{A}$ such that

$$q(g_1, g_2) = C_{g_1 g_2} = (Cg_1, Cg_2).$$

This proves parts (i) and (ii) of Theorem 3.1.

Suppose now G is unimodular; then we can see that

$$q(U(y)g_1, U(y)g_2) = \frac{1}{\|f\|^2} \int \overline{(U(xy)g_1, f)}(U(xy)g_2, f)dx$$
$$= \frac{1}{\|f\|^2} \int \overline{(U(x)g_1, f)}(U(x)g_2, f)d(xy^{-1})$$
$$= q(g_1, g_2).$$

This implies

$$(U(y)^{-1}CU(y)g_1, U(y)^{-1}CU(y)g_2) = (Cg_1, Cg_2),$$

and then $U(y)^{-1}CU(y) = C$ on $\mathscr{D}$.

By the remark at the end of the proof of Proposition A.1, it follows that C is a multiple of the identity since C is closed, $\mathscr{D}$ dense in $\mathscr{H}(U)$ and stable under U.

This proves point (iii) of Theorem 3.1.

In particular, if G is unimodular, then $\mathscr{A} = \mathscr{H}(U)$; that is, if one vector is admissible, then all vectors are admissible.

[1] J. R. Klauder and E. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968).
[2] J. R. Klauder and B. S. Skagerstam, *Coherent States. Applications in Physics and Mathematical Physics* (World Scientific, Singapore, 1984).
[3] G. Warner, *Harmonic Analysis on Semi-simple Lie Groups. I and II* (Springer, Berlin, 1972).
[4] S. A. Gaal, *Linear Analysis and Representation Theory* (Springer, Berlin, 1973).
[5] E. W. Aslaksen, and J. R. Klauder, J. Math. Phys. **10**, 2267 (1969).
[6] A. Grossmann and J. Morlet, SIAM J. Math. Anal. **15**, 723 (1984).
[7] A. Grossmann and J. Morlet, "Decomposition of functions into wavelets of constant shape and related transforms," to appear in *Mathematics + Physics, Lectures on Recent Results* (World Scientific, Singapore).
[8] T. Paul, J. Math. Phys. **25**, 3252 (1984).
[9] T. Paul, "Affine coherent states for the radial Schrödinger equation," in preparation.
[10] A. O. Barut and L. Girardello, Commun. Math. Phys. **21**, 41 (1972).
[11] A. M. Perelomov, Commun. Math. Phys. **26**, 222 (1972).
[12] M. Duflo and C. C. Moore, J. Funct. Anal. **21**, 209 (1976).
[13] A. L. Carey, Bull. Austral. Math. Soc. **15**, 1 (1976).
[14] E. Hewitt and K. A. Ross, *Abstract Harmonic Analysis. I.* (Springer, Berlin, 1963).
[15] H. Meschkowsky, *Hilbertsche Räume mit Kernfunktion* (Springer, Berlin, 1962).
[16] R. Young, *An Introduction to Non-harmonic Fourier Series* (Academic, New York, 1980).
[17] A. A. Kirillov, *Elements of the Theory of Representations* (Springer, Berlin, 1976).
[18] E. Hewitt and K. Stromberg, *Real and Abstract Analysis* (Springer, Berlin, 1969).
[19] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. I. Functional Analysis* (Academic, New York, 1980).
[20] T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1976).

On a computation formula for the representation matrices of U(n)

Takayoshi Maekawa
Department of Physics, Kumamoto University, Kumamoto 860, Japan

(Received 8 April 1985; accepted for publication 6 June 1985)

A d-matrix element of U(n), which is called the semi-highest-weight d-matrix element and plays an essential role for the D-matrix elements, is explicitly determined. By using the result, a formula of computing the D-matrix elements of U(n) is given in terms of lowering operators corresponding to those of Nagel and Moshinsky.

I. INTRODUCTION

In a previous paper,[1] the Haar measure of SU(n) [U(n)] is explicitly given in terms of the Euler-like parameters and then the differential operators of the first and the second parameter groups are computed. The D-matrix elements, together with the d-matrix elements, are introduced and their orthonormal and the completeness properties are given in a general form. It is, however, not given how to calculate the D-matrix elements or the d-matrix elements contrary to the case of SO(n),[2] although the method by means of the Young tableau is known.[3]

The irreducible representation matrix elements for the infinitesimal generators of U(n) have been known since Gel'fand and Tsetlin[4] and studied by many authors.[5] Though the explicit forms for the D-matrix elements are known for $n=3$ (Ref. 6) as well as $n=2$, and studied in the cases of any n (Ref. 7), those of U(n) seem not to be known in detail as for SO(n). On the other hand, for SO(n) a computation formula of the irreducible representation matrix elements for any n, in which lowering operators are used, is known.[2] Though the D-matrix elements of SO(n) are calculated by operating the lowering operators on the highest-weight D-matrix element which contains the highest-weight d-matrix element,[2] in order to obtain the D-matrix elements of U(n) it is necessary to know the highest-weight D-matrix element which contains a semi-highest-weight d-matrix element different from the highest-weight d-matrix element. It is, therefore, essential for the D-matrix elements of U(n) to determine the semi-highest-weight d-matrix element. Then the D-matrix elements will be found by operating the differential operators with respect to the Euler-like parameters on the highest-weight D-matrix element.

The pupose of this paper is to obtain the semi-highest-weight d-matrix element and then to give a computation formula to the D-matrix elements of U(n) through successive applications of the lowering operators to the highest-weight D-matrix element, as in the case of SO(n).[2]

In Sec. II, notations and some results needed in the following section are introduced, together with the definition of the semi-highest-weight d-matrix element. In Sec. III, the semi-highest-weight d-matrix element is obtained explicitly and then a computation formula of the D-matrix elements is given in a general form.

II. PRELIMINARIES

In this section, some notations and the results necessary in the next section are introduced, together with the definition of a semi-highest-weight d-matrix element, according to Ref. 1.

The generators e_{ij} of U(n) having the matrix elements $(e_{ij})_{kl}=\delta_{ik}\delta_{jl}$ satisfy the commutation relations

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

From these generators, we construct the generators of SU(n) as follows:

$$\begin{aligned}
&E_{ij}=\tfrac{1}{2}(e_{ij}+e_{ji})=E_{ji}, \quad \text{for} \quad i\neq j,\\
&E_{(ij)}=(i/2)(e_{ij}-e_{ji})=-E_{(ji)}, \quad \text{for} \quad i<j,\\
&E_{jj}=\frac{1}{\sqrt{2j(j+1)}}\left(\sum^{j}e_{kk}-je_{j+1\,j+1}\right),\\
&\quad \text{for} \quad j=1,2,\ldots,n-1.
\end{aligned} \tag{2.2}$$

The generators of U(n) are obtained by adding the unit generator $E=\Sigma e_{jj}$ to the above n^2-1 generators of SU(n) and this fact is, of course, due to the relation $\mathrm{U}(n)\sim\mathrm{U}(1)\times\mathrm{SU}(n)$. The n^2-1 generators E_{ij}'s may be arranged from 1 to n^2-1 as follows:

$$\begin{aligned}
&E_{(j-1)^2+2i-2}=E_{ij},\\
&E_{(j-1)^2+2i-1}=E_{(ij)}\\
&\quad (i=1,2,\ldots,j-1;\quad j=2,3,\ldots,n),\\
&E_{j^2-1}=E_{j-1\,j-1}\quad (j=2,3,\ldots,n).
\end{aligned} \tag{2.3}$$

The commutation relations for these E_j's may be written as follows:

$$[E_j,E_k]=i\sum f_{jkl}E_l, \tag{2.4}$$

where f_{ijk}'s are the structure constants which are totally antisymmetric with respect to their indices because of the normalization of $\mathrm{tr}\,(E_jE_k)=\delta_{jk}/2$.

The elements of SU(n) can be parametrized as follows:

$$\begin{aligned}
&g^{(n)}=g^{(n-1)}S^{(n)},\\
&S^{(n)}=\left(\prod_{j=n}^{2}e^{i\phi_{nn-j+1}\Gamma_{j-1}}e^{i\theta_{nn-j+1}E_{j^2-2}}\right)e^{i\psi_{nn-1}\Gamma_1}\\
&g^{(2)}=e^{i\phi_{21}\Gamma_1}e^{i\theta_{21}E_2}e^{i\psi_{21}\Gamma_1},
\end{aligned} \tag{2.5}$$

where

$$\begin{aligned}
\Gamma_j&=(e_{jj}-e_{j+1\,j+1})/2\\
&=(-\sqrt{j-1}E_{j^2-1}+\sqrt{j+1}E_{(j+1)^2-1})/\sqrt{2j}.
\end{aligned}$$

The parameters take the values

$$0\leqslant\theta_{jk}\leqslant\pi,\quad 0\leqslant\phi_{jk}<4\pi,\quad 0\leqslant\psi_{jj-1}<2\pi. \tag{2.6}$$

 0022-2488/85/102480-05$02.50

The Haar measure of SU(n) with respect to the above parameters is given by[1]

$$dV_n = \left[\prod_{i=1}^{n-1} \cos\frac{\theta_{ni}}{2}\left(\sin\frac{\theta_{ni}}{2}\right)^{2(n-i)-1} \times d\theta_{ni}\, d\phi_{ni}\right] d\psi_{nn-1}\, dV_{n-1},$$

$$dV_2 = \cos\frac{\theta_{21}}{2}\sin\frac{\theta_{21}}{2}\, d\theta_{21}\, d\phi_{21}\, d\psi_{21}. \tag{2.7}$$

The volume of SU(n) is given by $V_n = (4\pi)^n V_{n-1}/[2(n-1)!]$ and $V_2 = (4\pi)^2/2$. The measure of U(n) is given by multiplying (2.7) by $d\phi$ and the volume of U(n) is given by $2\pi V_n$.

The representation operators D_j corresponding to the generators E_j satisfy the same commutation relations as (2.4)

$$[D_j, D_k] = i\sum f_{jkl} D_l. \tag{2.8}$$

The representation D matrix of SU(n) corresponding to the parametrization (2.5) is given by

$$D^{(n)}(g^{(n)}) \equiv D^{(n-1)}(g^{(n-1)}) H(S^{(n)}),$$
$$H(S^{(n)}) = \left(\prod_{j=n}^{2} e^{i\phi_{nn-j+1}T_{j-1}} e^{i\theta_{nn-j+1}E_{j-2}}\right) e^{i\psi_{nn-1}T_1}, \tag{2.9}$$

where T_j is the representation operator corresponding to Γ_j.

The following relations are noticed[1]:

$$D_j D^{(n)}(g^{(n)}) = \overline{J}_j D^{(n)}(g^{(n)}),$$
$$D^{(n)}(g^{(n)}) D_j = J_j D^{(n)}(g^{(n)}), \tag{2.10}$$

where $\overline{J}_j$ and J_j are the differential operators of the first and the second parameter groups corresponding to D_j, whose explicit expressions are omitted here.[1] It is noted that the J_j's satisfy the same commutation relations as (2.8), but the $\overline{J}_j$'s satisfy the relations with a minus sign on the right side of (2.8).

The matrix elements of the unitary irreducible representations (UIR's) of U(n) are given by the Gel'fand and Tsetlin basis[4]

$$|m_{jk}\rangle = |\lambda_n,\ldots\lambda_1\rangle, \tag{2.11}$$

where λ_j denotes $(m_{1j},m_{2j},\ldots,m_{jj})$. The non-negative integers m_{jk} are subject to the conditions $m_{jk+1} \geqslant m_{jk} \geqslant m_{j+1k+1}$. The dimension of the UIR of U(n) is given as follows[8]:

$$N(\lambda_n) = D(l_1,\ldots,l_n)/D(n-1,\ldots,0),$$
$$D(l_1,\ldots,l_n) = \prod_{j<k}(l_j - l_k), \quad l_j = m_{jn} + n - j. \tag{2.12}$$

The action of D_j on the basis (2.11) is given by[4,5]

$$(D_{k^2-3} + iD_{k^2-2})|m_{jk}\rangle = \sum_{j=1}^{k-1} A^j_{k-1k}|m_{jk-1}+1\rangle,$$
$$(D_{k^2-3} - iD_{k^2-2})|m_{jk}\rangle = \sum_{j=1}^{k-1} B^j_{kk-1}|m_{jk-1}-1\rangle,$$
$$D(e_{kk})|m_{jk}\rangle = \left(\sum_{j=1}^{k} m_{jk} - \sum_{j=1}^{k-1} m_{jk-1}\right)|m_{jk}\rangle, \tag{2.13}$$

where the matrix elements A and B have the explicit forms

$$A^j_{k-1k} = \left[-\frac{\Pi_{i=1}^{k}(m_{ik}-m_{jk-1}-i+j)\Pi_{i=1}^{k-2}(m_{ik-2}-m_{jk-1}-i+j-1)}{\Pi_{i\neq j}(m_{ik-1}-m_{jk-1}-i+j-1)(m_{ik-1}-m_{jk-1}-i+j)}\right]^{1/2},$$
$$B^j_{kk-1} = \left[-\frac{\Pi_{i=1}^{k}(m_{ik}-m_{jk-1}-i+j+1)\Pi_{i=1}^{k-2}(m_{ik-2}-m_{jk-1}-i+j)}{\Pi_{i\neq j}(m_{ik-1}-m_{jk-1}-i+j+1)(m_{ik-1}-m_{jk-1}-i+j)}\right]^{1/2}.$$

The basis (2.11) may be obtained by operating the lowering operators on the highest-weight basis as follows[9]:

$$|m_{jk}\rangle = \mathscr{L}_{(n)}(D_j)|\lambda_n,\{\Lambda_{n-1}\}\rangle, \tag{2.14}$$

where $\{\Lambda_{n-1}\} = (\Lambda_{n-1},\ldots,\Lambda_1)$ denotes the highest possible values of the λ_j's; that is, $\Lambda_j = (m_{1n},\ldots,m_{jn})$ $(j = 1,2,\ldots,n-1)$, and $\mathscr{L}_{(n)}(D_j)$, whose explicit expression is omitted here,[9] depends on the representation operators D_j.

The representation D-matrix elements are defined by sandwiching (2.9) between the basis (2.11) as follows:

$$D^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \langle\lambda_n\{\lambda'_{n-1}\}|D^{(n)}(g^{(n)})|\lambda_n\{\lambda_{n-1}\}\rangle. \tag{2.15}$$

The d-matrix elements are defined by

$$d^{(\lambda_n)}_{\lambda'_{n-1}\{\lambda_{n-2}\}\lambda_{n-1}}(\theta) = \langle\lambda_n\lambda'_{n-1}\{\lambda_{n-2}\}|e^{i\theta D_{n^2-2}}|\lambda_n\{\lambda_{n-1}\}\rangle. \tag{2.16}$$

As is seen from (2.9) and the definition (2.15), the D-matrix elements are determined by the d-matrix elements (2.16) provided that those of SU($n-1$) [U($n-1$)] are known. Furthermore, it follows from (2.14) and (2.15) that the highest-weight D-matrix element is determined from the d-matrix element with a semi-highest weight which is defined by putting $\lambda'_{n-1} = \lambda_{n-1} = \Lambda_{n-1}$ in (2.16), i.e.,

$$d^{(\lambda_n)}_{\Lambda_{n-1}\{\lambda_{n-2}\}\Lambda_{n-1}}(\theta) = \langle\lambda_n\Lambda_{n-1}\{\lambda_{n-2}\}|e^{i\theta D_{n^2-2}}|\lambda_n\Lambda_{n-1}\{\lambda_{n-2}\}\rangle. \tag{2.17}$$

We can calculate the D-matrix elements by operating the lowering operators on the highest-weight D-matrix element if we can give an explicit expression to the semi-highest-weight d-matrix element (2.17). It is noted that the representation matrix elements of U(n) and SU(n) are obtained from each other by a simple replacement of the numbers m_{jk},[10] though the matrix elements (2.15) and (2.16) should be named those of U(n) because of the use of the U(n) basis (2.11).

Before obtaining the expression for (2.17), we give the

orthogonality relation of the D-matrix elements as follows[1,11]:

$$\int_{SU(n)} dV_n \, \overline{D^{(\lambda_n)}_{\{\lambda_{n-1}\}\{\lambda'''_{n-1}\}}(g^{(n)})} D^{(\lambda'_n)}_{\{\lambda'_{n-1}\}\{\lambda'''_{n-1}\}}(g^{(n)}) = \delta_{\{\lambda_n\}\{\lambda'_n\}} \delta_{\{\lambda'''_{n-1}\}\{\lambda''_{n-1}\}} [V_n/N(\lambda_n)], \tag{2.18}$$

where $\delta_{\{\lambda_n\}\{\lambda'_n\}}$ stands for a product of Kronecker deltas in each of the indices. Equation (2.18) leads to the orthogonality for the d-matrix elements as follows[1]:

$$\sum_{\lambda_{n-2}} N(\lambda_{n-2}) \int_0^\pi d\theta \cos\frac{\theta}{2}\left(\sin\frac{\theta}{2}\right)^{2n-3} \times d^{(\lambda_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta) d^{(\lambda'_n)}_{\lambda_{n-1}(\lambda_{n-2})\lambda'_{n-1}}(\theta) = \delta_{\lambda_n\lambda'_n} \frac{1}{n-1} \frac{N(\lambda_{n-1})N(\lambda'_{n-1})}{N(\lambda_n)}. \tag{2.19}$$

III. DETERMINATION OF THE SEMI-HIGHEST-WEIGHT d-MATRIX ELEMENT AND A COMPUTATION FORMULA OF THE D-MATRIX ELEMENTS

In this section, the semi-highest-weight d-matrix element defined by (2.17) is determined explicitly and then a computation formula of the D-matrix elements is given in terms of the lowering operators.

From (2.17) and (2.13), we obtain by considering $(\lambda_{n-1} \to \Lambda_{n-1})$

$$d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\Lambda_{n-1}}(\theta) = 1 - (\theta^2/2!)\langle \lambda_n \Lambda_{n-1}\{\lambda_{n-2}\}|(D_{n^2-2})^2 \times |\lambda_n \Lambda_{n-1}\{\lambda_{n-2}\}\rangle + \cdots. \tag{3.1}$$

The matrix element of the second term becomes from (2.13)

$$4\langle \lambda_n \Lambda_{n-1}\{\lambda_{n-2}\}|(D_{n^2-2})^2|\lambda_n \Lambda_{n-1}\{\lambda_{n-2}\}\rangle = \sum_{j=1}^{n-1} (B^j_{nn-1})^2|_{\lambda_{n-1}=\Lambda_{n-1}} = -\sum_{j=1}^{n-1} \frac{\Pi_{i=1}^n (m_{in} - m_{jn} - i + j + 1)\Pi_{i=1}^{n-2}(m_{in-2} - m_{jn} - i + j)}{\Pi_{i\neq j}^{n-1}(m_{in} - m_{jn} - i + j + 1)(m_{in} - m_{jn} - i + j)}. \tag{3.2}$$

The right side of (3.2) is easily calculated and given as follows (see Appendix):

$$4\langle \lambda_n \Lambda_{n-1}\{\lambda_{n-2}\}|(D_{n^2-2})^2|\lambda_n \Lambda_{n-1}\{\lambda_{n-2}\}\rangle = \sum_{i=1}^{n-1} m_{in} - \sum_{i=1}^{n-2} m_{in-2} - m_{nn}. \tag{3.3}$$

We can, therefore, deduce the following expression to the semi-highest-weight d-matrix element (2.17):

$$d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\Lambda_{n-1}}(\theta) = (\cos(\theta/2))^{\Sigma_{i=1}^{n-1} m_{in} - \Sigma_{i=1}^{n-2} m_{in-2} - m_{nn}}. \tag{3.4}$$

The validity of (3.4) can easily be seen by checking the integral (2.19), i.e.,

$$\sum_{\lambda_{n-2}} N(\lambda_{n-2}) \int_0^\pi d\theta \cos\frac{\theta}{2}\left(\sin\frac{\theta}{2}\right)^{2n-3} \times d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\Lambda_{n-1}}(\theta) d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\Lambda_{n-1}}(\theta) = [1/(n-1)][(N(\Lambda_{n-1}))^2/N(\lambda_n)], \tag{3.5}$$

where the summation over λ_{n-2} means the sums over m_{in-2} from m_{i+1n} to m_{in} $(i = 1,2,\ldots,n-2)$. By using the well-known formula

$$\int_0^{\pi/2} \sin^p\theta \cos^q\theta \, d\theta = \frac{1}{2} B\left(\frac{p+1}{2}, \frac{q+1}{2}\right), \quad \mathrm{Re}\, p, \mathrm{Re}\, q > -1, \quad B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q),$$

and (3.4), (3.5) gives the relation

$$\sum_{\lambda_{n-2}} N(\lambda_{n-2}) \times \frac{\Gamma(\Sigma_{i=1}^{n-1} m_{in} - \Sigma_{i=1}^{n-2} m_{in-2} - m_{nn} + 1)\Gamma(n-1)}{\Gamma(\Sigma_{i=1}^{n-1} m_{in} - \Sigma_{i=1}^{n-2} m_{in-2} - m_{nn} + n)} = [1/(n-1)][(N(\Lambda_{n-1}))^2/N(\lambda_n)], \tag{3.6}$$

whose validity is easily seen (see Appendix). It, therefore, follows that the expression (3.4) gives the semi-highest-weight d-matrix element.

A formula of computing the representation matrix elements can easily be given. It follows that the D-matrix element with the highest weight is given by

$$\langle \lambda_n\{\Lambda_{n-1}\}|D^{(n)}(g^{(n)})|\lambda_n\{\Lambda_{n-1}\}\rangle = \sum_{\{\lambda_{n-2}\}} \langle \{\Lambda_{n-1}\}|D^{(n-1)}(g^{(n-1)})|\Lambda_{n-1}\{\lambda_{n-2}\}\rangle \times e^{(1/2)i\phi_{n1}(\Sigma_{i=1}^{n-1} m_{in} - \Sigma_{i=1}^{n-2} m_{in-2} - m_{nn})} \times d^{(\lambda_n)}_{\Lambda_{n-1}(\lambda_{n-2})\Lambda_{n-1}}(\theta_{n1}) \times \langle \Lambda_{n-1}\{\lambda_{n-2}\}|H^{(n)}_{n-1}|\{\Lambda_{n-1}\}\rangle, \tag{3.7}$$

where

$$H^{(n)}_{n-1} = \left(\prod_{j=n-1}^{2} e^{i\phi_{nn-j+1}T_{j-1}} e^{i\theta_{nn-j+1}E_{f-2}}\right) e^{i\psi_{nn-1}T_1}.$$

As the matrix elements of $H^{(n)}_{n-1}$ are given by $\phi_{n-1j} \to \phi_{nj}$, $\theta_{n-1j} \to \theta_{nj}$, $\psi_{n-1n-2} \to \psi_{nn-1}$, and the other angles $\to 0$ in those of $D^{(n-1)}(g^{(n-1)})$, and the semi-highest-weight d matrix on (3.7) is given by (3.4), the matrix element with the highest weight is completely determined. Therefore, by taking into account the relations (2.10) and (2.14), the general D-

matrix elements are obtained by operating the lowering operators on the D-matrix element (3.7)

$$D^{(\lambda_n)}_{\{\lambda'_{n-1}\}\{\lambda_{n-1}\}}(g^{(n)}) = \mathscr{L}_{(n)}(\overline{J}_j)\mathscr{L}_{(n)}(J_j) \times \langle \lambda_n\{\Lambda_{n-1}\}|D^{(n)}(g^{(n)})|\lambda_n\{\Lambda_{n-1}\}\rangle, \tag{3.8}$$

where the normalized lowering operators $\mathscr{L}_{(n)}(J_j)$ and $\mathscr{L}_{(n)}(\overline{J}_j)$ are obtained, respectively, by replacing the generators D_j appearing in the lowering operators of Nagel and Moshinsky[9] by the corresponding differential operators J_j and $\overline{J}_j$, whose expressions are obtained by the formula given in Ref. 1.

Thus we conclude that the representation matrix elements of U(n) [SU(n)] are given by (3.8) through the action of the lowering operators on the D-matrix element (3.7) with the highest weight. The only task for us is to express J_j and $\overline{J}_j$ of (2.10) in terms of the differential operators and then to construct the lowering operators by these. The calculation of the formula for J_j and $\overline{J}_j$ is elementary but lengthy.[1]

APPENDIX: PROOF OF THE RELATIONS (3.3) AND (3.6)

In this appendix, it is shown that the sum (3.2) is given by the right side of (3.3) and the relation (3.6) holds.

The right side on (3.2) can be rewritten as follows:

$$\sum_{j=1}^{n-1}(m_{nn}-m_{jn}-n+j+1)\frac{\Pi_{i=1}^{n-2}(m_{in-2}-m_{jn}-i+j)}{\Pi_{i\neq j}^{n-1}(m_{in}-m_{jn}-i+j)}. \tag{A1}$$

By considering the jth and k th terms in (A1), there appears a denominator $m_{kn}-m_{jn}-k+j$ $(k>j)$ with different sign only in these terms. Taking into account the common denominator and putting $m_{kn}-m_{jn}-k+j=0$ on the numerators, we get the following form of the numerators:

$$\begin{aligned}&(m_{nn}-m_{jn}-n+j+1)\prod_{i=1}^{n-2}(m_{in-2}-m_{jn}-i+j)\\ &\times\prod_{l\neq j,k}^{n-1}(m_{ln}-m_{kn}-l+k)\\ &-(m_{nn}-m_{kn}-n+k+1)\\ &\times\prod_{i=1}^{n-2}(m_{in-2}-m_{kn}-i+k)\prod_{l\neq j,k}^{n-1}(m_{ln}-m_{jn}-l+j),\end{aligned} \tag{A2}$$

which becomes zero for $m_{kn}-m_{jn}+j-k=0$. This holds for each pair and it follows that the numerator contains the factor of the form

$$\prod_{i\neq j}^{n-1}(m_{in}-m_{jn}-i+j). \tag{A3}$$

The degree of the numerator of (A1) is $(n-1)$th and that of the denominator is $(n-2)$th. It, therefore, follows that (A1) is a linear form in m_{jn} $(j=1,\ldots,n)$ and m_{jn-2} $(j=1,2,\ldots,n-2)$ because of the cancellation of the denominator of (A1) by (A3). As (A1) becomes zero for $m_{jn}=0$ $(j=1,\ldots,n)$ and $m_{jn-2}=0$ $(j=1,\ldots,n-2)$, it must be linear and homogeneous, i.e.,

$$\sum_{i=1}^{n}a_i m_{in}+\sum_{i=1}^{n-2}b_i m_{in-2}. \tag{A4}$$

Equation (A1) becomes zero for $m_{jn}=m_{jn-2}$ $(j=1,\ldots,n-2)$ and $m_{n-1n}=m_{nn}$. From this, it follows that $b_i=-a_i$ $(i=1,\ldots,n-2)$ and $a_{n-1}=a_n$. Furthermore, by considering the case of $m_{jn-2}=m_{j+1n}$ $(j=1,\ldots,n-2)$, we get $m_{nn}-m_{1n}$ for (A1) and $a_1=a_{n-1}$. In this way, by considering special cases for m_{in} and m_{in-2}, we easily obtain

$$-\sum_{i=1}^{n-1}m_{in}+\sum_{i=1}^{n-2}m_{in-2}+m_{nn}, \tag{A5}$$

which gives (3.3).

Let us show that (3.6) is valid. Equation (3.6) together with (2.12) becomes

$$\sum_{m_{1n-2}=-m_{2n}}^{m_{1n}}\cdots\sum_{m_{n-2n-2}=m_{n-1n}}^{m_{n-2n}}\frac{\Pi_{j<k}^{n-2}(m_{jn-2}-m_{kn-2}-j+k)}{\Pi_{j=1}^{n-1}(\Sigma_{i=1}^{n-1}m_{in}-\Sigma_{i=1}^{n-2}m_{in-2}-m_{nn}+j)}=\frac{1}{(n-2)!}\frac{\Pi_{j<k}^{n-1}(m_{jn}-m_{kn}-j+k)}{\Pi_{j=1}^{n-1}(m_{jn}-m_{nn}-j+n)}, \tag{A6}$$

which may be rewritten as follows:

$$\sum_{p_1=\xi_2+1}^{\xi_1}\cdots\sum_{p_{n-2}=\xi_{n-1}+1}^{\xi_{n-2}}\frac{\Pi_{j<k}^{n-2}(p_j-p_k)}{\Pi_{j=1}^{n-1}(\Sigma_{i=1}^{n-1}\xi_i-\Sigma_{i=1}^{n-2}p_i+n+j-1)}=\frac{1}{(n-2)!}\frac{\Pi_{j<k}^{n-1}(\xi_j-\xi_k)}{\Pi_{j=1}^{n-1}(\xi_j+n)}, \tag{A7}$$

where

$$p_j=m_{jn-2}-m_{nn}-j\quad(j=1,\ldots,n-2),\quad \xi_j=m_{jn}-m_{nn}-j\quad(j=1,\ldots,n-1).$$

It can be proved that the following relation holds:

$$\sum_{p_1=\xi_2+1}^{\xi_1}\cdots\sum_{p_{n-2}=\xi_{n-1}+1}^{\xi_{n-2}}\frac{\Pi_{j<k}^{n-2}(p_j-p_k)}{\Pi_{j=1}^{n-1}(z-\Sigma_{i=1}^{n-2}p_i+n+j-1)}=\frac{1}{(n-2)!}\frac{\Pi_{j<k}^{n-1}(\xi_j-\xi_k)}{\Pi_{j=1}^{n-1}(z-\Sigma_{i=1}^{n-1}\xi_i+\xi_j+n)}, \tag{A8}$$

where z may be any possible number and (A8) gives (A7) in the case of $z=\Sigma_{i=1}^{n-1}\xi_i$. The numerator on the left side of (A8) can be represented in the determinant form

$$\prod_{j<k}^{n-2}(p_j-p_k)=\begin{vmatrix} p_1^{n-3}, & p_1^{n-4}, & \cdots, & p_1, & 1\\ p_2^{n-3}, & p_2^{n-4}, & \cdots, & p_2, & 1\\ \vdots & \vdots & & \vdots & \vdots\\ p_{n-2}^{n-3}, & p_{n-2}^{n-4}, & \cdots, & p_{n-2}, & 1\end{vmatrix}. \tag{A9}$$

Making use of the relations

$$B(\zeta,n-1)=\frac{(n-2)!}{\zeta(\zeta+1)\cdots(\zeta+n-2)}=\int_0^1 x^{\zeta-1}(1-x)^{n-2}\,dx,\quad \mathrm{Re}\,\zeta>0,$$

$$\sum_{p_i=\xi_{i+1}+1}^{\xi_i} p_i^l x^{-p_i}=\frac{\partial^l}{\partial t^l}\left[\frac{e^{\xi_i t}(1/x)^{\xi_i}-e^{\xi_{i+1}t}(1/x)^{\xi_{i+1}}}{1-xe^{-t}}\right]\Big|_{t=0}\equiv f_i^{(l)}(x)-f_{i+1}^{(l)}(x), \tag{A10}$$

where the first relation corresponds to a special case of that below (3.5) and

$$f_i^{(l)}(x)=\frac{\partial^l}{\partial t^l}\left[\frac{e^{\xi_i t}}{1-xe^{-t}}\right]\Big|_{t=0}\left(\frac{1}{x}\right)^{\xi_i}=\frac{\xi_i^l}{1-x}\left(\frac{1}{x}\right)^{\xi_i}+\sum_{s=1}^{l}\binom{l}{s}\xi_i^{l-s}\frac{\partial^s}{\partial t^s}\times(1-xe^{-t})^{-1}\Big|_{t=0}\left(\frac{1}{x}\right)^{\xi_i}, \tag{A11}$$

we rewrite the left side of (A8) as follows:

$$\frac{1}{(n-2)!}\int_0^1 dx\,x^{z+n-1}(1-x)^{n-2}\begin{vmatrix} f_1^{(n-3)}-f_2^{(n-3)}, & \cdots, & f_1^{(1)}-f_2^{(1)}, & f_1^{(0)}-f_2^{(0)} \\ f_2^{(n-3)}-f_3^{(n-3)}, & \cdots, & f_2^{(1)}-f_3^{(1)}, & f_2^{(0)}-f_3^{(0)} \\ \vdots & & \vdots & \vdots \\ f_{n-2}^{(n-3)}-f_{n-1}^{(n-3)}, & \cdots, & f_{n-2}^{(1)}-f_{n-1}^{(1)}, & f_{n-2}^{(0)}-f_{n-1}^{(0)} \end{vmatrix}$$

$$=\frac{1}{(n-2)!}\int_0^1 dx\,x^{z+n-1}(1-x)^{n-2}\begin{vmatrix} f_1^{(n-3)}, & f_1^{(n-4)}, & \cdots, & f_1^{(0)}, & 1 \\ f_2^{(n-3)}, & f_2^{(n-4)}, & \cdots, & f_2^{(0)}, & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ f_{n-1}^{(n-3)}, & f_{n-1}^{(n-4)}, & \cdots, & f_{n-1}^{(0)}, & 1 \end{vmatrix}, \tag{A12}$$

where the argument x of the f_i's is omitted.

Substituting the expressions (A11) for $f_i^{(l)}$ into (A12) and taking into account the fact that only the first term (A11) remains in each column due to the property of the determinant, we get after a simple procedure

$$\frac{1}{(n-2)!}\int_0^1 dx\,x^{z-\sum_{i=1}^{n-1}\xi_i+n-1}\times\begin{vmatrix} \xi_1^{n-3}, & \xi_1^{n-4}, & \cdots, & \xi_1, & 1, & x^{\xi_1} \\ \xi_2^{n-3}, & \xi_2^{n-4}, & \cdots, & \xi_2, & 1, & x^{\xi_2} \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \xi_{n-1}^{n-3}, & \xi_{n-1}^{n-4}, & \cdots, & \xi_{n-1}, & 1, & x^{\xi_{n-1}} \end{vmatrix}. \tag{A13}$$

After simple integrals and the use of fundamental properties of the determinant, (A13) gives

$$\frac{1}{(n-2)!}\begin{vmatrix} \xi_1^{n-2}, & \xi_1^{n-3}, & \cdots, & \xi_1, & 1 \\ \xi_2^{n-2}, & \xi_2^{n-3}, & \cdots, & \xi_2, & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \xi_{n-1}^{n-2}, & \xi_{n-1}^{n-3}, & \cdots, & \xi_{n-1}, & 1 \end{vmatrix}\times\frac{1}{\sum_{j=1}^{n-1}(z-\sum_{i=1}^{n-1}\xi_i+\xi_j+n)}, \tag{A14}$$

which agrees with (A8).

[1] T. Maekawa, J. Math. Phys. **26**, 1902 (1985).
[2] T. Maekawa, J. Math. Phys. **16**, 334 (1975); **19**, 2028 (1978); K. B. Wolf, J. Math. Phys. **12**, 197 (1971).
[3] H. Boerner, *Representations of Groups with Special Consideration for the Needs of Modern Physics* (North-Holland, Amsterdam, 1970).
[4] I. M. Gel'fand and M. L. Tsetlin, Dokl. Akad. Nauk SSSR **71**, 825 (1950).
[5] M. Moshinsky, J. Math. Phys. **4**, 1128 (1963); G. E. Baird and L. C. Biedenharn, *ibid.* **4**, 1449 (1963); J. G. Nagel and M. Moshinsky, *ibid.* **6**, 682 (1965).
[6] D. F. Holland, J. Math. Phys. **10**, 531 (1969).
[7] I. M. Gel'fand and M. I. Graev, Izv. Akad. Nauk SSSR Ser. Mat. **29**, 1329 (1965); D. F. Holland, J. Math. Phys. **10**, 1903 (1969); J. D. Louck, Am. J. Phys. **38**, 3 (1970).
[8] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1950).
[9] J. G. Nagel and M. Moshinsky, Ref. 5.
[10] G. E. Baird and L. C. Biedenharn, Ref. 5.
[11] N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations, Mathematical Monographs*, Vol. 22, translated by V. N. Singh (Am. Math. Soc., Providence, RI, 1968).

Satake diagrams, Iwasawa decompositions, and representations of the exceptional Lie group $F_4(-20)$

Veena Sharma and K. C. Tripathy
Department of Physics and Astrophysics, University of Delhi, Delhi-110 007, India

(Received 22 February 1984; accepted for publication 21 March 1985)

The parabolic subgroups of the noncompact, exceptional Lie group $F_4(-20)$ are computed from a systematic analysis of Iwasawa and Langlands decompositions. Satake diagrams have been used to determine the involutive automorphisms of $F_4(-20)$ which facilitate the Iwasawa decompositions. The polarizations associated with noncompact orbits with parabolic subgroups are computed. The representations so obtained for these polarizations using Kostant's induction scheme yield the principal series representations of $F_4(-20)$.

I. INTRODUCTION

Ever since the advent of noncompact groups, specially the SO(m,n) and SU(m,n) groups in particle physics, the study of unitary irreducible representations (UIR's) of such groups has generated brisk activities for mathematicians and physicists as well.[1] The exceptional Lie groups and their graded forms have only recently found much applicability when attempts were made to unify the fundamental interactions of particle theory with gravitation. However, very little attention has been paid to the study of noncompact exceptional Lie groups. In the present paper, we specifically focus on the exceptional Lie group $F_4(-20)$ motivated by its rich structure to accomodate many degrees of freedom (quantum numbers). The other noncompact exceptional Lie groups may be analyzed in an analogous fashion. Here, we discuss the construction of principal series representations of $F_4(-20)$. The exceptional Lie group F_4 with compact form $F_4(-52)$ has two associated noncompact forms: (1) $F_4(-20)$ with maximal compact subgroup SO(9), and (2) $F_4(+4)$ with maximal compact subgroup Sp(3) $\otimes$ SO(2). The plan of the paper is as follows.

In Sec. II, we give a brief resume of Iwasawa and Langlands decomposition to compute the parabolic subalgebras. Section III has two parts. In Sec. III A Satake diagrams[2] are defined and used to compute the inner automorphisms of $\mathcal{f}_4(-20)$, the Lie algebra of $F_4(-20)$. In Sec. III B, we discuss the Iwasawa decompositions of $\mathcal{f}_4(-20)$. In Sec. IV, parabolic subalgebras of $\mathcal{f}_4(-20)$ are computed. It is shown that $\mathcal{f}_4(-20)$ has two parabolic subalgebras, one being minimal and the other, the Lie algebra itself. The polarizations associated with noncompact orbits with parabolic subalgebras are computed. The representations obtained for these polarizations using Kostant's induced scheme[3,4] yield the principal series representation of $F_4(-20)$.

II. RESUME OF IWASAWA DECOMPOSITIONS AND LANGLANDS DECOMPOSITIONS[5-7]

Let $\mathcal{g}$ be a real Lie algebra and let $\mathcal{g}$ be generated by its compact real form $\mathcal{g}_k$ by an involutive automorphism σ defined with respect to the Cartan subalgebra $\mathcal{h}$ of $\mathcal{g}_c$, $\mathcal{g}_c$ being the complexification of $\mathcal{g}$. The following commutation relations are satisfied by the elements of $\mathcal{g}_c$:

$$[e_\alpha,h] = \alpha(h)e_\alpha, \quad h\in\mathcal{h}, \quad \alpha\in\Delta,$$

$$[e_\alpha,e_\beta] = \begin{cases} N_{\alpha\beta}e_{\alpha+\beta}, & \text{if } \alpha+\beta \text{ is a root,} \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

$$[e_\alpha,e_{-\alpha}] = h_\alpha, \quad h_\alpha\in\mathcal{h}.$$

Here, Δ denotes the set of roots of $\mathcal{g}_c$ with respect to $\mathcal{h}$ and the Killing form is defined as $B(e_\alpha,e_{-\alpha}) = -1$. The compact real form $\mathcal{g}_k$ may be taken to consist of

$$\{ih_\alpha, \text{ for } \alpha=\alpha_1,\alpha_2,\ldots,\alpha_l, \ l=\dim\mathcal{h};$$
$$(e_\alpha+e_{-\alpha}),\ (e_\alpha-e_{-\alpha}), \text{ for all } \alpha\}.$$

Let $\mathcal{k}$ be the maximal compact subalgebra of $\mathcal{g}$ defined such that $a\in\mathcal{k}$, iff $a\in\mathcal{g}$ and $\sigma a=a$, where σ is the involutive automorphism. Let $\mathcal{p}$ be the subspace of $\mathcal{g}$ such that $a\in\mathcal{p}$ iff $a\in\mathcal{g}$ and $\sigma a=-a$. Thus the basis of $\mathcal{k}$ consists of

$$\{ih_\alpha,\ (e_\alpha+e_{-\alpha}),\ i(e_\alpha-e_{-\alpha}), \text{ for all } \alpha$$
$$\text{such that } \exp\alpha(h)=1\}$$

and the basis of $\mathcal{p}$ consists of

$$\{i(e_\alpha+e_{-\alpha}),(e_\alpha-e_{-\alpha}), \text{ for all } \alpha$$
$$\text{such that } \exp\alpha(h)=-1\}.$$

Here, $\alpha(h)=B(h,h_\alpha)$, $B(x,y)$ being the Killing form of $\mathcal{g}_c$.

Let $\mathcal{a}$ be the maximal Abelian subalgebra of $\mathcal{p}$ with dimension $|m_1|$ and $\mathcal{m}$ the centralizer of $\mathcal{a}$ in $\mathcal{k}$. Complexification of $(\mathcal{a}\oplus\mathcal{m})$ gives a Cartan subalgebra $\mathcal{h}'$ of $\mathcal{g}_c$ with basis $H_1',H_2',\ldots,H_l'$. Then there exists an inner automorphism V that maps $\mathcal{h}'$ into $\mathcal{h}$ such that

$$H_j = VH_j', \quad \text{where } V=\prod_\alpha V_\alpha, \quad j=1,\ldots,l. \tag{2.2}$$

Let Δ^+ denote the set of positive roots, i.e., a positive subspace of Δ.

Then,

$$h_\alpha = \sum_{j=1}^{l} b_j(\alpha)H_j, \quad \alpha\in\Delta^+, \quad \text{if } b_j(\alpha)>0$$

and (2.3)

if j is the least index, then $b_j(\alpha)\neq 0$.

Correspondingly if the α' are the roots defined with respect to $\mathcal{h}'$ then

$$h' = \sum_{j=1}^{l} c_j(\alpha')H_j', \quad \forall\, h'\in\mathcal{h}', \quad h'_{\alpha'}\in\mathcal{h}'. \tag{2.4}$$

Here, $h_\alpha = Vh'_{\alpha'}$ and $c_j(\alpha') = b_j(\alpha)$. The positive roots can again be divided into the two following classes:

 0022-2488/85/102485-06$02.50

(i) $P_+ = \{\alpha:\ \alpha \in \Delta^+, \alpha(h) \neq \alpha(V\sigma V^{-1}h)\}$;

(ii) $P_- = \{\alpha:\ \alpha \in \Delta^+, \alpha(h) = \alpha(V\sigma V^{-1}h),\ \forall\ h \in \mathcal{h}\}$. (2.5)

Let the subalgebra $\tilde{n}$ be spanned by elements $V^{-1}e_\alpha$ for $\alpha \in P_+$ and $n = \tilde{n} \cap \mathcal{g}$, where $\tilde{n}$ and n are the nilpotent subalgebras of $\mathcal{g}_c$ and $\mathcal{g}$, respectively. Thus the Iwasawa decomposition of $\mathcal{g}$ is given as

$$\mathcal{g} = \mathcal{h} \oplus a \oplus n. \tag{2.6}$$

Now we discuss the Langlands decomposition. For the sake of completeness of computational convenience we briefly recapitulate some of the essential features of the Langlands decomposition of parabolic subalgebras.

A minimal parabolic subalgebra is defined to be any subalgebra that is conjugate to

$$\mathcal{h}_1 = m \oplus a \oplus n. \tag{2.7}$$

Any subalgebra of $\mathcal{g}$ containing a minimal parabolic subalgebra is a general parabolic subalgebra. There exist $2^{|m_1|}$ conjugacy classes of parabolic subalgebras of $\mathcal{g}$ and in each class there is a standard parabolic subalgebra which can be obtained as follows.

Let Σ be the set of roots for a and let ψ be the set of positive roots in Σ. Let θ denote the subset of ψ. Let $\langle\theta\rangle$ denote the set of roots in Σ which arises as linear combinations of roots in θ. We define

$$\langle\theta\rangle_+ = \Sigma_+ \cap \langle\theta\rangle, \quad \langle\theta\rangle_- = \Sigma_- \cap \langle\theta\rangle, \tag{2.8}$$

where Σ_+, Σ_- denote the positive and negative roots in Σ.

Let $n_+(\theta), n_-(\theta), n(\theta)$ denote the subspace of a corresponding to $\langle\theta\rangle_+, \langle\theta\rangle_-$ and $\{\Sigma_+ - \langle\theta\rangle_+\}$. Now, let us define

$$a_\theta = \{a \in a,\ \lambda(a) = 0, \forall\ \lambda \in \theta\}, \tag{2.9}$$

and let $a(\theta)$ be the orthogonal complement of a_θ in a with respect to the Cartan–Killing form.

Then

$$p_\theta = m_\theta \oplus a_\theta \oplus n_\theta \tag{2.10}$$

is a parabolic subalgebra of $\mathcal{g}$, where

$$m_\theta = m \oplus n_+(\theta) \oplus n_-(\theta) \oplus a(\theta). \tag{2.11}$$

A real Cartan subalgebra $\mathcal{h}$ is said to be σ invariant if

$$\mathcal{h} = (\mathcal{h} \cap \mathcal{k}) \oplus (\mathcal{h} \cap \mathcal{p}). \tag{2.12}$$

A parabolic subalgebra p_θ is said to be cuspidal if there exists a σ-invariant real Cartan subalgebra $\mathcal{h}$ such that

$$a_\theta = \mathcal{h} \cap \mathcal{p}.$$

This shows that the minimal parabolic subalgebra is cuspidal.

III. SATAKE DIAGRAMS AND IWASAWA DECOMPOSITION OF $\mathcal{f}_4(-20)$

A. Satake diagram and computation of inner automorphisms for $\mathcal{f}_4(-20)$

Let $\mathcal{Y}$ be a compact semisimple symmetric space, S the group of displacements, $\mathcal{U}$ the isotropy group of the base point, E a maximal torus in $\mathcal{Y}$, and $T \supset Q(E)$ a maximal torus in S. Let R be the root system of S relative to T and R the root system of $\mathcal{Y}$ relative to E. For $\alpha \in R$, let

$$\overline{\alpha} = 2\alpha|\mathcal{U} = \alpha - \sigma(\alpha),$$

where $\mathcal{U}$ is the Abelian subalgebra of the Lie algebra $\mathcal{g}$ and σ is the involutive automorphism of α. Then,

$$R_- = \{\overline{\alpha}: \overline{\alpha} \neq 0,\ \alpha \in R\}. \tag{3.1}$$

Also, let

$$R_0 = \{\alpha \in R: \overline{\alpha} = 0\}. \tag{3.2}$$

Let w_- be a Weyl chamber in $\mathcal{U}$ and w a Weyl chamber in $\mathcal{g}$ such that $\overline{w}_- \subset \overline{w}$. Let B_- (resp. B) denote the basis of R_- (resp. R) corresponding to w_- (resp. w). Then, B_0 will be the basis of R_0, where $B_0 = B \cap R_0$. Let $B/B_0 = \{\alpha_1,...,\alpha_r\}$, and $B_0 = \{\beta_1,...,\beta_s\}$. Then, it can be shown that[2]

$$-\sigma(\alpha_i) = \alpha_{\pi(i)} + \sum_l n_{il}\beta_l, \tag{3.3}$$

where π is the involutive permutation of $\{1,2,...,r\}$ and the n_{il} are non-negative integers.

We have

$$B_- = \{\overline{\alpha}:\ \alpha \in B/B_0\},$$

and the rank of $\mathcal{Y}$ equals the number of cycles of π. We now associate with B its Satake diagram Σ as follows. In the Dynkin diagram of B, denote the roots α_i by ○ as usual (white roots) and the roots β by ● (black roots). If $\pi(i) = k$ indicate this by ○ ○. The Satake diagrams determine the involution σ of R uniquely. As we know,

$$\sigma(\beta_i) = \beta_i.$$

The maximal root for $\mathcal{f}_4$ is $W = 2\alpha_1 + 4\alpha_2 + 3\alpha_3 + 2\alpha_4$. In the case of a Satake diagram corresponding to a symmetric space of rank one, $B_- = \{\lambda\}$. If α_i is a white root, then $\overline{\alpha}_i = \lambda$. Here $R_- = \{\pm\lambda\}$, or $R_- = \{\pm\lambda, \pm 2\lambda\}$, and since $\overline{W} \in R_-$, we have only two possibilities:

$$\text{case (i):}\ -\sigma(\alpha_1) = \alpha_1 + 3\alpha_2 + 2\alpha_3 + \alpha_4,$$
$$\text{case (ii):}\ -\sigma(\alpha_4) = 2\alpha_1 + 4\alpha_2 + 3\alpha_3 + \alpha_4. \tag{3.4}$$

These are shown in Fig 1. For the given roots of $\mathcal{f}_4$ we see that in case (ii) $\alpha = 2\alpha_1 + 3\alpha_2 + 2\alpha_3 + \alpha_4 \in R$. Also $\alpha + \sigma(\alpha) = 2\alpha_1 + 2\alpha_2 + \alpha_3 \in R$, which cannot be possible. Hence Fig. (I) (ii) is ruled out. Thus the Satake diagram for $\mathcal{f}_4$ is for $\mathcal{f}_4(-20)$ and is given by Fig. 1 (i). From this diagram we have the following information:

$$-\sigma(\alpha_1) = \alpha_1 + 3\alpha_2 + 2\alpha_3 + \alpha_4,$$
$$\sigma(\alpha_2) = \alpha_2, \quad \sigma(\alpha_3) = \alpha_3, \quad \sigma(\alpha_4) = \alpha_4). \tag{3.5}$$

The automorphisms for the other roots of $\mathcal{f}_4$ are determined from the relation

$$\phi(A + B) = \phi(A) + \phi(B),$$

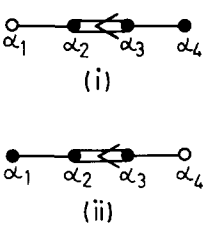


FIG. 1. Satake diagrams for $\mathcal{f}_4(-20)$.

ϕ being the automorphism. Thus,

$$\begin{aligned}
&\sigma(\alpha_1+\alpha_2) = -(\alpha_1+2\alpha_2+2\alpha_3+\alpha_4),\\
&\sigma(\alpha_2+\alpha_3) = \alpha_2+\alpha_3,\\
&\sigma(\alpha_3+\alpha_4) = \alpha_3+\alpha_4,\\
&\sigma(2\alpha_2+\alpha_3) = 2\alpha_2+\alpha_3,\\
&\sigma(\alpha_1+\alpha_2+\alpha_3) = -(\alpha_1+2\alpha_2+\alpha_3+\alpha_4),\\
&\sigma(\alpha_2+\alpha_3+\alpha_4) = \alpha_2+\alpha_3+\alpha_4,\\
&\sigma(\alpha_1+\alpha_2+\alpha_3+\alpha_4) = -(\alpha_1+2\alpha_2+\alpha_3),\\
&\sigma(2\alpha_1+2\alpha_2+\alpha_3) = -(2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4),\\
&\sigma(2\alpha_2+\alpha_3+\alpha_4) = 2\alpha_2+\alpha_3+\alpha_4,\\
&\sigma(2\alpha_1+2\alpha_2+\alpha_3+\alpha_4) = -(2\alpha_1+4\alpha_2+3\alpha_3+\alpha_4),\\
&\sigma(2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4) = -(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4),\\
&\sigma(2\alpha_1+3\alpha_2+2\alpha_3+\alpha_4) = -(2\alpha_1+3\alpha_2+2\alpha_3+\alpha_4).
\end{aligned} \tag{3.6}$$

B. Iwasawa decomposition for $\mathcal{L}_4(-20)$

The Dynkin diagram for the algebra $\mathcal{L}_4$ can be seen in Fig. 2. The roots of $\mathcal{L}_4$ are determined as follows.

If α and β are two simple roots, then $2[(\alpha,\beta)/(\alpha,\alpha)] = m$ (an integer) and $\beta - 2[(\beta,\alpha)/(\alpha,\alpha)]\alpha$ will be a root. The α string of roots containing β is given as $\beta - r\alpha,\ldots,\beta + q\alpha$. The roots for $\mathcal{L}_4$ are

$$\begin{aligned}
\Delta = \{&\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_1+\alpha_2,\alpha_2+\alpha_3,2\alpha_2+\alpha_3,\alpha_3+\alpha_4,\\
&\alpha_1+\alpha_2+\alpha_3,\alpha_2+\alpha_3+\alpha_4,\\
&\alpha_1+\alpha_2+\alpha_3+\alpha_4,2\alpha_1+2\alpha_2+\alpha_3,\\
&2\alpha_2+\alpha_3+\alpha_4,\alpha_1+2\alpha_2+\alpha_3,2\alpha_1+2\alpha_2+\alpha_3+\alpha_4,\\
&2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4,2\alpha_2+2\alpha_3+\alpha_4,\\
&\alpha_1+2\alpha_2+\alpha_3+\alpha_4,\alpha_1+2\alpha_2+2\alpha_3+\alpha_4,\\
&\alpha_1+3\alpha_2+2\alpha_3+\alpha_4,2\alpha_1+4\alpha_2+3\alpha_3+\alpha_4,\\
&2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4,2\alpha_1+3\alpha_2+2\alpha_3+\alpha_4,\\
&2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4\}.
\end{aligned} \tag{3.7}$$

From Fig. 2 we see that $\exp\alpha(h) = +1$, for

$$\begin{aligned}
\alpha = \{&\alpha_2,\alpha_3,\alpha_4,\alpha_1+\alpha_2,\alpha_2+\alpha_3,\alpha_3+\alpha_4,\alpha_2+\alpha_3+\alpha_4,\\
&2\alpha_2+\alpha_3+\alpha_4,2\alpha_2+\alpha_3,2\alpha_2+2\alpha_3+\alpha_4,\\
&\alpha_1+\alpha_2+\alpha_3,2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4,\\
&2\alpha_1+4\alpha_2+3\alpha_3+\alpha_4,2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4,\\
&\alpha_1+3\alpha_2+2\alpha_3+\alpha_4,\alpha_1+\alpha_2+\alpha_3+\alpha_4\}
\end{aligned} \tag{3.8}$$

and $\exp\alpha(h) = -1$, for

$$\begin{aligned}
\alpha = \{&\alpha_1,\alpha_1+2\alpha_2+\alpha_3,2\alpha_1+2\alpha_2+\alpha_3,\\
&\alpha_1+2\alpha_2+\alpha_3+\alpha_4,\alpha_1+2\alpha_2+2\alpha_3+\alpha_4,\\
&2\alpha_1+2\alpha_2+\alpha_3+\alpha_4,2\alpha_1+3\alpha_2+2\alpha_3+\alpha_4,\\
&2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4\}.
\end{aligned} \tag{3.9}$$

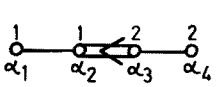


FIG. 2. Dynkin diagram for $\mathcal{L}_4$.

Thus, $\mathcal{K}$ has a basis consisting of ih_α for $\alpha = \alpha_1,\alpha_2,\alpha_3,\alpha_4$ and $(e_\alpha + e_{-\alpha}), i(e_\alpha - e_{-\alpha})$ for α given by (3.8). Also $\mathcal{K} = \{i(e_\alpha + e_{-\alpha}),(e_\alpha - e_{-\alpha})$, where α are the roots given by (3.9).

We select a, a maximal Abelian subalgebra of $\mathcal{K}$ spanned by

$$H_1' = i(e_\alpha + e_{-\alpha}), \quad \alpha = 2\alpha_1+2\alpha_2+\alpha_3+\alpha_4. \tag{3.10}$$

Then, m, the centralizer of a in $\mathcal{K}$, is given by $\{-iH_2', -iH_3', -iH_4'\}$ with

$$\begin{aligned}
&-iH_j' = ih_\beta,\\
&\quad \text{for } \beta = \{\alpha_3+\alpha_4, 2\alpha_2+\alpha_3+\alpha_4,\\
&\qquad 2\alpha_1+4\alpha_2+3\alpha_3+\alpha_4\}.
\end{aligned} \tag{3.11}$$

The inner automorphism

$$V = \prod_\alpha V_\alpha$$

is such that

$$H_j = VH_j' = -\{2/(\delta_j,\delta_j)\}^{1/2}h_{\delta_j}, \tag{3.12}$$

where

$$\begin{aligned}
\delta_j = \{&2\alpha_1+2\alpha_2+\alpha_3+\alpha_4,\alpha_3+\alpha_4,2\alpha_2+\alpha_3+\alpha_4,\\
&2\alpha_1+4\alpha_2+3\alpha_3+\alpha_4\}.
\end{aligned}$$

We have

$$\begin{aligned}
H_1 &= -\{2h_{\alpha_1}+2h_{\alpha_2}+h_{\alpha_3}+h_{\alpha_4}\},\\
H_2 &= -\{h_{\alpha_3}+h_{\alpha_4}\},\\
H_3 &= -\{2h_{\alpha_2}+h_{\alpha_3}+h_{\alpha_4}\},\\
H_4 &= -\{2h_{\alpha_1}+4h_{\alpha_2}+3h_{\alpha_3}+h_{\alpha_4}\}.
\end{aligned} \tag{3.13}$$

Using (2.3), we have

$$\begin{aligned}
\Delta^+ = \{&-\alpha_1,\alpha_2,\alpha_3,-\alpha_4,-(\alpha_1+\alpha_2),\alpha_2+\alpha_3,\\
&-(\alpha_3+\alpha_4),\alpha_1+\alpha_2+\alpha_3,-(\alpha_2+\alpha_3+\alpha_4),\\
&-(\alpha_1+\alpha_2+\alpha_3+\alpha_4),2\alpha_2+\alpha_3,\\
&-(2\alpha_2+\alpha_3+\alpha_4),2\alpha_2+2\alpha_3+\alpha_4,\\
&\alpha_1+2\alpha_2+\alpha_3,-(2\alpha_1+2\alpha_2+\alpha_3),\\
&-(2\alpha_1+2\alpha_2+\alpha_3+\alpha_4),\\
&-(\alpha_1+2\alpha_2+2\alpha_3+\alpha_4),\\
&-(\alpha_1+2\alpha_2+\alpha_3+\alpha_4),\\
&-(2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4),\\
&-(2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4),\\
&-(2\alpha_1+4\alpha_2+3\alpha_3+\alpha_4),\\
&-(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4),\\
&-(\alpha_1+3\alpha_2+2\alpha_3+\alpha_4),\\
&-(2\alpha_1+3\alpha_2+2\alpha_3+\alpha_4)\}.
\end{aligned} \tag{3.14}$$

Using (2.5), we have

$$\begin{aligned}
P_+ = \{&-\alpha_1,\alpha_3,-\alpha_4,-(\alpha_1+\alpha_2),\alpha_2+\alpha_3,-(\alpha_3+\alpha_4),\\
&-(\alpha_1+\alpha_2+\alpha_3+\alpha_4),2\alpha_2+\alpha_3,2\alpha_2+2\alpha_3+\alpha_4,\\
&-(2\alpha_1+2\alpha_2+\alpha_3),-(2\alpha_1+2\alpha_2+\alpha_3+\alpha_4),\\
&-(2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4),
\end{aligned}$$

$$
\begin{aligned}
&-(\alpha_1+2\alpha_2+\alpha_3+\alpha_4),\\
&-(2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4),\\
&-(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4)\},
\end{aligned}
\tag{3.15}
$$

and

$$
\begin{aligned}
P_- = \{&\alpha_2,\alpha_1+\alpha_2+\alpha_3,-(\alpha_2+\alpha_3+\alpha_4),\\
&-(2\alpha_2+\alpha_3+\alpha_4),\alpha_1+2\alpha_2+\alpha_3,\\
&-(\alpha_1+2\alpha_2+2\alpha_3+\alpha_4),\\
&-(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4),\\
&-(\alpha_1+3\alpha_2+2\alpha_3+\alpha_4),\\
&-(2\alpha_1+3\alpha_2+2\alpha_3+\alpha_4)\}.
\end{aligned}
\tag{3.16}
$$

The basis elements of $\tilde{n}\subset\mathcal{G}_c$ are given by $V^{-1}e_a$ for $\alpha\in P_+$ and are given by[7]

$$
\begin{aligned}
&V_\alpha^{-1}e_{-\alpha} = -\tfrac{1}{2}(e_\alpha - e_{-\alpha}) - \tfrac{1}{2}i[2/(\alpha,\alpha)]^{1/2}h_\alpha,\\
&\quad \text{for } \alpha = 2\alpha_1+2\alpha_2+\alpha_3+\alpha_4,\\
&V_\alpha^{-1}e_{2\alpha_2+\alpha_3} = 2^{-1/2}e_{2\alpha_2+\alpha_3} - i2^{-1/2}\\
&\qquad\times(\operatorname{sgn} N_{\alpha,2\alpha_2+\alpha_3})e_{2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4},\\
&V_\alpha^{-1}e_{-(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4)}\\
&\quad = -i2^{-1/2}\operatorname{sgn}(N_{\alpha_1,-(2\alpha_2+\alpha_3)})e_{-(2\alpha_2+\alpha_3)}\\
&\qquad + 2^{-1/2}e_{-(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4)},\\
&V_\alpha^{-1}e_{\alpha_3} = 2^{-1/2}e_{\alpha_3} - i2^{-1/2}(\operatorname{sgn} N_{\alpha,\alpha_3})e_{2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4},\\
&V_\alpha^{-1}e_{-(2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4)}\\
&\quad = -i2^{-1/2}(\operatorname{sgn} N_{\alpha,-\alpha_3})e_{-\alpha_3}\\
&\qquad + 2^{-1/2}e_{-(2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4)},\\
&V_\alpha^{-1}e_{-\alpha_1} = 2^{-1/2}e_{-\alpha} - i2^{-1/2}\\
&\qquad\times(\operatorname{sgn} N^{-\alpha}_{\alpha,\alpha_1})e_{-(\alpha_1+2\alpha_2+\alpha_3+\alpha_4)},\\
&V_\alpha^{-1}e_{-(2\alpha_1+2\alpha_2+\alpha_3+\alpha_4)}\\
&\quad = -i2^{-1/2}(\operatorname{sgn} N_{\alpha,\alpha_1})e_{\alpha_1} + 2^{-1/2}e_{-(\alpha_1+2\alpha_2+\alpha_3+\alpha_4)},\\
&V_\alpha^{-1}e_{-\alpha_4} = 2^{-1/2}e_{-\alpha_4} - i2^{-1/2}\\
&\qquad\times(\operatorname{sgn} N_{\alpha,-\alpha_4})e_{-(2\alpha_1+2\alpha_2+\alpha_3)},\\
&V_\alpha^{-1}e_{(2\alpha_1+2\alpha_2+\alpha_3)}\\
&\quad = -i2^{-1/2}(\operatorname{sgn} N_{\alpha,\alpha_4})e_{\alpha_4} + 2^{-1/2}e_{(2\alpha_1+2\alpha_2+\alpha_3)},\\
&V_\alpha^{-1}e_{-(\alpha_1+\alpha_2)}\\
&\quad = 2^{-1/2}e_{-(\alpha_1+\alpha_2)} - i2^{-1/2}\\
&\qquad\times(\operatorname{sgn} N_{\alpha,(-(\alpha_1+\alpha_2))})e_{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4)},\\
&V_\alpha^{-1}e_{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4)}\\
&\quad = -i2^{-1/2}(\operatorname{sgn} N_{\alpha,\alpha_1+\alpha_2})e_{\alpha_1+\alpha_2}\\
&\qquad + 2^{-1/2}e_{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4)},\\
&V_\alpha^{-1}e_{(2\alpha_2+2\alpha_3+\alpha_4)}\\
&\quad = 2^{-1/2}e_{2\alpha_2+2\alpha_3+\alpha_4} - i2^{-1/2}\\
&\qquad\times\operatorname{sgn} N_{\alpha,2\alpha_2+2\alpha_3+\alpha_4}e_{2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4},\\
&V_\alpha^{-1}e_{-(2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4)}\\
&\quad = -i2^{-1/2}\operatorname{sgn} N_{\alpha,-(2\alpha_2+2\alpha_3+\alpha_4)}e_{-(2\alpha_2+2\alpha_3+\alpha_4)}\\
&\qquad + 2^{-1/2}e_{-(2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4)},\\
&V_\alpha^{-1}e_{\alpha_2+\alpha_3} = e_{\alpha_2+\alpha_3},\quad V_\alpha^{-1}e_{-(\alpha_3+\alpha_4)} = e_{-(\alpha_3+\alpha_4)}.
\end{aligned}
\tag{3.17}
$$

Now n is calculated from $\tilde{n}\cap\mathcal{G}$ and it has the following elements:

$$
\begin{aligned}
&-\tfrac{1}{2}\{e_{2\alpha_1+2\alpha_2+\alpha_3+\alpha_4} - e_{-(2\alpha_1+2\alpha_2+\alpha_3+\alpha_4)}\} - \tfrac{1}{2}i[2h_{\alpha_1}+2h_{\alpha_2}+h_{\alpha_3}+h_{\alpha_4}],\\
&2^{-1/2}(e_{2\alpha_2+\alpha_3}+e_{-(2\alpha_2+\alpha_3)}) - i2^{-1/2}(\operatorname{sgn} N_{\alpha,2\alpha_2+\alpha_3})(e_{2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4}+e_{-(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4)}),\\
&2^{-1/2}\operatorname{sgn} N_{\alpha,-(2\alpha_2+\alpha_3)}(e_{2\alpha_2+\alpha_3}-e_{-(2\alpha_2+\alpha_3)}) + i2^{-1/2}(e_{2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4}-e_{-(2\alpha_1+4\alpha_2+2\alpha_3+\alpha_4)}),\\
&2^{-1/2}(e_{\alpha_3}+e_{\alpha_3}) - i2^{-1/2}\operatorname{sgn} N_{\alpha,\alpha_3}(e_{2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4}+e_{-(2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4)}),\\
&2^{-1/2}\operatorname{sgn} N_{\alpha,\alpha_3}(e_{\alpha_3}-e_{-\alpha_3}) + i2^{-1/2}(e_{2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4}-e_{2\alpha_1+2\alpha_2+2\alpha_3+\alpha_4}),\\
&i2^{-1/2}(e_{\alpha_1}-e_{-\alpha_1}) + 2^{-1/2}\operatorname{sgn} N_{\alpha,-\alpha_1}(e_{\alpha_1+2\alpha_2+\alpha_3+\alpha_4}-e_{(\alpha_1+2\alpha_2+\alpha_3+\alpha_4)}),\\
&2^{-1/2}\operatorname{sgn} N_{\alpha,\alpha_1}(e_{\alpha_1}-e_{-\alpha_1}) + i2^{-1/2}(e_{\alpha_1+2\alpha_2+\alpha_3+\alpha_4}-e_{(\alpha_1+2\alpha_2+\alpha_3+\alpha_4)}),\\
&i2^{-1/2}(e_{\alpha_4}-e_{-\alpha_4}) + 2^{-1/2}\operatorname{sgn} N_{\alpha,-\alpha_4}(e_{2\alpha_1+2\alpha_2+\alpha_3}-e_{-(2\alpha_1+2\alpha_2+\alpha_3)}),\\
&2^{-1/2}\operatorname{sgn} N_{\alpha,\alpha_4}(e_{\alpha_4}-e_{-\alpha_4}) + i2^{-1/2}(e_{2\alpha_1+2\alpha_2+\alpha_3}-e_{-(2\alpha_1+2\alpha_2+\alpha_3)}),\\
&i2^{-1/2}(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}) + 2^{-1/2}\operatorname{sgn} N_{\alpha,-(\alpha_1+\alpha_2)}(e_{\alpha_1+\alpha_2+\alpha_3+\alpha_4}-e_{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4)}),\\
&-i2^{-1/2}\operatorname{sgn} N_{\alpha,\alpha_1+\alpha_2}(e_{\alpha_1+\alpha_2}-e_{-(\alpha_1+\alpha_2)}) + i2^{-1/2}(e_{\alpha_1+\alpha_2+\alpha_3+\alpha_4}-e_{-(\alpha_1+\alpha_2+\alpha_3+\alpha_4)}),\\
&2^{-1/2}(e_{2\alpha_2+2\alpha_3+\alpha_4}+e_{-(2\alpha_2+2\alpha_3+\alpha_4)}) - i2^{-1/2}\operatorname{sgn} N_{\alpha,2\alpha_2+2\alpha_3+\alpha_4}(e_{2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4}+e_{-(2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4)}),\\
&2^{-1/2}\operatorname{sgn} N_{\alpha-(2\alpha_2+2\alpha_3+\alpha_4)}(e_{2\alpha_2+2\alpha_3+\alpha_4}-e_{-(2\alpha_2+2\alpha_3+\alpha_4)}) + i2^{-1/2}(e_{2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4}-e_{-(2\alpha_1+4\alpha_2+3\alpha_3+2\alpha_4)}),\\
&e_{\alpha_2+\alpha_3}+e_{-(\alpha_2+\alpha_3)},\quad i(e_{\alpha_3+\alpha_4}-e_{-(\alpha_3+\alpha_4)}).
\end{aligned}
\tag{3.18}
$$

IV. PRINCIPAL SERIES REPRESENTATIONS OF $F_4(-20)$

A. Parabolic subgroups of $F_4(-20)$

As noted earlier, there are $2^{|m_1|}$ classes of parabolic subalgebras of $\mathcal{G}$, $|m_1|$ being the dimension of a. For $f_4(-20)$ $\dim|m_1| = 1$. Therefore, there will be two parabolic subalgebras: one being the minimal parabolic and the other the Lie algebra itself.[7] The minimal parabolic subalgebra

$$p_1 = m\oplus a\oplus n$$

is given by the sum of

$$m = \{ih_\alpha, \alpha = \alpha_3 + \alpha_4, 2\alpha_2 + \alpha_3 + \alpha_4, 2\alpha_1 + 4\alpha_2 + 3\alpha_3 + \alpha_4\},$$

and

$$a = \{i(e_\alpha + e_{-\alpha}), \alpha = 2\alpha_1 + 2\alpha_2 + \alpha_3 + \alpha_4\},$$

and n is given by Eq. (3.18). The minimal parabolic subalgebra being the cuspidal parabolic subalgebra leads to the following definition of a series of representations[8]:

$$\rho = \mathrm{ind}_{P_1}^{G}(\sigma \times \tau), \quad \sigma \in \hat{M}, \quad \tau \in \hat{A}.$$

We have that $P_1 = MAN$, a cuspidal parabolic subgroup, defines the principal series representations. Since P_1 is minimal parabolic and $\sigma = 1_M$, M being Abelian, ρ is irreducible. Note here that P_1, M, A, N, and G are the corresponding groups of the Lie algebras f_1, m, a, n, and g, respectively.

B. Polarization and Kostants induced representation for $F_4(-20)$ (principal series)

Let g be semisimple and let h be a σ-invariant split Cartan subalgebra, i.e.,

$$h = (h \cap f) \oplus (h \cap e), \tag{4.1}$$

$g = h \oplus f$ being the Cartan decomposition. Let Δ denote the set of roots relative to (g, h).

Let $x \in x = \mathbf{h} \cap \mathbf{p}$, i.e., $x = i(e_\alpha + e_{-\alpha})$, for $\alpha = 2\alpha_1 + 2\alpha_2 + \alpha_3 + \alpha_4$ for $f_4(-20)$. Let us define

$$\langle \eta \rangle = \{\alpha \in \Delta : \alpha(x) = 0\}. \tag{4.2}$$

This condition, using (3.7), gives us

$$\begin{aligned}\langle \eta \rangle = \{ & \pm\alpha_2, \pm(\alpha_3 + \alpha_4), \pm(\alpha_1 + \alpha_2 + \alpha_3), \\ & \pm(\alpha_1 + 2\alpha_2 + \alpha_3), \\ & \pm(\alpha_2 + \alpha_3 + \alpha_4), \pm(2\alpha_2 + \alpha_3 + \alpha_4), \\ & \pm(\alpha_1 + 2\alpha_2 + 2\alpha_3 + \alpha_4), \\ & \pm(2\alpha_1 + 2\alpha_2 + \alpha_3 + \alpha_4), \\ & \pm(\alpha_1 + 3\alpha_2 + 2\alpha_3 + \alpha_4), \\ & \pm(2\alpha_1 + 4\alpha_2 + 3\alpha_3 + \alpha_4)\}.\end{aligned} \tag{4.3}$$

Let G_x be the isotropy group and g_x the corresponding Lie algebra. Let χ be the character on G_x associated with x. Then g_x can be written as[4]

$$g_x = h + n_1, \tag{4.4}$$

when

$$n_1 = \text{lin span } \{ g_\alpha : \alpha \in \langle \eta \rangle \} \cap g. \tag{4.5}$$

Let Δ^+ denote the positive subspace of Δ. Let x be quantizable. There exists a positive polarization $\mathcal{S}_x$ at x which is of the form[3,4]

$$\mathcal{S}_x = h + n_1 + n_2, \tag{4.6}$$

where

$$n_2 = \{ g_\beta, \beta \in \Delta^+, \quad \beta \notin \langle \eta \rangle \}$$

and the β's are given by

$$\begin{aligned}\beta = \{ & -\alpha_1, \alpha_3, -\alpha_4, -(\alpha_1 + \alpha_2), \alpha_2 + \alpha_3, 2\alpha_2 + \alpha_3, \\ & -(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4), 2\alpha_2 + 2\alpha_3 + \alpha_4, \\ & -(2\alpha_1 + 2\alpha_2 + \alpha_3), -(2\alpha_1 + 2\alpha_2 + 2\alpha_3 + \alpha_4), \\ & -(\alpha_1 + 2\alpha_2 + \alpha_3 + \alpha_4), \\ & -(2\alpha_1 + 4\alpha_2 + 3\alpha_3 + 2\alpha_4), \\ & -(2\alpha_1 + 4\alpha_2 + 2\alpha_3 + \alpha_4), \\ & -(2\alpha_1 + 3\alpha_2 + 2\alpha_3 + \alpha_4)\}.\end{aligned} \tag{4.7}$$

We note here that the polarization $\mathcal{S}$ satisfies the following properties[3,4]: (i) $\langle B, [\mathcal{S}, \mathcal{S}] \rangle = 0$, where B is extended to $g_c \times g_c$ by complex linearity and $\mathcal{S}$ is maximal w.r.t. this condition, (ii) $\dim_c \mathcal{S} = \frac{1}{2}(\dim_R g + \dim g_x)$; (iii) $\mathcal{S} + \overline{\mathcal{S}}$ is a Lie subalgebra of g_c, where the bar indicates the complex conjugation; and (iv) $g_x \subseteq \mathcal{S}$ and $\mathcal{S}$ is ad G_x stable, i.e.,

$$\mathrm{ad}_s \mathcal{S} \subseteq \mathcal{S}, \quad \forall\ s \in G_x,$$

$\mathcal{S}$ is said to be positive if

$$iB(x, [\bar{z}, z]) \geqslant 0, \quad \forall\ z \in \mathcal{S}. \tag{4.8}$$

Let us define[3,4]

$$d = \mathcal{S} \cap g, \quad e = (\mathcal{S} + \overline{\mathcal{S}}) \cap g, \tag{4.9}$$

$D = D_0 G_x$, and $E = E_0 G_x$, where D_0 and E_0 are the analytic subgroups corresponding to

$$d = \mathcal{S}_x, \quad e = g. \tag{4.10}$$

We extend χ from G_x to D. The Auslander–Kostant induction scheme[3,4] reduces to

$$\sigma = \widetilde{\mathrm{ind}}_{S_x}^{G_c}(\chi). \tag{4.11}$$

We note that the representation (4.11) yields the principal series representations of $F_4(-20)$, which was defined through the induction scheme from the cuspidal parabolic subgroup P_1 discussed in Sec. IV A.

V. CONCLUSIONS

(1) To summarize our analysis, we have presented the Iwasawa and Langlands decompositions of $f_4(-20)$ in detail. Use of Satake diagrams facilitates the Iwasawa decompositions.

(2) (a) The principal series representations are obtained by induction from the known character associated with the cuspidal parabolic subgroup P_1; and (b) the same result is achieved by resorting to Kostant's induction scheme.

(3) The noncompact exceptional Lie groups were recently used in supergravity theories.[9] So it is interesting to study the structure of the $F_4(-20)$ group that may open new vistas in elementary particle physics.

(4) The other noncompact exceptional Lie groups could be studied in a similar fashion as $F_4(-20)$.

We wish to report on other classes of UIR's of $F_4(-20)$ in a separate paper.

[1] A. O. Barut and R. Raczka, *Theory of Group Representations and Applications* (PWN–Blish Scientific, Warsaw, 1977); R. Hermann, *Lie Groups for Physicists* (Benjamin, New York, 1966).

[2] O. Loos, *Symmetric Spaces* (Benjamin, New York, 1965), Vol. II.

[3] L. Auslander and B. Kostant, Invent.Math. **14**, 255 (1971); L. P. Rothschild and J. A. Wolf, Ann. Sci. Ec. Norm. Sup. 7, 155 (1974).

[4]B. R. Sitaram and K. C. Tripathy, J. Math. Phys. **23**, 206,481,484 (1982).
[5]H. Chandra, Trans. Am. Math. Soc. **75**, 185 (1953); **76**, 26 (1954); **76**, 34 (1954).
[6]K. Iwasawa, Ann. Math. **50**, 507 (1949); R. P. Langlands, "Problems in the theory of automorphic forms," *Lecture Notes in Mathematics, Vol. 170* (Springer, Berlin, 1970).
[7]J. F. Cornwell, J. Math. Phys. **16**, 1992 (1975); **20**, 547 (1979).
[8]R. L. Lipsman, "Group Representations," *Lecture Notes in Mathematics, Vol. 388* (Springer, Berlin, 1974).
[9]K. C. Tripathy and V. Sharma, *Proceedings of the 13th International Colloquium on Group Theoretical Methods in Physics,* Maryland, 1984 (World Scientific, Singapore, 1984), p. 562; M. Günaydin, G. Sierra, and P. K.Townsend, Phys. Lett. B **133**, 72 (1983); J. Ellis, M. K. Gaillard, M. Günaydin, and B. Zumino, "Sypersymmetry and non-compact groups in Supergravity," SLAC-PUB-3065, Feb. 1983.

Thermodynamics of dimers on a rectangular $L \times M \times N$ lattice

Alain J. Phares and Francis J. Wunderlich
Department of Physics, Villanova University, Villanova, Pennsylvania 19085

(Received 18 February 1985; accepted for publication 31 May 1985)

The exact closed-form analytic solution of the problem of dimers on infinite two-dimensional and three-dimensional lattices is obtained. Entropy, isothermal compressibility, and constant pressure heat capacity of the system are given in terms of the normalized number density of dimers. The absolute activity of dimers is also given in terms of the normalized number density; it exhibits a behavior near close packing with a critical exponent exactly equal to 2, and with an amplitude $1/(4\phi)$, where ϕ is the molecular freedom per dimer at close packing.

I. INTRODUCTION

Two previous articles,[1] referred to as papers I and II, were devoted to the study of dimers on two-dimensional lattices irrespective of their orientation. Paper I presented a new general mathematical technique for dealing with such problems. The technique was based on the work of McQuistan *et al.*[2] Essentially, recurrence relations were developed which needed decoupling. The systematic decoupling of these linear relations involving constant coefficients was achieved by one of us.[3] Paper I also made the connection between the partition function of dimers with absolute activity x and the largest z root of a polynomial $P(x,z)$. This polynomial depends on the size of the lattice $L \times M \times N$ where L and M are fixed and N is allowed to become infinite. However, the calculations performed in paper I were limited to the value $x = 1$. Paper II is an extension of paper I; results were obtained for values of the activity in the range 0–10 for the partition function, the grand potential, the number density, the entropy, the isothermal compressibility, and the constant pressure specific-heat capacity. An approximate expression was derived for the partition function in terms of the absolute activity for dimers on a square lattice, based on the assumption that the z roots of the polynomial $P(x,z)$ increase exponentially with the number M of compartments in any given row of the $1 \times M \times N$ lattice. No mathematical proof of this fact was supplied beyond the simple observation that this is true with a very good approximation in the range $M = 1$–4. Results derived numerically by Gaunt[4] for the locations of the maxima of entropy, isothermal compressibility, and constant pressure specific-heat capacity were recovered using the approximate expression.

This article has several objectives.

(1) We obtain the thermodynamic properties of aligned dimers on finite and infinite two- and three-dimensional lattices, explicitly in terms of the absolute activity and the normalized number density. For aligned dimers, the largest z root of polynomials $P(x,z)$ is shown to exhibit the exponential behavior with the size M of the lattice. Therefore, this paper gives the first rigorous mathematical justification for the calculations performed in papers I and II.

(2) Following an approach completely different from Kasteleyn's,[5] we calculate the molecular freedom per dimer at close packing Φ for planar $1 \times M \times N$ lattices $(N \to \infty)$, and we recover Kasteleyn's results for $M = 1$, 2, 3, and 4. However, Kasteleyn's formula is much more general since it allows one to obtain Φ for any value of M and it predicts its exact value when both M and N are infinite. The advantage of our technique is that it can be extended to three-dimensional lattices $L \times M \times N$ $(N \to \infty)$. As a special case, we calculate the exact value of Φ for the $2 \times 2 \times N$ lattice, a problem investigated by Hock and McQuistan.[6]

(3) We show that it is possible to obtain an analytical fit to all the thermodynamical quantities for dimers on various square lattices we have calculated in paper II. This is achieved by using the exact analytic results for aligned dimers, where the molecular freedom per dimer at close packing Φ (which is 1 for aligned dimers) is replaced by the value calculated exactly by Kasteleyn,[5] and a parameter Q (which is 2 for aligned dimers) by 4. Agreement with the numerical values of paper II is astonishingly good.

(4) The success of analytical fit for various finite two-dimensional square lattices is extended to obtain *exact* analytical results for the *infinite* two- and three-dimensional lattices based on the knowledge of the molecular freedom per dimer at close packing. In support of the approximate analytical fit for finite three-dimensional lattices, we calculate the thermodynamical properties of dimers on a $2 \times 2 \times N$ lattice, a problem investigated recently by Hock and McQuistan.[6]

The general outline of this article is as follows: Section II develops the solution of dimers of an $L \times M \times N$ lattice parallel to the M axis; dimensions L and M are fixed while N is allowed to become infinitely large. We think of the (L,M) plane as "horizontal," and we refer to this case as the case of horizontal dimers. Here, there is perfect symmetry between dimers parallel to the L axis and dimers parallel to the M axis. The case of dimers parallel to the N axis (or "vertical" dimers) is distinct from the case of horizontal dimers, since the dimers are parallel to the N dimension which is allowed to become infinitely large. For this reason we discuss in a separate section, Sec. III, the case of vertical dimers, or dimers parallel to the N axis. The purpose of Sec. III is to confirm the internal consistency of our method. Indeed, although the point of view is different, the results for vertical dimers should be the same as those for horizontal dimers, when both N and M are allowed to become infinite. In both cases, one discovers the exponential behavior of the largest root mentioned above. In Sec. IV, we derive the values of the thermodynamical quantities explicitly in terms of the normalized number density; we also give their graphical representations and make the comparison with those obtained in

 0022-2488/85/102491-09$02.50

papers I and II for dimers irrespective of their orientation. Another achievement of this section is to obtain the closed-form analytic expression of the absolute activity in terms of the normalized number density, showing the behavior near close packing with a critical exponent exactly equal to two as conjectured by Gaunt.[4] Section V develops a method for calculating the molecular freedom per dimer at close packing based on the knowledge of polynomial $P(x,z)$. The method is applied for dimers on lattices $1\times1\times N$, $1\times2\times N$, $1\times3\times N$, $1\times4\times N$, and $2\times2\times N$. Section VI discusses the analytical fit of the data available for dimers on a two-dimensional lattice and extends the analysis to the three-dimensional case. Finally, Sec. VII is the conclusion.

II. HORIZONTAL DIMERS

The method of McQuistan *et al.* as generalized in paper I calls for generating truncated lattices from the original one. Consider an $L\times M\times N$ lattice space. The horizontal dimers are assumed to be aligned and parallel to the M axis. We view our lattice space as being made of N arrays, each array containing $L\times M$ cells. We now focus on the N th array. Since all possible dimers occupying the lattice are parallel to the M axis, dimers whose one end is in the N th array will therefore have their other end within the same array. Using the notation of paper I, this means that one can only generate lattices of the A type (nontruncated).

The next step is to consider filling the $L\times M\times N$ lattice with q dimers parallel to the M axis, where L and M are fixed and N is allowed to become infinitely large. Let $A(q-p;L,M,N-1)$ be the total number of possible arrangements of $(q-p)$ dimers on the $L\times M\times(N-1)$ lattice when the N th array is occupied by p dimers. Also, let $C(p;L,M)$ be the total number of possible arrangements of p dimers on the N th array containing $L\times M$ cells. It is obvious that the total number of distinct arrangements of q dimers on the whole lattice when p of these dimers are located in the N th array is given by the product.

$$C(p;L,M)A(p-q;L,M,N-1).$$

It is also obvious that the total number of arrangements of the q dimers of the $L\times M\times N$ lattice is the sum of all possible arrangements when the N th array is empty ($p=0$), occupied by only one dimer ($p=1$), by two dimers ($p=2$), etc., until it is fully occupied ($p=p_{\max}$). In other words, one has

$$A(q;L,M,N)=\sum_{p=0}^{p_{\max}}C(p;L,M)A(q-p;L,M,N-1). \quad (1)$$

Upon introducing the bivariant generating function

$$G(x,y)=\sum_{N=0}^{\infty}\sum_{q=0}^{q_{\max}}x^q y^N A(q;L,M,N), \quad (2)$$

and using it in Eq. (1), one obtains

$$G(x,y)=1+\sum_{p=0}^{p_{\max}}x^p y C(p;L,M)G(x,y). \quad (3)$$

This, in turn, gives the closed-form expression of the bivariant function, as the inverse of a polynomial $D(x,y)$ which is of first order in the variable y. Following the method of paper I, one sets $y=1/z$ and searches for the largest z root of the polynomial

$$P(x,z)=D(x,1/z), \quad (4)$$

which, in this case, has only one root, namely,

$$R_3(x;L,M)=\sum_{p=0}^{p_{\max}}x^p C(p;L,M). \quad (5)$$

One next step is to calculate the closed-form expression of the root $R_3(x;L,M)$. This we will achieve by recognizing that $C(p;L,M)$ is the total number of arrangements of fitting p dimers parallel to the M axis on the two-dimensional $L\times M$ lattice. The same approach is to be used over again. We think of this lattice as being made of L rows, each row made up of M cells. We single out the L th row. Let r be the number of aligned dimers filling the L th row, thus allowing $(p-r)$ dimers to fill the remaining $(L-1)$ rows. Then $C(p-r;L-1,M)$ is the total number of distinct arrangements of $(p-r)$ dimers parallel to the M axis in the $(L-1)$ rows of cells. Let $F(r;M)$ be the number of arrangements of r dimers in the L th row containing M cells. As before, it then follows that

$$C(p;L,M)=\sum_{r=0}^{r_{\max}}F(r;M)C(p-r;L-1,M). \quad (6)$$

Again, one introduces the bivariant generating function

$$G_c(x,y)=\sum_{L=0}^{\infty}\sum_{p=0}^{p_{\max}}x^p y^L C(p;L,M). \quad (7)$$

Combining Eqs. (6) and (7), and using again the technique outlined in paper I, one is able to show that

$$\sum_{p=0}^{p_{\max}}x^p C(p;L,M)=\left(\sum_{r=0}^{r_{\max}}x^r F(r;M)\right)^L. \quad (8)$$

We have now reached the last step in our derivation, which is finding the expression of the right-hand side of the above equation. This is the problem of dimers in one-dimensional space, solved already by McQuistan *et al.* and rederived differently in paper I. It is straightforward to show that

$$R_2(x;M)=\sum_{r=0}^{r_{\max}}x^r F(r;M)$$
$$=(z_1^{M+1}-z_2^{M+1})/(z_1-z_2). \quad (9)$$

In Eq. (9), z_1 and z_2 are the z roots of the quadratic equation

$$z_2-z-x=0, \quad (10)$$

namely,

$$z_1=\tfrac{1}{2}\{1+(1+4x)^{1/2}\},\quad z_2=\tfrac{1}{2}\{1-(1+4x)^{1/2}\}. \quad (11)$$

Finally, the root to be used for calculating all the thermodynamical quantities is

$$R_3(x;L,M)=(R_2(x;M))^L$$
$$=((z_1^{M+1}-z_2^{M+1})/(z_1-z_2))^L. \quad (12)$$

As anticipated in the Introduction, Eq. (12) exhibits an exact exponential behavior in L and an almost exact exponential behavior in M, this behavior in M becoming more and more accurate as M becomes larger and larger. This follows from the fact that the positive root z_1 is, in absolute value, larger than the negative root z_2; thus, given the approximation for large values of M

$$R_3(x;L,M)=(z_1(x))^{LM}. \quad (13)$$

We leave to Sec. IV the derivation of the thermodynamic quantities explicitly in terms of the normalized number density, including their graphical representations.

III. VERTICAL DIMERS

We now consider the case of dimers parallel to the N axis. For simplicity, we will assume that $L = 1$. The analysis that follows can be easily extended to any value of L. However, since the purpose of this section is to exhibit the internal consistency of our method, there is very little to be gained by doing our discussion for any value of L. The case $M = 1$ is the one-dimensional problem whose solution is already known (e.g., paper I). We will study the cases $M = 2$ and 3, with the constraint that the dimers are parallel to the N axis.

There is no point in repeating the lengthy proofs similar to those presented in paper I. The interested reader could get the computational details directly from the authors. We will simply state that the bivariant generating function associated with the number of distinct arrangements, $A(q;L = 1,M = 2,N)$, of q dimers on the $1\times2\times N$ lattice, is calculated to be

$$G(x, y) = (1 - xy)/D(x, y), \tag{14}$$

where $D(x, y)$ is a polynomial in x and y. Replacing y by $1/z$, one looks for the z root of

$$D(x,1/z) = y^3(z^3 - (x + 1)z^2 - x(x + 1)z + x^3). \tag{15}$$

The z roots of this cubic polynomial are easily found to be

$$(z_1)^2, \quad (z_2)^2, \quad \text{and} \quad -x, \tag{16}$$

where z_1 and z_2 are the values listed in Eq. (11). The first root listed above is the largest root.

In the case $M = 3$, the bivariant generating function associated with the number of arrangements $A(q;L = 1,M = 3,N)$ of q dimers parallel to the N axis on a $1\times3\times N$ lattice is

$$G(x, y) = (1 - x^3y^2)(1 - xy - x^3y^2)/D(x, y). \tag{17}$$

Here $D(x,1/z)$ is a sixth-order polynomial in z, namely,

$$\begin{aligned} D(x,1/z) = y^6(z^6 &- (x + 1)z^5 - x(2 + 5x + 3x^2)z^4 \\ &- x^2(1 + x - 2x^2)z^3 + x^4(2 + 5x + 3x^2)z^2 \\ &- x^6(x + 1)z - x^9). \end{aligned} \tag{18}$$

This polynomial has four real roots and two complex roots. The root of largest modulus is real and positive. We were able to obtain analytically the value of this root as well as another real root by making possible the following factorization:

$$\begin{aligned} D(x,1/z) = y^3(z - z_1{}^3)(z + z_2{}^3)(z^4 &- 2xz^3 - x(2x - 1)z^2 \\ &- 2x^4z + x^6). \end{aligned} \tag{19}$$

The largest root is (z_1) raised to the third power. There is no need of repeating the calculations for higher values of M. The pattern is very clear, the largest z root of the polynomial $D(x,1/z)$ is (z_1) raised to the M th power. Thus, in the case of vertical dimers, the exponential behavior of the largest root is an exact behavior. Without denying the intrinsic value of the mathematical proof, it is worth recognizing that this exponential behavior for vertical dimers can be predicted on physical grounds. Indeed, consider the grand canonical partition function of dimers on a $1\times M\times N$ square lattice with the restriction that all dimers are parallel to the N axis. Think of this lattice as made of M rows, each row containing N cells. Fitting dimers in any of the M rows should not affect the arrangements of the dimers on the other rows. In other words, we are dealing with independent probabilities. Therefore, the grand canonical partition function for large values of N should be the partition function for one row made of an infinite number of cells multiplied by itself M times, or,

$$\Delta_M(x) = (\Delta_1(x))^M. \tag{20}$$

We have already shown in paper I that the partition function for dimers of absolute activity x in one dimensions is $z_1(x)$. The exponential behavior of the z root follows. From this analysis, supported by the mathematical checking, it becomes clear that the thermodynamical properties of vertical dimers are the same as the thermodynamical properties of dimers in one dimension.

IV. THERMODYNAMICAL QUANTITIES

This section gives the values of the thermodynamical quantities of horizontal dimers and vertical dimers in the limit as N becomes infinitely large, while both L and M are fixed. We call $Z(x)$, $\Gamma(x)$, $\rho(x)$, and ρ_0 the partition function, the grand potential, the number density, and the close-packing density, respectively. If one site on the lattice is occupied by one end of a dimer, let Q be the number of neighboring sites that the other end of the dimer can occupy. In the case of aligned dimers, the value of Q would be 2. The general expressions of the thermodynamic quantities can all be calculated from the knowledge of the partition function (see, for example, Ref. 4); namely, one has

$$\Gamma(x) = (2/Q)\ln Z, \tag{21}$$

$$\rho(x) = x\frac{d\Gamma}{dx}. \tag{22}$$

In the limit of an infinitely large lattice, the number density at close packing is given in terms of Q as

$$\rho_0 = (1/Q). \tag{23}$$

The isothermal compressibility $K_T(x)$, the entropy per unit volume $S_V(x)$, and the constant-pressure specific-heat capacity per unit volume $C_V(x)$ will automatically follow:

$$K(x) = x\frac{d\rho}{dx} = k_B T\rho^2 K_T(x), \tag{24}$$

$$S(x) = -\rho\ln(x) + \Gamma = S_V/k_B, \tag{25}$$

$$C(x) = K(\Gamma/\rho)^2 = C_V/k_B. \tag{26}$$

In the above equations k_B is Boltzmann's constant, and T is the absolute temperature.

We now list the results for horizontal dimers on the $L\times M\times N$ lattice, where N is infinitely large, L can take on any positive value, and M is restricted to be larger than 1 to allow dimers parallel to the M axis to fit on the lattice. Following paper I, the partition function is the largest (and in this case the only) root computed is Sec. II raised to the power $(1/LM)$. This root is given by Eq. (12). It then follows that the partition function for horizontal dimers is given ex-

plicitly in terms of $z_1(x)$ and $z_2(x)$, Eq. (11), which are the roots of the quadratic equation, Eq. (10), namely,

$$Z(x) = \{(z_1^{M+1} - z_2^{M+1})/(z_1 - z_2)\}^{1/M}. \tag{27}$$

There is nothing to be gained in writing down the expressions of the various thermodynamic quantities explicitly in terms of x. However, we intend to discuss to some extent the close-packing situation. For this purpose we give the expression of the number density

$$\rho_H(x) = \frac{(M+1)x(z_1^M + z_2^M)}{M(1+4x)^{1/2}(z_1^{M+1} - z_2^{M+1})} - \frac{2x}{M(1+4x)}. \tag{28}$$

Index H refers to the number density for horizontal dimers. However, the number density at close packing is not exactly $1/Q = \frac{1}{2}$; it really depends on whether M is even or odd. Indeed when M is even, any row with M cells and full of dimers contains exactly $(M/2)$ dimers. But, if M is odd, any row reaching close packing would have one cell unoccupied; in this case the total number of dimers is $(M-1)/2$; consequently,

$$\rho_0 = \tfrac{1}{2}, \quad \text{for } M \text{ even}, \qquad \rho_0 = \tfrac{1}{2}(1 - 1/M), \quad \text{for } M \text{ odd}. \tag{29}$$

For vertical dimers on an $L \times M \times N$ lattice (dimers parallel to the N axis), we discovered that all the thermodynamical quantities are the same as those of dimers in one dimension. These quantities are calculated from $z_1(x)$. The partition function $Z(x)$ turns out to be precisely $z_1(x)$ (see paper I). Using Eq. (11), we obtain the number density

$$\rho(x) = (2/Q)(1 - 1/\sqrt{1+4x}). \tag{30}$$

For the case of vertical dimers $Q = (1/\rho_0) = 2$. Let θ be the normalized number density ρ/ρ_0. It is then possible to use Eq. (31) to express x explicitly in terms of θ, and one finds

$$x = \theta(2-\theta)/4(1-\theta)^2. \tag{31}$$

All thermodynamical quantities may now be calculated directly in terms of θ. The results are

$$\Gamma(\theta) = (2/Q)\ln(Z) = (2/Q)\ln((2-\theta)/(2-2\theta)), \tag{32}$$

$$S(\theta) = -\frac{\theta}{Q}\ln\frac{\theta(2-\theta)}{4(1-\theta)^2} + \frac{2}{Q}\ln\frac{2-\theta}{2-2\theta}, \tag{33}$$

$$K(\theta) = \theta(2-\theta)(1-\theta)/2Q, \tag{34}$$

$$C(\theta) = \frac{2(1-\theta)(2-\theta)}{Q\theta}\ln^2\left(\frac{2-\theta}{2(1-\theta)}\right). \tag{35}$$

By setting $x = 1$ in Eq. (31) one obtains the value of θ that maximizes the entropy, namely,

$$\theta_m = 1 - \frac{1}{\sqrt{5}} \quad \text{and} \quad S_{\max} = \frac{2}{Q}\ln\left(\frac{1+\sqrt{5}}{2}\right). \tag{36}$$

The quantity $K(\theta)$ related to the isothermal compressibility maximizes for $\theta = 1 - 1/\sqrt{3}$ and its value is

$$K_{\max} = 1/3Q\sqrt{3}. \tag{37}$$

Finally, the specific-heat capacity maximizes for the solution of the transcendental equation

$$\ln[(2-\theta)/2(1-\theta)] = 2\theta/(2-\theta^2), \tag{38a}$$

which is found to be

$$\theta = 0.844\,518\,622. \tag{38b}$$

From the knowledge of the entropy, one can also calculate the molecular freedom per dimer at close packing[4]

$$\phi = \exp(QS(\theta = 1)) = 1. \tag{39}$$

In Fig. 1, we plotted entropy curves for a system of dimers on an $L \times M \times N$ lattice parallel to the M axis, with N allowed to become infinitely large. These curves are labeled by the corresponding value of M and are, as predicted analytically, independent of the value of L (the number of $M \times N$ layers). Figures 2 and 3 give K and C for the same system. These plots are in terms of the normalized number density θ.

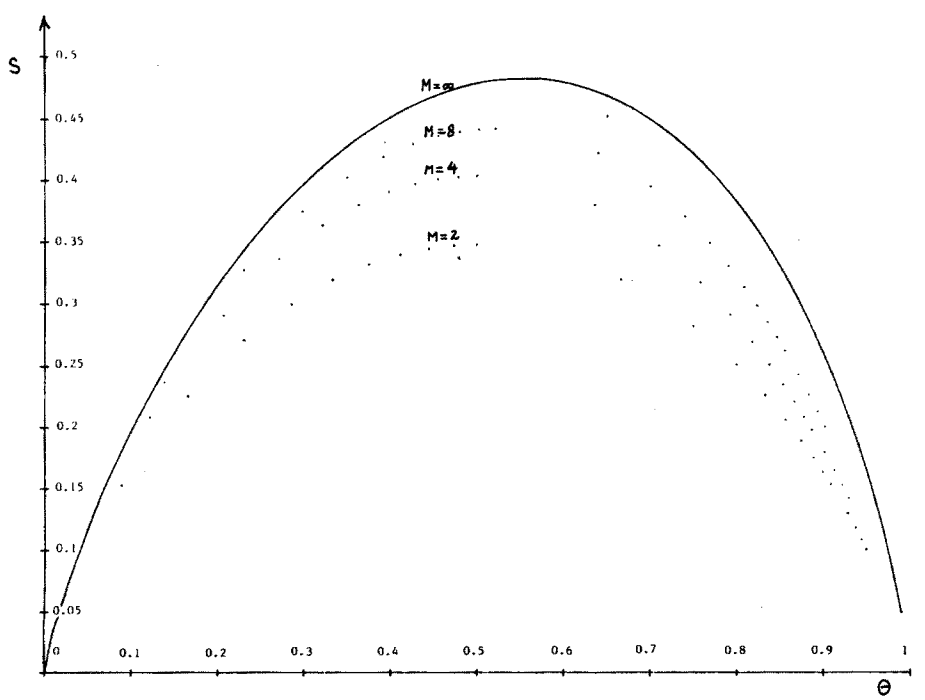


FIG. 1. Plot of S vs θ for horizontal dimers. Here S is (S_V/k_B), where S_V is the entropy per unit volume and k_B is Boltzmann's constant. θ is the normalized number density. The limiting curve labeled $M = \infty$ is the same for both vertical dimers on a finite lattice and horizontal dimers on the infinite lattice.

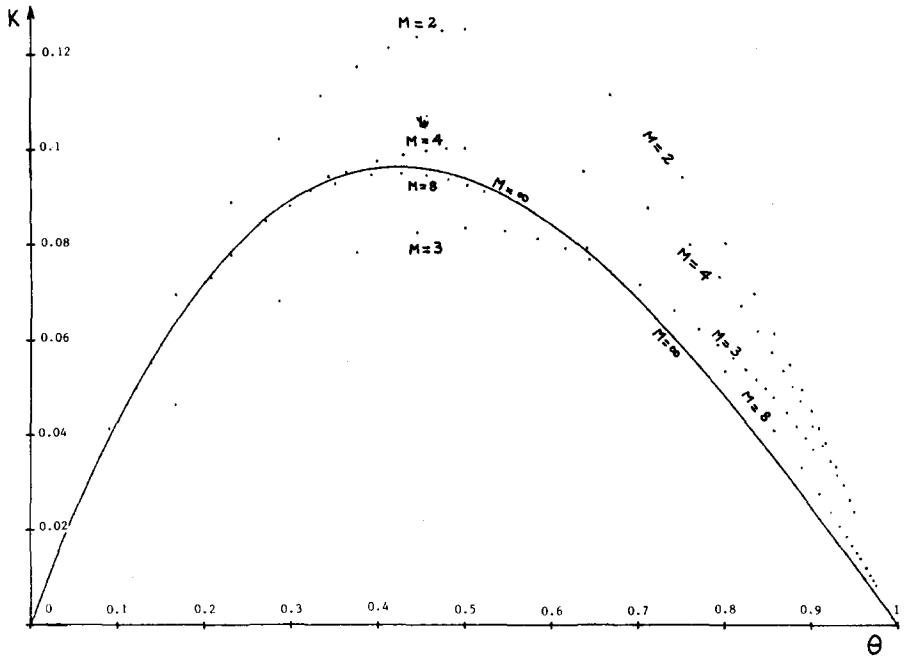


FIG. 2. Plot of K vs θ for horizontal dimers. Here K is $k_B T\rho^2 K_T$, where K_T is the isothermal compressibility per unit volume. θ is the normalized number density. The limiting curve labeled $M = \infty$ is the same for both vertical dimers on a finite lattice and for horizontal dimers on the infinite lattice.

One finds that these curves can be divided into two sets of curves, corresponding to even and odd values of M, respectively. In the limit as M becomes large, the limiting curve for "even" and "odd" series of curves is precisely the curve that one obtains for aligned dimers on the infinite lattice, which is also the curve calculated for the one-dimensional problem, as shown in Sec. III on vertical dimers. The limiting curve is represented by a continuous line.

We end this section with a discussion on close packing. We consider first the case of horizontal dimers. Near close-packing conditions we chose to look for the leading terms in the expression of the number density $\rho(x)$, given by Eq. (28). One finds that even and odd values of M lead to slightly different results, namely,

$$x \underset{\rho\to\rho_0}{\sim} A(M)/(1-(\rho/\rho_0))^{\gamma}, \tag{40}$$

where the critical exponent γ is exactly 2 and the amplitude $A(M)$ depends on the parity of M

$$A(M) = ((M+1)/2M)^2, \quad \text{for } M \text{ even},$$
$$A(M) = (M/2(M-1))^2, \quad \text{for } M \text{ odd}. \tag{41}$$

These results are to be compared with Eq. (31) for vertical dimers. The critical exponent is again 2 and the amplitude $\frac{1}{4}$ agrees with the value for horizontal dimers only for large values of M, as one might expect. This critical behavior was anticipated a long time ago based on various numerical approximations (see, for example, Ref. 4). We now have an analytical expression predicting this behavior without rely-

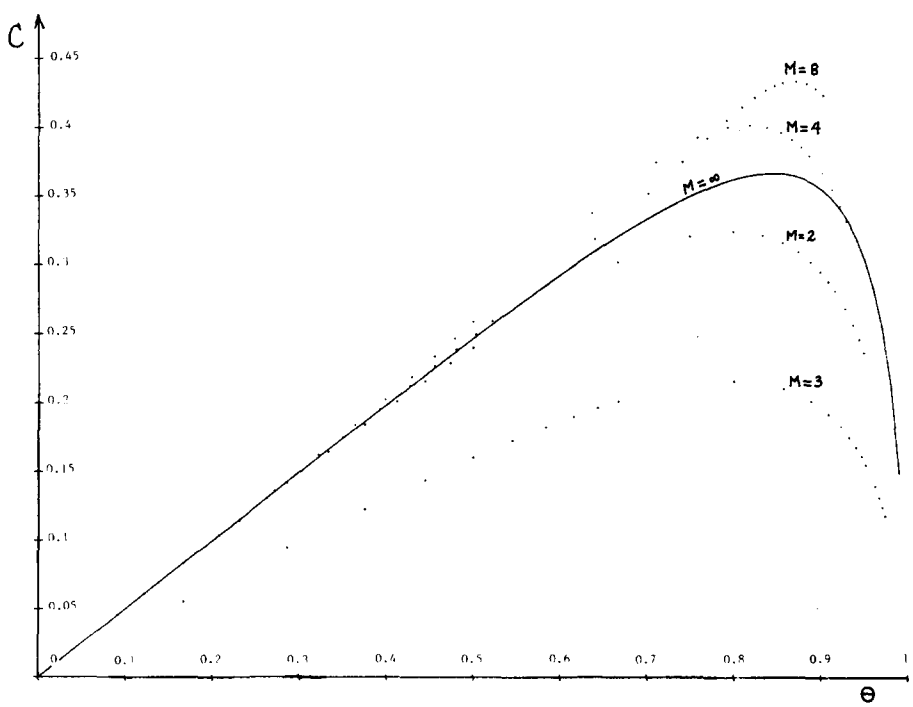


FIG 3. Plot of $C = C_V/k_B$ vs θ. Here C_V is the constant pressure specific-heat capacity per unit volume. θ is the normalized number density of dimers. The limiting curve labeled $M = \infty$ is the same for both vertical dimers on a finite lattice and for horizontal dimers on the infinite lattice.

ing on any approximation method for the case of aligned dimers.

V. MOLECULAR FREEDOM PER DIMER AT CLOSE PACKING

In the previous section, the molecular freedom per aligned dimer at close packing was (not surprisingly) found to be 1. Here, we intend to show that the molecular freedom per dimer at close packing Φ can be calculated from the knowledge of polynomial $P(x,z)$. Let z be the largest z root of $P(x,z)$ for an $L \times M \times N$ lattice. The partition function in the large N limit is given by[1]

$$Z(x) = (z(x))^{1/LM}. \tag{42}$$

One combines this equation and Eqs. (21) and (25) to obtain

$$S(x) = -\rho \ln(x) + (2/Q)\ln(z(x))^{1/LM}. \tag{43}$$

One multiplies Eq. (43) by Q, replaces $(Q\rho)$ by θ, the normalized number density, and exponentiates both sides of the resulting equation:

$$\exp(QS(x)) = (z(x))^{2/LM}/x^{\theta}. \tag{44}$$

In the close-packing limit, the absolute activity x becomes increasingly large, while the normalized number density θ approaches unity. In this limit, the left-hand side of Eq. (44) is identified as the molecular freedom per dimer at close packing, namely,

$$\Phi = \lim_{x\to\infty} [(z(x))^{2/LM}/x]. \tag{45}$$

In other words, the largest z root $z(x)$ of polynomial $P(x,z)$ should behave at large values of x like

$$z(x) = \underset{x\to\infty}{\sim} (\Phi x)^{LM/2}. \tag{46}$$

This observation allows one to calculate Φ in the following manner.

(1) Set $z = \lambda x^{LM/2}$ in the expression of the polynomial $P(x,z)$.

(2) Factor out x^{α}, where α is the highest exponent of x in the expression of $P(x,z)$, and obtain $V(\lambda,x)$ as

$$V(\lambda,x) = P(x,\lambda x^{LM/2})/x^{\alpha}. \tag{47}$$

(3) Take the limit of $V(\lambda,x)$ as x approaches infinity, namely,

$$V\lambda = \lim_{x\to\infty} V(\lambda,x). \tag{48}$$

(4) Find the roots of $V(\lambda)$, i.e., solve

$$V(\lambda) = 0. \tag{49}$$

(5) The largest root of $V(\lambda)$, say λ_1, gives the molecular freedom at close packing, namely,

$$\Phi = \lambda_1^{2/LM}. \tag{50}$$

We use this procedure to calculate Φ for the lattices mentioned in the Introduction: $1\times1\times N$, $1\times2\times N$, $1\times3\times N$, $1\times4\times N$, and $2\times2\times N$. The polynomials $P(x,z)$ for two-dimensional lattices are found in paper II (Table I). The polynomial $P(x,z)$ for the three-dimensional lattice $2\times2\times N$ was derived in Refs. 2 and 3. Listed in Table I are the expressions of $V(\lambda)$ for various lattices and their associated roots. As expected, we recover the results of Kasteleyn,[5] namely,

Lattice	$1\times1\times N$	$1\times2\times N$	$1\times3\times N$	$1\times4\times N$
Φ	1	$\frac{1}{2}(1+\sqrt{5})$	$(2+\sqrt{3})^{1/3}$	1.685 389 03

We also obtain a result not predicted by Kasteleyn's formula, the molecular freedom per dimer at close packing for the three-dimensional lattice $2\times2\times N$

$$\Phi = (2+\sqrt{3})^{1/2} \quad (2\times2\times N). \tag{51}$$

(See Table I.)

VI. GENERALIZATION

This section is concerned with the generalization of the results obtained for aligned dimers to the case of dimers on two- and three-dimensional lattices irrespective of their orientation. In performing this generalization, we are guided by the following facts.

(1) The partition function is given by Eq. (42).

(2) For increasing values of L and M, the root $z(x)$ has an almost exact exponential behavior in L and M. This behavior becomes more accurate with increasing values of L or M, as shown in the case of horizontal dimers.

TABLE I. This table gives a listing of polynomials $V(\lambda)$ for various lattices. The positive and negative roots of $V(\lambda)$ are obtained analytically and listed separately. The quantity μ stands for $\frac{1}{2}(1+\sqrt{5})$.

Lattice	Polynomial $V(\lambda)$	Roots	
$1\times1\times N$	λ^2-1	$+1$	-1
$1\times2\times N$	$\lambda^3-2\lambda^2+1$	$\mu=\frac{1}{2}(1+\sqrt{5}),1$	$\frac{1}{2}(1-\sqrt{5})$
$1\times3\times N$	$\lambda^6-6\lambda^5+7\lambda^4+8\lambda^3$	$2+\sqrt{3}$	-1
		$1+\sqrt{2}$	$1-\sqrt{2}$
	$-9\lambda^2-2\lambda+1$	1	
		$2-\sqrt{3}$	
$1\times4\times N$	$\lambda^{14}-2\lambda^{13}-10\lambda^{12}$	$\frac{1}{4}(1+\mu(\sqrt{6-\mu}+\sqrt{5+\mu})$	$\frac{1}{2}(\mu\sqrt{5+\mu})$,
	$+14\lambda^{11}+41\lambda^{10}-34\lambda^9$	$-\sqrt{5+\mu}(1-\sqrt{6-\mu}))$,	$\frac{1}{4}(1-\mu(-\sqrt{6-\mu}+\sqrt{5+\mu})$
	$-82\lambda^8+38\lambda^7+86\lambda^6$	$\frac{1}{2}(\mu+\sqrt{5+\mu}),1,1,1,$	$+\sqrt{5+\mu}(1-\sqrt{6-\mu}))$,
	$-20\lambda^5-47\lambda^4+4\lambda^3$	$\frac{1}{2}(1-\mu+\sqrt{6-\mu})$,	$\frac{1}{2}(1-\mu-\sqrt{6-\mu}),-1,-1,-1,$
	$+12\lambda^2-1$	$\frac{1}{4}(1-\mu(\sqrt{6-\mu}+\sqrt{5+\mu})$	$\frac{1}{4}(1-\mu\sqrt{6-\mu}-\sqrt{5+\mu})$
		$+\sqrt{5-\mu}(1+\sqrt{6-\mu}))$,	$-\sqrt{5+\mu}(1+\sqrt{6-\mu}))$,

TABLE II. Thermodynamical properties of dimers on a $2\times2\times N$. The values of θ, S, K, and C are calculated for selected values of the absolute activity x, using Eqs. (31) and (33)–(35).

X	θ	S	K	C
0.01	0.037 41	0.035 16	0.005 84	0.006 24
0.02	0.070 37	0.058 37	0.010 35	0.011 74
0.03	0.099 75	0.076 48	0.013 90	0.016 64
0.05	0.150 12	0.103 67	0.019 04	0.025 09
0.07	0.192 06	0.123 40	0.022 48	0.032 15
0.09	0.227 78	0.138 47	0.024 85	0.038 19
0.12	0.272 76	0.155 43	0.027 17	0.045 84
0.16	0.321 26	0.171 39	0.028 91	0.054 141
0.20	0.360 60	0.182 67	0.029 79	0.060 92
0.30	0.433 98	0.199 87	0.030 30	0.073 67
0.50	0.525 53	0.214 37	0.029 14	0.089 69
0.70	0.582 80	0.219 40	0.027 55	0.099 67
0.90	0.623 31	0.220 96	0.026 08	0.106 60
2.00	0.735 60	0.215 75	0.020 69	0.124 53
5.00	0.832 60	0.197 56	0.014 75	0.135 73
7.00	0.860 45	0.189 32	0.012 86	0.137 20
10.0	0.886 00	0.180 29	0.011 05	0.137 21
13.0	0.902 43	0.173 64	0.009 84	0.136 17
19.0	0.923 01	0.164 21	0.008 26	0.132 97

(3) The behavior at large values of x of the root $z(x)$, for a given lattice size, is given by Eq. (46), or, more precisely, by

$$z(x) \underset{x\to\infty}{\sim} (\Phi(L,M)x)^{LM/2}, \tag{52}$$

where the molecular freedom per dimer at close packing, referred to as $\Phi(L,M)$, depends on the lattice size. For two-dimensional lattices, Kasteleyn's formula gives[5]

$$\Phi(1,M) = \prod_{l=1}^{[\frac{1}{2}M]}\left(\cos\frac{l\pi}{M+1} + \left(1+\cos^2\frac{l\pi}{M+1}\right)^{1/2}\right) \tag{53}$$

and for the infinite two-dimensional lattice,[5]

$$\Phi(1,\infty) = \exp[2G/\pi] = 1.791\,622\,812\ldots, \tag{54}$$

where G is Catalan's constant. Several people computed the molecular freedom per dimer at close packing on the infinite three-dimensional lattice. Nagle's[7] value is

$$\Phi(\infty,\infty) = 2.4423. \tag{55}$$

As reported in the previous section, we computed exactly $\Phi(2,2)$ as

$$\Phi(2,2) = (2+\sqrt{3})^{1/2}. \tag{56}$$

Based on all these facts, it is safe to assume that the root $z(x)$ for an $L\times M\times N$ lattice is approximately given by

$$z(x) = (z_1(x))^{LM} = (\tfrac{1}{2}(1+\sqrt{1+4x}))^{LM}, \tag{57}$$

where $z_1(x)$ is the root for the one-dimensional lattice, namely, the one listed in Eq. (11). Unfortunately, Eq. (57) leads to a molecular freedom per dimer at close packing equal to 1. Clearly, this approximation is extremely bad for large values of x. One is able to improve on this by requiring Eq. (52) to be satisfied also. This is possible if one replaces Eq. (57) by

$$z(x) = [\tfrac{1}{2}(1+\sqrt{1+4\Phi(L,M)x})]^{LM}. \tag{58}$$

The above approximation becomes more accurate with increasing values of L or M. With this in mind, we compute all the thermodynamic quantities for dimers on an $L\times M\times N$ lattice ($N\to\infty$)

$$Z(x) = \tfrac{1}{2}(1+\sqrt{1+4\Phi(L,M)x}), \quad \Gamma(x) = (2/Q)\ln Z(x), \tag{59}$$

$$\rho(x) = (2/Q)(1 - 1/\sqrt{1+4\Phi(L,M)x}), \tag{60}$$

$$x(\theta) = \theta(2-\theta)/4\Phi(L,M)(1-\theta)^2, \tag{61}$$

$$S(\theta) = -\frac{\theta}{Q}\ln\frac{\theta(2-\theta)}{4\Phi(L,M)(1-\theta)} + \frac{2}{Q}\ln\frac{2-\theta}{2-2\theta}, \tag{62}$$

$$K(\theta) = \theta(2-\theta)(1-\theta)/2Q, \tag{63}$$

$$C(\theta) = \frac{2(1-\theta)(2-\theta)}{Q\theta}\ln^2\left(\frac{2-\theta}{2-2\theta}\right). \tag{64}$$

Here $Q = 2$, 4, or 6 for one-, two-, or three-dimensional lattices, respectively, all these closed-form analytic expressions approaching the exact thermodynamic quantities in the limit of the infinite one-dimensional, infinite two-dimensional, or infinite three-dimensional lattices. We note that $K(\theta)$ and $C(\theta)$ are independent of $\Phi(L,M)$; they maximize for the same conditions specified by Eqs. (37) and (38). However, this is not the case for $S(\theta)$. The value of θ that maximizes the entropy is given by

$$\theta_m = 1 - 1/\sqrt{1+4\Phi(L,M)}, \tag{65}$$

a relation whose accuracy increases with increasing values of L and M and which becomes exact for $(L=1, M=\infty)$ and $(L=\infty, M=\infty)$, i.e., for the infinite square lattice and the infinite simple cubic lattice, respectively. The case $L=1$ and $M=1$ is the infinite one-dimensional problem.

As a check of these mathematical expressions and their accuracy, we find them to agree with the exact numerical results derived in paper II for ($L=1$; $M=2$, 3 and 4) within 1.7% for the partition function Z. The partition function for ($L=1, M=2$) computed from the approximate relation (59) agrees within 1.5% with the exact values we compute numerically using the method of papers I and II.

We plotted in Figs. 4–6, $S(\theta)$, $K(\theta)$, and $C(\theta)$ for infinite one-, two-, and three-dimensional lattices (referred to as the 1D, 2D, and 3D curves, respectively). These curves are obtained from Eqs. (62)–(64). For comparison, we also plotted the points obtained by extrapolation in paper II for dimers on the infinite two-dimension lattice. In Fig. 4, the points obtained by extrapolation for $S(\theta)$ are astonishingly close to the exact analytical fit. This is consistent with the fact that (in paper II) two different methods of extrapolation lead to results agreeing within a 3% deviation. However, the extrapolation of paper II is not as good for $K(\theta)$ and $C(\theta)$ as made explicit in Figs. 5 and 6, respectively. The reason for this comes from the cumulative effect of errors in the numerical computation of K and C.

VII. SUMMARY CONCLUSION

Guided by the study of aligned dimers and the results previously obtained in papers I and II, we were able to obtain an analytical fit of the thermodynamical quantities of dimers on a rectangular $L\times M\times N$ lattice, where N is allowed to become infinite and L and M are fixed. The closed-form analytical expressions, Eqs. (59)–(64), are good to better than 1.7% error for low values of L and M, but are *exact* for the *infinite* one-, two-, and three-dimensional lattices. In addition, other major results are the following.

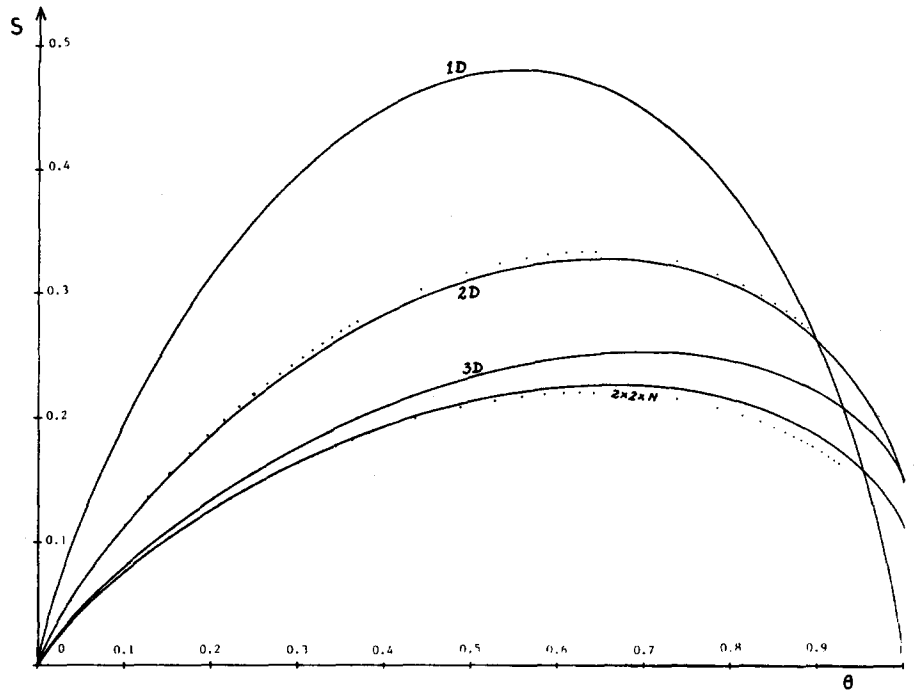


FIG. 4. Plot of S vs θ for dimers irrespective of their orientations. Curve 1D corresponds to the infinite one-dimensional case. Curve 2D correponds to the infinite two-dimensional case, and the data points nearby are those obtained by extrapolation in paper II. Curve 3D corresponds to the infinite three-dimensional case. Curve $2\times2\times N$ is the approximate analytical fit of the data points nearby for dimers on a $2\times2\times N$ lattice. These points are obtained exactly using the techniques of papers I and II.

(1) As exhibited by Eq. (61), the behavior of the absolute activity near close packing has a critical exponent

$$\gamma = 2,$$

with the associated amplitude

$$A = \frac{\theta(2-\theta)}{4\Phi(L,M)} \underset{\theta\to 1}{\longrightarrow} \frac{1}{4\Phi(L,M)}. \tag{66}$$

(2) Suprisingly enough, $K(\theta)$ and $C(\theta)$ are independent of the molecular freedom of dimers at close packing. The value of the occupation density θ for which K is maximum is $\theta = 1 - 1/\sqrt{3}$, and $K_{\max} = 1/12\sqrt{3}$ for the infinite square lattice and $1/18\sqrt{3}$ for the infinite cubic lattice. The value of θ for which C is maximum is $\theta = 0.844\,518\,622$. Finally, the value of C at the maximum is 0.183 259 44 for the infinite square lattice and 0.122 172 96 for the infinite cubic lattice.

(3) The entropy per unit volume divided by Boltzmann's constant $S(\theta)$ maximizes at $\theta = 1 - 1/\sqrt{5} = 0.552\,786\,404$ for the infinite one-dimensional lattice, at $\theta = 0.650\,069\,136$

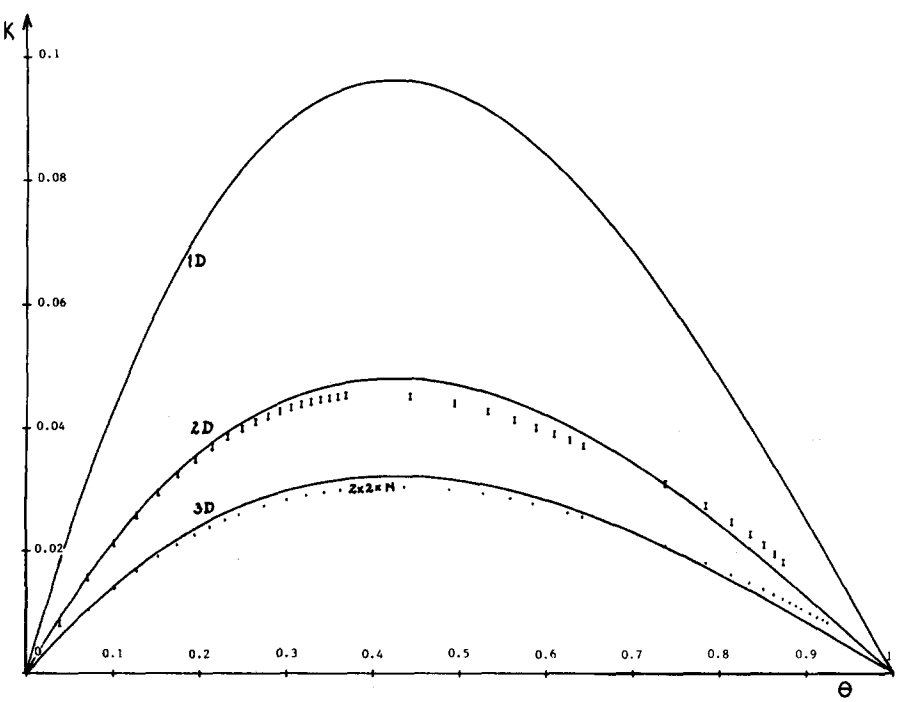


FIG. 5. Plot of K vs θ for dimers irrespective of their orientations. Curve 1D corresponds to the infinite one-dimensional case. Curve 2D corresponds to the infinite two-dimensional case, and the data points nearby are those obtained by extrapolation in paper II. Curve 3D corresponds to the infinite three-dimensional case. Curve $2\times2\times N$ is the approximate analytical fit of the data points nearby for dimers on a $2\times2\times N$ lattice. These points are obtained exactly using the techniques of papers I and II.

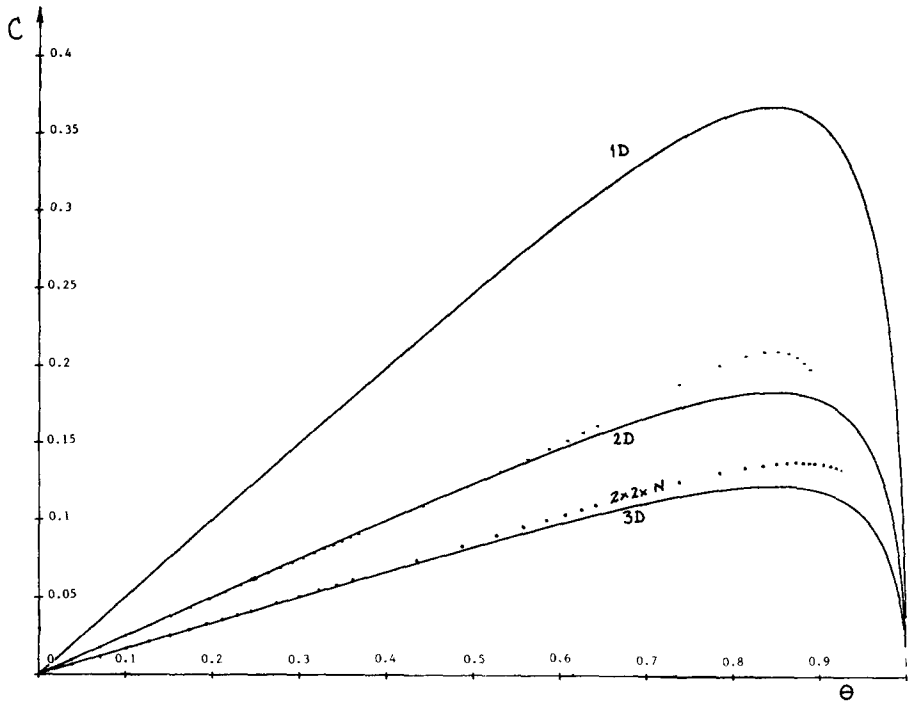


FIG. 6. Plot of C vs θ for dimers irrespective of their orientations. Curve 1D corresponds to the infinite one-dimensional case. Curve 2D corresponds to the infinite two-dimensional case, and the data points nearby are those obtained by extrapolation in paper II. Curve 3D corresponds to the infinite three-dimensional case. Curve $2\times2\times N$ is the approximate analytical fit of the data points nearby for dimers on a $2\times2\times N$ lattice. These points are obtained exactly using the techniques of papers I and II.

for the infinite square lattice, and at $\theta = 0.695\,27$ for the infinite cubic lattice. The corresponding values of S are 0.481 211 826, 0.328 462 937, 0.253 73, respectively. These results are to be compared with previous numerical approximations.[1,4]

(4) We developed a new method for calculating molecular freedom of dimers on three-dimensional lattices. However, at this stage, it does not give a closed-form expression similar to Kasteleyn's formula on two-dimensional lattices.

ACKNOWLEDGMENTS

We are indebted to Dr. Donald E. Shaw for his help in various aspects of this work. We are also grateful to David J. Hutnick for developing the computer software we used for plotting of the data using the IBM-PC and the 7470A-HP plotter.

[1] A. J. Phares, J. Math. Phys. **25**, 1756 (1984); A. J. Phares, D. E. Shaw, and F. J. Wunderlich, *ibid.* **26**, 1762 (1985).
[2] D. Lichtman and R. B. McQuistan, J. Math. Phys. **8**, 2441 (1967); R. B. McQuistan, Nuovo Cimento B **58**, 86 (1968); R. B. McQuistan and S. J. Lichtman, J. Math. Phys. **11**, 3095 (1970); R. B. McQuistan and J. L. Hock, Fibonacci Quart. **21**, 196 (August 1983).
[3] A. J. Phares, J. Math. Phys. **25**, 2169 (1984).
[4] D. S. Gaunt, Phys. Rev. **179**, 174 (1969).
[5] P. W. Kasteleyn, Physica **27**, 1209 (1961).
[6] J. L. Hock and R. B. McQuistan, J. Math. Phys. **24**, 1859 (1983).
[7] J. F. Nagle, Phys. Rev. **152**, 190 (1966).

Solution of the eigenvalue problem of an integral equation with the help of its associated differential equation applied to the calculation of diffraction losses in confocal resonators

H. J. W. Müller-Kirsten and E. Sattel
Department of Physics, University of Kaiserslautern, 6750 Kaiserslautern, West Germany

(Received 17 January 1985; accepted for publication 2 May 1985)

A problem in optics is chosen in order to develop a general method for solving the eigenvalue problem of a homogeneous integral equation with the help of the eigenfunctions of the associated differential equation. The problem chosen is that of the modes of optical resonators with circular, confocal mirrors which are given by the solutions of a homogeneous Fredholm integral equation which can be derived from Kirchhoff 's diffraction formula. This integral equation can be converted into a hyperspheroidal differential equation supplemented by appropriate boundary conditions. The solutions and eigenvalues of this differential equation are studied in detail for both small and large values of a parameter called the Fresnel number. These eigenfunctions are then used for the computation of the eigenvalues of the original integral equation which measure the diffraction loss in the resonator. Throughout the same general perturbation method is used, and our emphasis is on the solution of the general problem of solving the eigenvalue problem of the homogeneous integral equation together with that of its related differential equation.

I. INTRODUCTION

Our objective in the following is to solve the eigenvalue problem of a homogeneous integral equation with the help of the solution of the eigenvalue problem of the corresponding differential equation. The particular system of equations that we choose is of relevance in the study of finite-dimensional optical resonators, although the method of solution as such is applicable to many other cases. We begin, therefore, by recapitulating the physical motivation of these equations.

The reflection of light in optical resonators is accompanied by diffraction at the edges of the finite-size mirrors. Multiple reflection then leads to a homogeneous, locally coherent field distribution in the space between the mirrors. The transverse eigenmodes of this field are called the modes of the resonator.

In the following we consider a resonator consisting of two identical confocal spherical mirrors. The segment radius of the identical mirrors is a and the distance between their centers is $L = b$, the length of the resonator, as shown in Fig. 1. The foci of the two mirrors coincide since for small spherical mirrors the focal length is half the radius of curvature.

We assume an arbitrary field on one of the mirrors and determine the field on the other mirror after its first passage through the resonator. The latter is then used to determine the field on the first mirror after the second passage and so on. After a sufficient number of passages, say q, the field distribution u_q stabilizes. We can use Kirchhoff's formula in order to express u_{q+1} in terms of u_q. Thus

$$u_{q+1} = \int_A u_q K\, ds, \quad K \equiv \frac{ik}{4\pi R}(1+\cos\theta)e^{-ikR}. \tag{1}$$

Here $k = 2\pi/\lambda$, where λ is the wavelength, R is the distance of a point on one mirror from one on the other mirror, θ is the angle between R and the normal to the latter, and A is the area of the mirror. After a sufficient number of passages the field distribution on the mirrors stabilizes, and the field on one differs from that on the other only by a factor measuring the loss of intensity, i.e.,

$$u_q = (1/\gamma)^q v. \tag{2}$$

Substituting this into (1) we obtain

$$v = \gamma \int_A Kv\, ds, \tag{3}$$

where v describes the modes of the resonator and γ determines both the damping and phase shift of the wave as it passes through the resonator. For the evaluation of the integral we assume radiation consisting of plane waves that propagate parallel to the optical axis. Then

$$u_{q+1}(r_2,\varphi_2) = \frac{ik}{4\pi}\int_0^a \int_0^{2\pi} u_q(r_1,\varphi_1) \times \frac{e^{-ikR}}{R}(1+\cos\theta) r_1\, dr_1\, d\varphi_1, \tag{4}$$

where $\cos\theta = b_1/R$,

$$R = [b_1^2 + r_1^2 + r_2^2 - 2r_1 r_2 \cos(\varphi_1 - \varphi_2)]^{1/2},$$
$$b_1 = b - \Delta_1 - \Delta_2, \tag{5}$$
$$\Delta_i = b - (b^2 - r_i^2)^{1/2}, \quad i = 1,2.$$

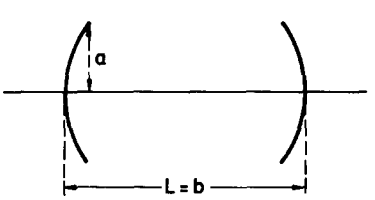


FIG. 1. Arrangement and radii of confocal mirrors.

 0022-2488/85/102500-10$02.50

For b/a large, we have $\Delta_i \approx r_i^2/2b$ and $b_1 \approx R \approx b$ so that $(1 + b_1/R) \approx 2$ and

$$R \approx b - (r_1 r_2/b)\cos(\varphi_1 - \varphi_2).$$

Then

$$v(r_2,\varphi_2) = \gamma \int_0^a \int_0^{2\pi} K(r_2,\varphi_2;r_1,\varphi_1) r_1 \, dr_1 \, d\varphi_1, \tag{6}$$

where (absorbing e^{-ikb} in γ)

$$K(r_2,\varphi_2;r_1,\varphi_1) = (ik/2\pi b) e^{ik(r_1 r_2/b)\cos(\varphi_1 - \varphi_2)}. \tag{7}$$

The solutions of (6) have the form

$$v(r,\varphi) = S_n(r) e^{-in\varphi}, \tag{8}$$

where n is an integer. Then

$$S_n(r_2) = \gamma \int_0^a K_n(r_2,r_1) S_n(r_1) r_1 \, dr_1, \tag{9}$$

where

$$K_n(r_1,r_2) = (i^{n+1} k/b) J_n(kr_1 r_2/b).$$

Equation (9) is a homogeneous linear integral equation of the second kind with a linear and continuous kernel. The eigenfunctions associated with different eigenvalues are orthogonal in the interval $(0,a)$, i.e.,

$$\int_0^a S_{nl}(r) S_{nm}(r) r \, dr = 0 \quad (l \neq m). \tag{10}$$

Hence, correspondingly

$$\int_0^{2\pi} \int_0^a v_{nm}(r,\varphi) v_{kl}(r,\varphi) r \, dr \, d\varphi = 0,$$

if either $n \neq k$ or $m \neq 1$.

The set of all eigenfunctions $S_{nm}(r)$ describes the field and hence its intensity at the mirrors. Here, n describes the order in the azimuthal direction and m that in the radial direction. It is customary[1,2] to write (9) in the form (setting $r_1 = ay$, $r_2 = ax$, $\gamma_{nm} = i^{n+1}\gamma$)

$$x^{1/2} S_{nm}(c,x) = \gamma_{nm} c \int_0^1 y^{1/2} S_{nm}(c,y) J_n(cxy)(xy)^{1/2} \, dy, \tag{11}$$

and $c = ka^2/L$, or more precisely, $c/2\pi$ is called the Fresnel number.

The existence of solutions of Eq. (3) has been investigated by several authors[3–5] for resonators consisting of plane-parallel or spherical mirrors.

An important quantity is the diffraction loss P_{nm} defined by

$$P_{nm} = 1 - |\gamma_{nm}|^{-2}. \tag{12}$$

In the literature, e.g., in Ref. 6, P_{nm} is sometimes defined as $1 - |\gamma_{nm}|^2$; in this case the eigenvalue γ_{nm} is, of course, a factor on the other side of the integral equation. It is important to observe that the diffraction loss does not depend so much on the specific geometry of the resonator as on the parameter

$$c = ka^2/L = 2\pi N, \quad N = a^2/L\lambda, \tag{13}$$

where N is called the Fresnel number. Resonators having the same Fresnel number are similar as far as their diffraction losses are concerned.

Equation (11) is a Fredholm integral equation. As such it can be shown to be equivalent to the following differential equation:

$$\left\{ \frac{d}{dx}(1 - x^2)\frac{d}{dx} + \Gamma_{nm} - c^2 x^2 - \frac{n^2 - \frac{1}{4}}{x^2} \right\} \times x^{1/2} S_{nm}(c,x) = 0, \tag{14}$$

together with the boundary conditions

$$S_{nm}(x), \quad \frac{d}{dx} S_{nm}(x) \quad \text{finite at } x = 0 \text{ and } 1. \tag{15}$$

The formal integral equation $\psi = \int I\psi \, dx$ is compatible with the differential equation $D\psi = E\psi$ if the integral operator I commutes with the differential operator D, i.e., $[D,I] = 0$. It can be shown that this is the case for (11) and (14) provided the conditions (15) are satisfied. Equation (14) is a generalized form of the prolate spheroidal wave equation.[2,7,8]

In the following we use the perturbation method explained, for instance, in Ref. 9 and applied to various types of differential equations in Refs. 10–14 for the perturbative solution of the hyperspheroidal differential equation (14). In Sec. II we derive expansions for the solutions and eigenvalues for small values of the Fresnel parameter c and in Sec. III for large values of c. In particular, we point out certain characteristics of these expansions which are similar to those of corresponding expansions for the solutions of Mathieu's equation.[8,11] In Sec. IV we present numerical calculations of the eigenvalues and demonstrate how both types of expansions supplement each other beyond their immediate individual domains of validity. In Sec. V we use the eigenfunctions of Sec. III for the calculation of the eigenvalues of the integral equation in the domain of small Fresnel parameters c, and in Sec. VI we use the eigenfunctions of Sec. IV for the calculation of these eigenvalues in the domain of large values of c.

II. SOLUTION OF THE HYPERSPHEROIDAL DIFFERENTIAL EQUATION FOR SMALL VALUES OF THE FRESNEL NUMBER

The approximate behavior of the eigenvalues Γ_{nm} of (14) can be found by considering a domain of the variable x in which the solutions can be approximated by normalized functions. Thus, setting

$$\begin{aligned} &x^{1/2} S(x) = x^{n+1/2} P(x), \\ &2x^2 - 1 = y, \\ &k = \tfrac{1}{2}(m - n), \quad k = 0,1,2,..., \end{aligned} \tag{16}$$

so that

$$k(k + n + 1) = \tfrac{1}{4}[-n(n+2) + m(m+2)], \tag{17}$$

Eq. (14) becomes

$$\begin{aligned} &(1 - y^2)P''(y) + [n - (n+2)y]P'(y) \\ &\quad + [k(k + n + 1) - \tfrac{1}{4}m(m+2) + \tfrac{1}{4}(\Gamma - \tfrac{3}{4})]P(y) \\ &\quad = (c^2/8)(y + 1)P(y). \end{aligned} \tag{18}$$

Comparing this with the differential equation of Jacobi polynomials $P_k^{(\alpha,\beta)}$, i.e.,

$$\begin{aligned} &(1 - y^2)P_k^{(\alpha,\beta)\prime\prime}(y) + [\beta - \alpha - (\alpha + \beta + 2)y]P_k^{(\alpha,\beta)\prime}(y) \\ &\quad + k(k + \alpha + \beta + 1)P_k^{(\alpha,\beta)}(y) = 0, \end{aligned} \tag{19}$$

for $\alpha = 0$ and $\beta = n$, we obtain

$$\Gamma_{nm} - \tfrac{3}{4} \sim m(m+2).$$

Thus

$$\Gamma_{nm} - \tfrac{3}{4} = m(m+2) - c^2\Delta, \tag{20}$$

where $\Delta = O(c^0)$, and

$$S_{nm}(x) = x^n P_k^{(0,n)}(2x^2 - 1) + O(c^2). \tag{21}$$

We can now rewrite (18) in the form

$$D_k^{(n)} P(y) = (c^2/8)(2\Delta + y + 1)P(y), \tag{22}$$

where

$$D_k^{(n)}(y) = (1-y^2)\frac{d^2}{dy^2} + [n - (n+2)y]\frac{d}{dy} + k(k+1). \tag{23}$$

Using the recurrence relation of Jacobi polynomials, we have

$$\begin{aligned} R_k^{(0)} &= (c^2/8)(2\Delta + y + 1)P_k^{(0,n)}(y) \\ &= (c^2/8)\{[k,k+1]P_{k+1}^{(0,n)} + [k,k]P_k^{(0,n)} \\ &\quad + [k,k-1]P_{k-1}^{(0,n)}\}, \end{aligned} \tag{24}$$

where

$$[k,k+1] = \frac{2(k+1)(k+n+1)}{(2k+n+1)(2k+n+2)},$$

$$[k,k] = \frac{n^2}{(2k+n)(2k+n+2)} + 1 + 2\Delta,$$

$$[k,k-1] = \frac{2k(k+n)}{(2k+n)(2k+n+1)} \quad (k,n \text{ not both } 0). \tag{25}$$

Now,

$$D_{k+d}^{(n)} = D_k^{(n)} + d(2k+d+n+1)$$

and

$$D_{k+d}^{(n)} P_{k+d}^{(0,n)} = 0,$$

so that

$$D_k^{(n)}(P_{k+d}^{(0,n)}/[-d(2k+d+n+1)]) = P_{k+d}^{(0,n)}. \tag{26}$$

Thus, in order to remove a term $aP_{k+d}^{(0,n)}$, $d \neq 0$ from the right-hand side of (22) and hence from (24), a term $\{a/[-d(2k+n+1+d)]\}P_{k+d}^{(0,n)}$ must be added to $P^{(0)}$. To the same order of approximation the coefficient of $P_k^{(0,n)}$ in $R_k^{(0)}$ must be equated to zero and yields the appropriate value of Δ. This type of perturbation method has been discussed extensively in previous applications; we therefore refer to the literature, Refs. 9–14, for further details. Calculating in this way contributions up to and including those of the fourth order, we obtain

$$\begin{aligned} \Gamma_{nm}(c) &= \frac{3}{4} + m(m+2) + \frac{c^2}{2}\left(\frac{n^2}{m(m+2)} + 1\right) + \frac{c^4}{2^6}\left(-\frac{(m-n+2)^2(m+n+2)^2}{(m+2)^3(m+1)(m+3)} + \frac{(m-n)^2(m+n)^2}{m^3(m+1)(m-1)}\right) \\ &\quad + \frac{c^6}{2^7}\left\{\frac{n^2}{m(m+2)(m+1)}\left(-\frac{(m-n+2)^2(m+n+2)^2}{(m+2)^4(m+3)(m+4)} + \frac{(m-n)^2(m+n)^2}{m^4(m-1)(m-2)}\right)\right\} \\ &\quad + \frac{c^8}{2^{10}}\left\{-\frac{(m-n+2)^2(m+n+2)^2}{(m+2)^3(m+1)(m+3)}\left(\frac{1}{2^5}\frac{(m-n+4)^2(m+n+4)^2}{(m+3)^2(m+2)(m+4)^2(m+5)} + \frac{4n^2}{(m+2)^4m^2(m+4)^2}\right)\right. \\ &\quad + \frac{(m-n)^2(m+n)^2}{m^3(m-1)(m+1)}\left(\frac{1}{2^5}\frac{(m-n-2)^2(m+n-2)^2}{(m-1)^2m(m-3)(m-2)^2} + \frac{4n^2}{m^4(m-2)^2(m+2)^2}\right) \\ &\quad - \frac{1}{(m+1)4^2}\left(\frac{(m-n+2)^2(m+n+2)^2}{(m+2)^4(m+3)} + \frac{(m-n)^2(m+n)^2}{(m-1)m^4}\right) \\ &\quad \left.\times\left(-\frac{(m-n+2)^2(m+n+2)^2}{(m+2)^3(m+1)(m+3)} + \frac{(m-n)^2(m+n)^2}{(m+1)m^3(m-1)}\right)\right\} + O(c^{10}) \end{aligned} \tag{27}$$

and

$$\begin{aligned} P(y) &= P^{(0)} + P^{(1)} + P^{(2)} + \cdots \\ &= P_k^{(0,n)}(y) + \frac{c^2}{2^4}\left\{-\frac{(m-n+2)(m+n+2)}{(m+1)(m+2)^2}P_{k+1}^{(0,n)}(y) + \frac{(m-n)(m+n)}{m^2(m+2)}P_{k-1}^{(0,n)}(y)\right\} + \frac{c^4}{2^9}\left\{\frac{(m-n+2)(m+n+2)}{(m+2)^2(m+1)}\right. \\ &\quad \times\left(\frac{(m-n+4)(m+n+4)}{(m+3)^2(m+4)}P_{k+2}^{(0,n)}(y) - \frac{16n^2}{m(m+2)^2(m+4)}P_{k+1}^{(0,n)}\right) + \frac{(m-n)(m+n)}{m^2(m+1)}\left(\frac{16n^2}{(m-2)m^2(m+2)}P_{k-1}^{(0,n)}(y)\right. \\ &\quad \left.\left.\left. + \frac{(m-n-2)(m+n-2)}{(m-2)(m-1)^2}P_{k-2}^{(0,n)}(y)\right)\right\} + O(c^6). \end{aligned} \tag{28}$$

If, in evaluating these expressions, terms $P_l^{(0,n)}$ arise with $l < 0$, these have to be put equal to zero since Jacobi polynomials are defined only for $l \geqslant 0$. It might seem that (28) becomes singular and hence undefined for certain values of m. However, this is not the case. As an example we consider

$$\frac{(m-n-2)(m+n-2)}{(m-2)(m-1)^2}P_{k-2}^{(0,n)}.$$

This term arises only if $k = (m-n)/2 \geqslant 2$, i.e., $m \geqslant 4+n$, $n = 0, 1, 2, \ldots$. In case $k = 2$ we have

$$\frac{1}{(m-2)(m-1)^2}P_0^{(0,n)} = \frac{P_0^{(0,n)}}{(n+2)(n+3)^2},$$

and there is no problem.

Looking at (27) we observe that the higher-order coefficients of this expansion are singular for integral values of m.

Thus the coefficient of c^4 contains $1/(m^2-1)$, that of c^6 contains the factors $1/(m^2-1)$, $1/(m^2-2)$, and so on. This aspect has been ignored by Heurtley[1] and others, but has its analogy in the study of similar solutions of Mathieu's equation.[8,11] A simple justifiable recipe for the evaluation of the expansion in these cases says the following: if such a singular term is encountered, it has to be ignored, the resulting expansion being the same as that calculated specifically for the singular value of m. As an example we consider the case $n=m=1$. Since $P_0^{(0,1)}=1$ and

$$P_1^{(\alpha,\beta)}(y)=\tfrac{1}{2}[(\alpha-\beta)+y(\alpha+\beta+2)],$$

we have

$$P_1^{(0,1)}=\tfrac{1}{2}(3y-1),$$

or

$$y=\tfrac{1}{3}[2P_1^{(0,1)}-P_0^{(0,1)}]. \tag{29}$$

Hence (24) yields

$$R_0^{(0)}=(c^2/8)(2\Delta+y+1)P_0^{(0,1)}$$
$$=(c^2/8)\{(2\Delta+\tfrac{4}{3})P_0^{(0,1)}-\tfrac{2}{3}P_1^{(0,1)}\}. \tag{30}$$

Proceeding now in our standard way[9–14] we obtain

$$\Gamma_{11}(c)=3+\tfrac{3}{4}+\tfrac{2}{3}c^2-6^{-3}c^4+O(c^6). \tag{31}$$

As a further example we write down explicitly the above expansion for the fundamental mode:

$$S_{00}(x)=P_0^{(0,0)}(y)-(c^2/16)P_1^{(0,0)}(y)$$
$$+(c^4/1152)P_2^{(0,0)}(y)+O(c^6), \tag{32}$$

in agreement with the result of Heurtley.[1]

So far we have derived only one solution of (22) and hence of the original differential equation. A second, linearly independent solution can be obtained by choosing for the zeroth-order approximation that solution of Jacobi's equation which together with the Jacobi polynomial forms a linearly independent pair. Alternatively, one can exploit the symmetry properties of (22) and obtain this second solution by replacing in the first m by $-m-2$ and c by $-c$. It is easily verified that these replacements leave the eigenvalue expansion unchanged.

Our expansions for small values of the Fresnel number c are certainly convergent in a domain $|c|<\text{const}\neq 0$. This convergence can be proved with the help of a theorem stated by Meixner and Schäfke.[8] The determination of the exact radius of convergence is complicated and is therefore not considered here in more detail. However, our expansions can also be used in the sense of asymptotic expansions beyond this radius of convergence. In this connection we observe that for $m^2\to\infty$, c finite,

$$\Gamma_{nm}(c)\to m(m+2)+\tfrac{3}{4}+c^2/2, \tag{33}$$

and it can, in fact, be shown that our expansions are asymptotic expansions for large values of m.

Finally, we write out the particular solution of (14) for small values of c and valid in the domain $0\leqslant x\leqslant 1$, which satisfies the boundary conditions (15). This is (apart from a normalization constant)

$$S_{nm}(x)=x^n\Big\{P_k^{(0,n)}(2x^2-1)+\sum_{j=1}^{\infty}\left(\frac{c^2}{8}\right)^j$$
$$\times\sum_{\substack{d=-j\\ d\neq 0}}^{j}\frac{[k,k+d]_j}{d(2k+n+1+d)}$$
$$\times P_{k+d}^{(0,n)}(2x^2-1)\Big\}, \tag{34}$$

where $[k,k+d]_j$ are the appropriate coefficients and $k=\frac{1}{2}(m-n)$.

Heurtley[1] obtained solutions of the hyperspheroidal wave equation in terms of the Zernike polynomials $R_m^n(x)$, which are related to Jacobi polynomials; thus

$$R_m^n(x)=x^nP_{(1/2)(m-n)}^{(0,n)}(2x^2-1).$$

Slepian[2] obtained expansions of the solutions in terms of polynomials

$$R_{N,n}(x)=F(-n,n+N+1;N+1;x).$$

These are related to Jacobi polynomials by the relation

$$R_{N,\tilde{n}}(x)=\binom{\tilde{n}-N}{\tilde{n}}^{-1}P_{\tilde{n}}^{(0,N)}(2x^2-1).$$

Again we find agreement with our results by observing that

$$R_{N,\tilde{n}}(x)=x^{1/2}S_{nm}(x),$$

with

$$N=n,\quad \tilde{n}=\tfrac{1}{2}(m-n).$$

III. SOLUTION OF THE HYPERSPHEROIDAL DIFFERENTIAL EQUATION FOR LARGE VALUES OF THE FRESNEL NUMBER

As before, we begin with the determination of the approximate behavior of the eigenvalues $\Gamma=\Gamma_{nm}(c)$ of (14) for large values of c. To this end we convert (14) to its normal form by setting

$$x^{1/2}S(x)=(1-x^2)^{-1/2}Y(x), \tag{35}$$

so that

$$Y''(x)+\{[\Gamma-c^2x^2-(n^2-\tfrac{1}{4})/x^2+1]/(1-x^2)$$
$$-x^2/(1-x^2)^2\}Y(x)=0. \tag{36}$$

Expanding the denominators of the coefficient of $Y(x)$ in rising powers of x, setting $z=x(2c)^{1/2}$ and considering $c\to\infty$, we obtain

$$\ddot{Y}(z)+\{A-l(l+1)/z^2-\tfrac{1}{4}z^2\}Y(z)=0,$$

where

$$A=(1/2c)(\Gamma+\tfrac{5}{4}-n^2) \tag{37}$$

and

$$l(l+1)=n^2-\tfrac{1}{4}. \tag{38}$$

The solution

$$Y(z)=z^{l+1}e^{-(1/4)z^2}\Phi(a,b;\tfrac{1}{2}z^2), \tag{39}$$

where

$$a=-\tfrac{1}{2}[A-(l+\tfrac{3}{2})],\quad b=l+\tfrac{3}{2}, \tag{40}$$

and Φ is the confluent hypergeometric function, exists for $z\to\infty$ and hence is normalized provided

$$a=-k,\quad k=0,1,2,\ldots. \tag{41}$$

We define a number m by the relation

$$k = \tfrac{1}{2}(m - n). \tag{42}$$

The behavior of the eigenvalues now follows from (41) together with (40) and (42). Thus

$$\Gamma_{nm}(c) = 2c(m+1) + n^2 - \tfrac{5}{4} + \Delta, \tag{43}$$

where $\Delta = O(c^0)$.

We now use the same general iteration technique in order to derive large $-c$ asymptotic expansions for the solutions and eigenvalues of the hyperspheroidal equation. Setting

$$S(x) = x^n e^{-(1/2)cx^2} T(x) \tag{44}$$

in (14), changing the variable to

$$z = cx^2, \tag{45}$$

and replacing Γ by (43), the equation can be written as

$$D_k^{(n)} T(z) = (1/4c)\{4z^2 \ddot{T}(z) + [4z(n+2) - 4z^2]\dot{T}(z) + [2(n+1) - \Delta - 2z(n+2) + z^2]T(z)\}, \tag{46}$$

where

$$D_k^{(n)} \equiv z\frac{d^2}{dz^2} + (n+1-z)\frac{d}{dz} + k \tag{47}$$

and

$$k = \tfrac{1}{2}(m - n).$$

To zeroth order the solution $T = T^{(0)}$ of (46) is a generalized Laguerre function, i.e.,

$$T^{(0)} = L_k^n(z).$$

Using well-known recurrence relations for these functions, we can rewrite the right-hand side of (46) for $T = T^{(0)}$ as

$$R_k^{(0)} = (1/4c)\{(k,k)L_k^n + (k,k-2)L_{k-2}^n + (k,k+2)L_{k+2}^n\}, \tag{48}$$

where

$$\begin{aligned}(k,k) &= -2k(k+n+1) - n(n+1) - \Delta,\\ (k,k-2) &= (n+k)(n+k-1),\\ (k,k+2) &= (k+1)(k+2).\end{aligned} \tag{49}$$

Since

$$D_{k+d}^{(n)} = D_k^{(n)} + d \tag{50}$$

and

$$D_{k+d}^{(n)} L_{k+d}^n = 0,$$

we have

$$D_k^{(n)}(L_{k+d}^n/(-d)) = L_{k+d}^n.$$

Thus a term $aL_{k+d}^n/d \neq 0$, on the right-hand side of (46) can be removed by adding to $T^{(0)}$ the new contribution $-(a/d)L_{k+d}^n$. The equation determining Δ and hence the eigenvalue Γ_{nm} is again obtained by equating to zero the sum of all contributions involving L_k^n. Proceeding in the familiar way,[9-14] we obtain

$$\begin{aligned}\Gamma_{nm}(c) = {} & 2c(m+1) - \tfrac{1}{2}[(m-n)(m+n+2) + 2n + \tfrac{5}{2}] + (1/2^7c)[(m+n)(m+n-2)(m-n)(m-n-2)\\ & - (m-n+2)(m+n+2)(m-n+4)(m+n+4)] + (1/2^8c^2)[(m+n)(m+n-2)(m-n)(m-n-2)(m-1)\\ & - (m-n+2)(m+n+2)(m-n+4)(m+n+4)(m+3)] + (1/2^{18}c^3)\{(m+n)(m-n)(m+n-2)(m-n-2)\\ & \times[2^9(m-1)^2 + (m+n-4)(m-n-4)(m+n-6)(m-n-6)] - (m-n+2)(m+n+2)(m-n+4)\\ & \times(m+n+4)[2^9(m+3)^2 + (m-n+6)(m+n+6)(m-n+8)(m+n+8)] - 2[(m+n)^2(m+n-2)^2\\ & \times(m-n)^2(m-n-2)^2 - (m-n+2)^2(m+n+2)^2(m-n+4)^2(m+n+4)^2]\} + O(c^{-4})\end{aligned} \tag{51}$$

and

$$\begin{aligned}T(z) = {} & T^{(0)} + T^{(1)} + T^{(2)} + \cdots\\ = {} & L_k^n(z) + (1/32c)\{(m+n)(m+n-2)L_{k-2}^n(z) - (m-n+2)(m-n+4)L_{k+2}^n(z)\}\\ & + (1/2^{11}c^2)\{(m-n+2)(m-n+4)(m-n+6)(m-n+8)L_{k+4}^n(z) - 2^5(m-n+2)(m-n+4)(m+3)L_{k+2}^n(z)\\ & + 2^5(m+n)(m+n-2)(m-1)L_{k-2}^n(z) + (m+n)(m+n-2)(m+n-4)(m+n-6)L_{k-4}^n(z)\} + O(c^{-3}).\end{aligned} \tag{52}$$

For the fundamental mode, (52) implies

$$\begin{aligned}S_{00}(x) = {} & e^{-(1/2)cx^2}\{L_0(cx^2) - (1/4c)L_2(cx^2) + (1/16c^2)[3L_4(cx^2) - 6L_2(cx^2)]\\ & + (1/64c^3)[-15L_6(cx^2) + 48L_4(cx^2) - 53L_2(cx^2)] + O(c^{-4})\}.\end{aligned} \tag{53}$$

These solutions are rapidly decreasing expansions provided $|x| \ll c$. The particular solution of the hyperspheroidal differential equation (14) which satisfies the boundary conditions (15) and represents an asymptotic expansion for large values of c valid in the domain $0 \leqslant x \leqslant c^{-1/4}$ (the upper bound being obtained as in Ref. 13) is therefore (apart from an overall normalization constant)

$$S_{nm}(x) = x^n e^{-(1/2)cx^2}\left[L_k^n(cx^2) + \sum_{p=1}^{\infty}\frac{1}{2^{2p}c^p} \times \sum_{\substack{j=-p \\ j\neq 0}}^{p} [k,k+j]_{2p} L_{k+j}^n(cx^2)\right], \tag{54}$$

where $[k,k+j]_{2p}$ are the appropriate coefficients and

$k = \frac{1}{2}(m - n)$. As before, a second, linearly independent solution is obtained by the replacements $m \to -m-2$, $c \to -c$, which again leave the expansion of the eigenvalue unchanged.

It may be noted that Heurtley[1] obtains the solution by solving a recurrence relation. His results agree with ours. Slepian,[2] on the other hand, gives the solution in terms of Laguerre polynomials but does not determine the expansion coefficients. However, he also investigates the WKB-like solutions in adjoining domains of validity, which we shall not discuss in the present investigation.

IV. COMPUTATION OF EIGENVALUES $\Gamma_{nm}(c)$ OF THE HYPERSPHEROIDAL DIFFERENTIAL EQUATION

We now use (27) for the calculation of $\Gamma_{nm}(c)$ for small values of c and similarly (51) for the calculation of $\Gamma_{nm}(c)$ for large values of c. Elsewhere[14] it has been shown that the perturbation procedure can be formulated in such a way that many more higher-order terms of these expansions can be obtained by an algebraic computer calculation. In the present investigation we restrict ourselves to the terms given in the above expansions. This implies, of course, that we can use these expansions reliably only for such values of c, m, and n, for which successive terms decrease sufficiently rapidly. Moreover, since the use of complicated convergence factor methods or Borel summation techniques is beyond the scope of the present considerations, we employ the well-known old Poincaré[15] recipe of terminating an asymptotic expansion with terms which are alternately positive and negative at the term preceding the least term and taking half the least term. If, on the other hand, the terms do not alternate in sign, the series is terminated at the least term. Except for borderline cases these considerations are unnecessary here in view of the rapid falloff of successive terms, e.g., $\Gamma_{02}(2) = 8.75 + 2 + 0.1 - 0.008\,44 + \cdots$.

In Table I we give eigenvalues calculated in this way for relatively small values of c and in Table II eigenvalues calculated for relatively large values of c. In Fig. 2 we demonstrate the matching of both types of expansions in the case of an eigenvalue for which comparison values are available from the work of Slepian.[2]

TABLE I. Eigenvalues $\Gamma_{nm}(c)$ of the differential equation obtained from (27) for small values of c.

c	Γ_{00}	Γ_{02}	Γ_{04}	Γ_{06}	
0.1	0.755	8.755	24.755	48.755	
0.5	0.874	8.875	24.875	48.875	
1	1.240	9.256	25.251	49.250	
2	2.586	10.841	26.772	50.760	
3	4.471	13.540	29.329	53.286	
4	...	17.550	32.940	56.830	
5	...	23.203	37.713	61.467	
6	...	30.800	43.714	67.200	
7	...	...	...	74.083	
8	...	...	...	82.172	
9	...	...	...	91.528	
c	Γ_{11}	Γ_{13}	Γ_{15}	Γ_{17}	Γ_{19}
0.1	3.756	15.755	35.755	63.755	99.755
0.5	3.916	15.883	35.878	63.877	99.876
1	4.412	16.284	36.265	64.258	100.255
2	6.342	17.909	37.820	65.789	101.775
3	9.375	20.682	40.444	68.359	104.320
5	...	30.101	49.115	76.743	112.567
6	...	37.062	55.318	82.647	118.328
10	...	...	95.312	119.262	153.315
c	Γ_{22}	Γ_{24}	Γ_{26}	Γ_{28}	$\Gamma_{2.10}$
0.1	8.757	24.755	48.755	80.755	120.755
0.5	8.937	24.896	48.885	80.881	120.879
1	9.498	25.334	49.292	81.275	121.266
2	11.712	27.087	50.923	82.854	122.820
3	15.310	30.018	53.655	85.498	125.417
5	26.035	39.475	62.524	94.055	133.801
6	...	49.045	68.732	100.025	139.801
10	...	...	...	136.143	174.568
c	Γ_{33}	Γ_{35}	Γ_{37}	Γ_{39}	
0.1	15.758	35.756	63.756	99.755	
0.5	15.950	35.907	63.892	99.886	
1	16.549	36.378	64.322	100.296	
2	18.928	38.261	66.037	101.934	
3	22.842	41.392	68.903	104.672	
5	34.916	51.348	84.489	113.487	
6	42.822	58.138	84.489	119.596	
10	82.416	96.756	122.187	155.913	

V. EIGENVALUES OF THE INTEGRAL EQUATION FOR SMALL VALUES OF c

The kernel of our original integral equation (11) does not possess a step function as in many well-known problems in quantum mechanics. As a consequence no simple relationship exists between the eigenvalues of the integral equation and the eigenvalues of the corresponding differential equation (for literature on this connection see Ref. 16). However, it is possible to calculate the eigenvalues $\gamma_{nm}(c)$ of the integral equation from a knowledge of the eigenfunctions $S_{nm}(x)$ of the differential equation. This is the procedure we follow here.

We can rewrite (11) as

$$\gamma_{nm}(c) = \left[\frac{x^{1/2} S_{nm}(x)}{c \int_0^1 y^{1/2} S_{nm}(y) J_n(cxy)(xy)^{1/2} dy} \right]_{x=0}. \tag{55}$$

For small values of c, $S_{nm}(x)$ is given by (34) or, explicitly, by

$$\begin{aligned} S_{nm}(x) = x^n \{ & P_k^{(0,n)} + c^2 (a_1 P_{k+1}^{(0,n)} + a_2 P_{k-1}^{(0,n)}) \\ & + c^4 (a_3 P_{k+2}^{(0,n)} + a_4 P_{k+1}^{(0,n)} + a_5 P_{k-1}^{(0,n)} \\ & + a_6 P_{k-2}^{(0,n)}) + O(c^6) \}, \end{aligned} \tag{56}$$

where $P_k^{(0,n)} = P_k^{(0,n)}(2x^2 - 1)$ and

$$\begin{aligned} a_1 &= -(m-n+2)(m+n+2)/16(m+1)(m+2)^2, \\ a_2 &= (m-n)(m+n)/16(m+2)m^2, \\ a_3 &= \frac{(m-n+2)(m+n+2)(m-n+4)(m+n+4)}{512(m+2)^2(m+1)(m+3)^2(m+4)}, \\ a_4 &= -\frac{(m-n+2)(m+n+2)n^2}{32(m+2)^4(m+1)m^2(m+4)}, \\ a_5 &= \frac{(m-n)(m+n)n^2}{32m^4(m+1)(m-2)(m+2)}, \\ a_6 &= \frac{(m-n)(m+n)(m-n-2)(m+n-2)}{512m^2(m+1)(m-2)(m-1)^2}. \end{aligned} \tag{57}$$

TABLE II. Eigenvalues $\Gamma_{nm}(c)$ of the differential equation obtained from (51) for large values of c.

c	Γ_{00}	Γ_{02}	Γ_{04}	Γ_{06}	Γ_{08}
4	6.58	16.92	...	...	...
5	8.62	23.40	...	...	...
6	10.65	29.69	...	...	...
7	12.66	35.88	53.01	...	...
8	14.68	42.01	63.61	...	...
9	16.69	48.11		...	...
10	18.69	54.19	84.85	108.18	124.16
15	28.71	84.40	135.31	180.82	220.17
20	38.72	114.50	185.88	251.75	312.74
c	Γ_{11}	Γ_{13}	Γ_{15}	Γ_{17}	Γ_{19}
5	17.330	...	...	...	...
7	25.474	45.92	...	...	...
8	29.516	54.21	...	...	...
10	37.570	70.59	...	...	...
15	57.637	111.03	159.75	201.92	237.94
20	77.667	151.23	220.05	283.65	341.51
c	Γ_{22}	Γ_{24}	Γ_{26}	Γ_{28}	
6	32.04	...	...	...	
8	44.26	...	...	...	
10	56.38	86.75	110.75	...	
15	86.52	137.53	183.15	222.65	
20	116.58	187.87	254.20	315.07	

For $J_n(cxy)$ we substitute its series expansion, i.e.,

$$J_n(cxy) = [\tfrac{1}{2}(cxy)^n/n!]\{1 + O((cxy)^2)\}.$$

Now,

$$P_k^{(0,n)}(-1) = (-1)^k(k+n)!/k!\,n!. \tag{58}$$

Inserting these expressions into (55) we obtain

$$\gamma_{nm}(c) = \frac{(-1)^k(k+n)!/k!\,n! + O(c^2)}{(c^{n+1}/2^n n!)\left[\int_0^1 y^{2n+1}P_k^{(0,n)}(2y^2-1)dy + O(c^2)\right]}. \tag{59}$$

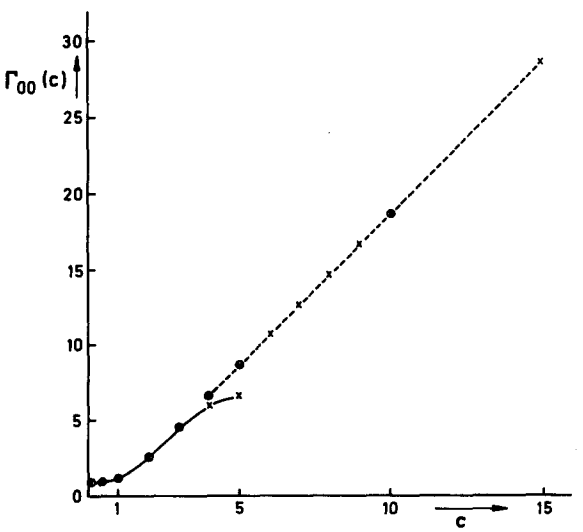


FIG. 2. Eigenvalues $\Gamma_{00}(c)$ obtained from the expansion for small c (——) and from the expansion for large c (- - - -). Exact values from Slepian.[2]

In the integral in the denominator we set

$$2y^2 - 1 = t.$$

Then

$$I = \int_0^1 y^{2n+1}P_k^{(0,n)}(2y^2-1)dy$$

$$= \frac{1}{2^{n+2}}\int_{-1}^1 (t+1)^n P_k^{(0,n)}(t)dt$$

$$= [1/2(n+1)]\,{}_3F_2(-k,n+k+1,1;1,n+2;1), \tag{60}$$

on using a formula of Ref. 17. The expression for the eigenvalue therefore becomes (for $k \geqslant 0$, $n \geqslant 0$)

$$\gamma_{nm}(c) = [(-1)^k 2^{n+1}(n+1)/c^{n+1}] \times \frac{\{(k+n)!/k! + O(c^2)\}}{\{{}_3F_2(-k,n+k+1,1;1,n+2;1) + O(c^2)\}}.$$

Taking into account the higher-order terms, this is

$$\gamma_{nm}(c) = [(-1)^k 2^{n+1}(n+1)/c^{n+1}](A/B), \tag{61}$$

TABLE III. Eigenvalues γ_{nm} and diffraction loss factor γ_{nm}^{-2} obtained from (61)–(63) for small values of c; comparison values taken from Slepian.[2]

c	γ_{00}	γ_{00}^{-2}	λ_{00} Slepian
0.1	2.001 25 − 1	2.496 88 − 3	2.496 877 5 − 3
0.2	1.002 50 − 1	9.950 16 − 3	...
0.3	6.704 21 0	2.224 87 − 2	...
0.4	5.050 11 0	3.921 01 − 2	...
0.5	4.062 72 0	6.058 52 − 2	6.058 534 8 − 2
0.6	3.408 71 0	8.606 37 − 2	...
0.7	2.945 24 0	1.152 81 − 1	...
0.8	2.600 89 0	1.478 28 − 1	...
0.9	2.335 99 0	1.832 56 − 1	...
1	2.126 74 0	2.210 91 − 1	2.211 148 7 − 1
1.5	1.526 71 0	4.290 32 − 1	4.295 190 6 − 1
2	1.263 92 0	6.259 81 − 1	6.296 304 5 − 1
2.5	1.139 69 0	7.698 90 − 1	...
3	1.088 65 0	8.437 74 − 1	8.870 503 6 − 1
3.5	1.083 53 0	8.517 60 − 1	...
4	1.111 36 0	8.096 37 − 1	
4.5	1.165 50 0	7.361 63 − 1	
5	1.242 50 0	6.477 50 − 1	
c	γ_{11}	γ_{11}^{-2}	λ_{10} Slepian
0.1	8.004 45 − 2	1.560 76 − 6	...
0.2	2.004 45 − 2	2.488 91 − 5	...
0.3	8.933 43 − 1	1.253 03 − 4	...
0.4	5.044 62 − 1	3.929 55 − 4	...
0.5	3.244 72 − 1	9.498 27 − 4	9.498 265 8 − 1
0.6	2.267 01 − 1	1.945 68 − 3	...
0.7	1.677 65 − 1	3.553 04 − 3	...
0.8	1.295 16 − 1	5.961 47 − 3	...
0.9	1.033 00 − 1	9.371 21 − 3	...
1	8.455 63 0	1.398 65 − 2	1.398 616 8 − 2
1.5	4.025 15 0	6.172 13 − 2	...
2	2.489 16 0	1.613 97 − 1	1.612 318 3 − 1
2.5	1.794 31 0	3.106 03 − 1	...
3	1.433 94 0	4.863 41 − 1	4.832 686 6 − 1
3.5	1.234 43 0	6.562 41 − 1	...
4	1.123 29 0	7.925 31 − 1	7.847 350 5 − 1
4.5	1.065 86 0	8.802 44 − 1	...
5	1.043 89 0	9.176 82 − 1	
6	1.043 89 0	8.749 69 − 1	

where

$$A = \frac{(k+n)!}{k!} - c^2\left[a_1\frac{(k+n+1)!}{(k+1)!} + a_2\frac{(k-1+n)!}{(k-1)!}\right] + c^4\left[a_3\frac{(k+2+n)!}{(k+2)!} - a_4\frac{(k+1+n)!}{(k+1)!} - a_5\frac{(k-1+n)!}{(k-1)!} + a_6\frac{(k-2+n)!}{(k-2)!}\right] + O(c^6), \quad (62)$$

and, for $k \geqslant 0$, $n \geqslant 0$,

$$B = F(k) + c^2[a_1F(k+1) + a_2F(k-1)] + c^4[a_3F(k+2) + a_4F(k+1) + a_5F(k-1) + a_6F(k-2)] + O(c^6), \quad (63)$$

where $F(k) \equiv {}_3F_2(-k, n+k+1, 1; 1, n+2; 1)$.

If in the evaluation of these terms a contribution arises which is infinite, then our earlier considerations apply [see comments following (28)] and it has to be ignored (i.e., a separate calculation is necessary). In particular, we have

$$\gamma_{00}(c) \simeq \frac{2}{c}\left\{\frac{1 - a_1c^2 + a_3c^4 + O(c^6)}{1 + O(c^6)}\right\}, \quad (64)$$

with $a_1 = -0.062\,50$, $a_3 = 0.000\,87$, $a_4 = 0$.

In Table III we show some eigenvalues calculated in this way. The quantity $1/\gamma_{nm}^2$ [see (12)], which is a measure of the diffraction loss, is also given and compared with values obtained by Slepian.[2]

VI. EIGENVALUES OF THE INTEGRAL EQUATION FOR LARGE VALUES OF c

In this case we again proceed as before except that now we use (54) for $S_{nm}(x)$, which is valid for $0 \leqslant x < c^{-1/4}$. Then

TABLE IV. Eigenvalues γ_{nm} and diffraction loss factor γ_{nm}^{-2} obtained from (71) for large values of c; comparison values taken from Slepian.[2]

c	γ_{00}	γ_{00}^{-2}	λ_{00} Slepian
20	1.000 02	0.999 96	...
10	1.003 48	0.993 08	0.999 999 57
9	1.005 76	0.988 58	...
8	1.009 58	0.981 11	...
7	1.000 72	0.998 56	...
6	1.004 08	0.991 89	...
5	1.013 12	0.974 26	...
4	1.035 90	0.931 89	0.974 951 17
3.5	1.056 80	0.895 40	...
3	1.089 98	0.841 71	0.887 050 36
2.5	1.142 69	0.765 84	...
2	1.229 76	0.661 24	0.629 304 5

$$x^{1/2}S_{nm}(x) = \gamma_{nm}c\left\{\int_0^{c^{-1/4}} y^{n+1/2}e^{-(1/2)cy^2}[L_k^n(cy^2) + O(c^{-1})]J_n(cxy)(xy)^{1/2}\,dy + \int_{c^{-1/4}}^1 (\cdots)\right\}.$$

The contribution $\int_{c^{-1/4}}^1$ would have to be calculated with the help of a new class of solutions, dubbed WKB-like in previous investigations.[9] However, these solutions are exponentially decreasing. Moreover, we can argue (as in many other cases dealing with similar integrals[10,13]) that the main contribution to the entire integral comes from the domain around the lower limit, and hence we integrate the first contribution over the whole domain of integration, thereby approximating the second integral effectively by the integral over the extended domain. It will be seen that our numerical results justify this procedure *a posteriori*. Thus,

$$\gamma_{nm}(c) = \left[\frac{e^{-(1/2)cx^2}\{L_k^n(cx^2) + O(c^{-1})\}}{(c^{n+1}/2^n n!)\{\int_0^1 y^{2n+1}e^{-(1/2)cy^2}L_k^n(cy^2)dy + O(c^{-1})\}}\right]_{x=0}. \quad (65)$$

With the change of variable

$$t = cy^2,$$

we have

$$I \equiv \int_0^1 y^{2n+1}e^{-(1/2)cy^2}L_k^n(cy^2)dy = \frac{1}{2c^{n+1}}\int_0^c t^n e^{-(1/2)t}L_k^n(t)dt \simeq \frac{1}{2c^{n+1}}\int_0^\infty t^n e^{-(t/2)t}L_k^n(t)dt = \frac{1}{c^{n+1}}\cdot\frac{(n+k)!(-1)^k2^n}{k!}, \quad (66)$$

using

$$\int_0^\infty e^{-st}t^\alpha L_n^\alpha(t)dt = \frac{\Gamma(\alpha+n+1)(s-1)^n}{n!s^{\alpha+n+1}}.$$

Inserting the result into (65) and using

$$L_k^n(0) = (n+k)!/k!n!, \quad (67)$$

we obtain

$$\gamma_{nm}(c) = [(n+k)!/n!k! + O(c^{-1})]/[(n+k)!(-1)^k/n!k! + O(c^{-1})]. \quad (68)$$

Thus for $c \to \infty$ the eigenvalues γ_{nm} approach 1, and the diffraction loss

$$P_{nm} = 1 - 1/|\gamma_{nm}|^2$$

approaches zero in agreement with what one expects on physical grounds for infinitely large mirrors. The calculation of higher-order contributions is difficult.

We have [with $L_k^n \equiv L_k^n(cx^2)$]

$$S_{nm}(x) = x^n e^{-(1/2)cx^2}[L_k^n + (1/c)\{a_1 L_{k-2}^n + a_2 L_{k+2}^n\} + (1/c^2)\{a_3 L_{k+4}^n + a_4 L_{k+2}^n + a_5 L_{k-2}^n + a_6 L_{k-4}^n\} + (1/c^3)\{a_7 L_{k+6}^n + a_8 L_{k+4}^n + \cdots + a_{12} L_{k-6}^n\} + O(1/c^4)], \tag{69}$$

where

$$a_1 = \tfrac{1}{32}(m+n)(m+n+2), \quad a_2 = -\tfrac{1}{32}(m-n+2)(m-n+4),$$
$$a_3 = (1/2^{11})(m-n+2)(m-n+4)(m-n+6)(m-n+8), \quad a_4 = -(1/2^6)(m-n+2)(m-n+4)(m+3),$$
$$a_5 = (1/2^6)(m+n)(m+n-2)(m-1), \quad a_6 = (1/2^{11})(m+n)(m+n-2)(m+n-4)(m+n-6),$$
$$a_7 = -(1/2^{16}\cdot 3)(m-n+2)(m-n+4)(m-n+6)(m-n+8)(m-n+10)(m-n+12),$$
$$a_8 = (1/2^{11})(m-n+2)(m-n+4)(m-n+6)(m-n+8)(m+4), \tag{70}$$
$$a_9 = -(1/2^{16})(m-n+2)(m-n+4)\{2^9(m+3)^2 + (m-n+6)(m-n+8)(m+n+8)(m+n+6) + 2[(m+n)(m+n-2)(m-n)(m-n-2) - (m-n+2)(m-n+4)(m+n+2)(m+n+4)]\},$$
$$a_{10} = (1/2^{16})(m+n)(m+n-2)\{2^9(m-1)^2 + (m+n-4)(m+n-6)(m-n-4)(m-n-6) - 2[(m+n)(m+n-2)(m-n)(m-n-2) - (m-n+2)(m-n+4)(m+n+2)(m+n+4)]\},$$
$$a_{11} = (1/2^{11})(m+n)(m+n-2)(m+n-4)(m+n-6)(m-2),$$
$$a_{12} = (1/2^{16}\cdot 3)(m+n)(m+n-2)(m+n-4)(m+n-6)(m+n-8)(m+n-10).$$

We then find with the approximations described at the beginning

$$\gamma_{nm} = C/(C+D), \tag{71}$$

where

$$C = \frac{(n+k)!}{k!} + \frac{1}{c}\left[\frac{a_1(k-2+n)!}{(k-2)!} + \frac{a_2(k+2+n)!}{(k+2)!}\right] + \frac{1}{c^2}\left[\frac{a_3(k+4+n)!}{(k+4)!} + \frac{a_4(k+2+n)!}{(k+2)!} + \frac{a_5(k-2+n)!}{(k-2)!} + \frac{a_6(k-4+n)!}{(k-4)!}\right] + \frac{1}{c^3}\left[\frac{a_7(k+6+n)!}{(k+6)!} + \frac{a_8(k+4+n)!}{(k+4)!} + \frac{a_9(k+2+n)!}{(k+2)!} + \frac{a_{10}(k-2+n)!}{(k-2)!} + \frac{a_{11}(k-4+n)!}{(k-4)!} + \frac{a_{12}(k-6+n)!}{(k-6)!}\right] + O(c^{-4})$$

and

$$D = -\frac{1}{2^{n+1}}\int_c^\infty t^n e^{-(1/2)t}\left\{L_k^n(t) + \frac{1}{c}[a_1 L_{k-2}^n + a_2 L_{k+2}^n] + \frac{1}{c^2}[a_3 L_{k+4}^n + a_4 L_{k+2}^n + a_5 L_{k-2}^n + a_6 L_{k-4}^n] + (1/c^3)[a_7 L_{k+6}^n + a_8 L_{k+4}^n + a_9 L_{k+2}^n + a_{10} L_{k-2}^n + a_{11} L_{k-4}^n + a_{12} L_{k-6}^n] + O(c^{-4})\}dt.$$

These integrals have to be evaluated individually. This has been done for $n = 0$, $k = 0$. The eigenvalues $\gamma_{00}(c)$ have then been calculated with the prescription of Poincaré[15] mentioned earlier. The results are shown in Table IV. The comparison values are again taken from the work of Slepian.[2]

We have thus shown how the eigenvalues of the original homogeneous integral equation can be calculated with the help of the eigenfunctions of the corresponding differential equation. Of course, for many numerical calculations it is more practical to obtain the eigenvalues of the integral equation by a numerical computer calculation.

Finally we mention for the sake of completeness that some authors use a different enumeration of resonator modes. Thus the modes (N, M) of Fox and Li[18] correspond to our modes $(n, (m-n)/2)$.

ACKNOWLEDGMENT

We thank Professor H. Weber of this department for bringing the problem of this investigation to our attention.

[1] J. C. Heurtley, Proceedings of the Symposium on Quasi-Optics, Polytechnic Institute, June 1964; also J. Opt. Soc. Am. **54**, 1400 (1964).
[2] D. Slepian, Bell Syst. Tech. J. **43**, 3009 (1964).
[3] H. Hochstadt, SIAM Rev. **8**, 62 (1966).
[4] J. A. Cochran, Bell Syst. Tech. J. **44**, 77 (1965).
[5] D. J. Newman and S. P. Morgan, Bell Syst. Tech. J. **43**, 113 (1964).
[6] D. J. Newman and S. P. Morgan, TEEE J. Quant. Electron. **54**, 100 (1966).
[7] C. Flammer, *Spheroidal Wave Equations* (Stanford U. P., Stanford, 1957).
[8] J. Meixner and F. W. Schäfke, *Mathieusche Funktionen und Sphäroidfunktionen* (Springer, Berlin, 1954).
[9] H. J. W. Müller-Kirsten, Phys. Rev. D **22**, 1952 (1980).
[10] R. B. Dingle and H. J. W. Müller, J. Reine Angew. Math. **211**, 11 (1962);

216, 123 (1964); H. J. W. Müller, J. Reine Angew. Math. **211**, 33, 179 (1962); **212**, 26 (1963).
[11]H. H. Aly, H. J. W. Müller-Kirsten, and N. Vahedi, J. Math. Phys. **16**, 961 (1975).
[12]S. K. Bose and H. J. W. Müller-Kirsten, J. Math. Phys. **20**, 2471 (1979); S. K. Bose, G. E. Hite, and H. J. W. Müller-Kirsten, J. Math. Phys. **20**, 1878 (1979).
[13]R. S. Kaushal and H. J. W. Müller-Kirsten, J. Math. Phys. **20**, 2233 (1979).
[14]R. Müller and H. J. W. Müller-Kirsten, Phys. Rev. D **20**, 2541 (1979); J. Math. Phys. **22**, 733 (1981).
[15]H. Poincaré cited by E. T. Whittaker and G. N. Watson, *A Course in Modern Analysis* (Cambridge U. P., Cambridge, England, 1927). See also T. Stieltjes, Ann. Sci. Éc. Norm. Sup. **3**, 201 (1886) and G. N. Watson, *Theory of Bessel Functions* (Cambridge U. P., Cambridge, England, 1952).
[16]W. von Koppenfels, Math. Ann. **112**, 24 (1936).
[17]I. S. Gradschteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1965).
[18]A. G. Fox and T. Li, Bell Syst. Tech. J. **40**, 453 (1960); IEEE J. Quant. Electron. **QE-2**, 774 (1966).

First integrals for the modified Emden equation $\ddot{q} + \alpha(t)\dot{q} + q^n = 0$

P. G. L. Leach
Department of Applied Mathematics, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg 2001, South Africa

(Received 18 April 1984; accepted for publication 1 March 1985)

It is shown that the modified Emden equation $\ddot{q} + \alpha(t)\dot{q} + q^n = 0$ possesses first integrals for functions $\alpha(t)$ other than kt^{-1}. The function $\alpha(t)$ is obtained explicitly in the case $n = 3$ and parametrically for other $n(\neq 2)$. The case $n = 2$ is seen to be particularly difficult to solve.

I. INTRODUCTION

The Emden equation of index n,

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\theta}{d\xi}\right) = -\theta^n,$$
$$\theta(0) = 1, \quad \theta'(0) = 0, \quad \xi \geqslant 0, \tag{1.1}$$

arises in the study of equilibrium configurations of a spherical gas cloud acting under the mutual attractions of its molecules and subject to the laws of thermodynamics. The index n is given by

$$n = (c_v - c)/(c_p - c_v), \tag{1.2}$$

where c_v is the specific heat at constant volume, c_p is the specific heat at constant pressure, and c is the assumed constant in the relationship between heat input dQ and temperature change dT, i.e.,

$$dQ = c\,dT. \tag{1.3}$$

Chandrasekhar[1] discusses the equation extensively and reports complete solutions in the cases $n = 0,1,5$.

The Emden equation may also be considered as an equation in dynamics, viz.,

$$\ddot{q} + (2/t)\dot{q} + q^n = 0, \tag{1.4}$$

which represents in general an anharmonic oscillator subject to damping dependent upon the velocity. In the case $n = 5$, Logan[2] illustrates the use of Noether's theorem with point transformations by applying the theorem to the variational integral

$$J = \int_{t_0}^{t_1} t^2\left(\frac{1}{2}\dot{q}^2 - \frac{1}{6}q^6\right)dt, \tag{1.5}$$

the integrand of which is a Lagrangian from which (1.4) may be obtained. He obtained the first integral

$$I = \tfrac{1}{6}t^3q^6 + \tfrac{1}{2}t^3\dot{q}^2 + \tfrac{1}{2}t^2q\dot{q}. \tag{1.6}$$

This same first integral was obtained by Sarlet and Bahar[3] by introducing a time-dependent integrating factor. They extended their treatment to the more general equation

$$\ddot{q} + \beta(t)\dot{q} + \alpha(t)q^m = 0, \quad m \neq -1 \tag{1.7}$$

and found that the first integral

$$I = \left(\dot{q}^2 + \frac{2\alpha}{m+1}q^{m+1}\right)\exp\left(2\int^t \beta(t')dt'\right)$$
$$\times\left[C + C_4\int^t \exp\left(-\int^{t'}\beta(t'')\,dt''\right)dt'\right]$$
$$- C_4q\dot{q}\exp\left(\int^t \beta(t')\,dt'\right) \tag{1.8}$$

exists when $\alpha(t)$ and $\beta(t)$ satisfy the relation

$$\alpha^{-2/(m+3)}\exp\left(-\frac{4}{m+3}\int^t\beta(t')\,dt'\right)$$
$$- C_4\int^t \exp\left(-\int^{t'}\beta(t'')\,dt''\right)dt' = C \tag{1.9}$$

and C and C_4 are constants. Furthermore, they showed that the result was equivalent to using the gauge-invariant Noether approach. The extension to the gauge-variant Noether approach was mentioned, but the calculations and integral were not given.

Moreira,[4] as an example of the application of the Lewis–Leach[5] direct method to Newtonian equations of motion, studied the equation

$$\ddot{q} + \alpha(t)\dot{q} + q^n = 0. \tag{1.10}$$

He found the first integral

$$I = \exp\left(m\int^t\alpha(t')\,dt'\right)\{\dot{q}^2 + \alpha q\dot{q}(2-m)$$
$$+ [2/(n+1)]\,q^{n+1} + \tfrac{1}{2}q^2(m-2)[(m-1)\alpha^2 + \dot{\alpha}]\}, \tag{1.11}$$

where $m = 2(n+1)/(n+3)$, provided that $\alpha(t)$ satisfied the differential equation

$$\ddot{\alpha} + (3m-2)\alpha\dot{\alpha} + \alpha^3m(m-1) = 0. \tag{1.12}$$

He did not provide the general solution of (1.12), but gave two particular solutions, viz.,

$$\alpha(t) = 2/mt, \quad \alpha(t) = [(m-1)t]^{-1}. \tag{1.13}$$

Feix and Lewis,[6] in their study of first integrals for dissipative nonlinear systems using rescaling (the equivalent in Newtonian mechanics of generalized canonical transformations in Hamiltonian mechanics[7]), examined the equation

$$\ddot{x} + \beta(t)\dot{x} + \frac{\partial\phi(x,t)}{\partial x} = 0. \tag{1.14}$$

As an example they treat the case in which

$$\phi(x,t) = \gamma(t)\,[x^{m+1}/(m+1)], \quad m \neq -1. \tag{1.15}$$

However, they specified $\beta(t)$ to be proportional to t^{-1}.

In this note we intend to apply Noether's theorem to an equation of Emden type. It is necessary to decide which of the various forms of the equation given above is to be used.

 0022-2488/85/102510-05$02.50

Our choice is that of Moreira (1.10). The reason for this is as follows. Consider the equation

$$\ddot{x} + \beta(t)\dot{x} + \gamma(t)x^n = 0, \tag{1.16}$$

which was discussed in Refs. 3 and 6. Under the transformation

$$(x,t) \to (X,T: X = x, T = f(t)), \tag{1.17}$$

(1.17) becomes

$$X'' + (\ddot{f}\dot{f}^{-2} + \beta\dot{f}^{-1})X' + \gamma\dot{f}^{-2}X^n = 0, \tag{1.18}$$

where prime denotes d/dT. If we set

$$\gamma = (\dot{f})^2, \quad \alpha(T(t)) = \ddot{f}\dot{f}^{-2} + \beta\dot{f}^{-1}, \tag{1.19}$$

we have an equation of type (1.10). Should it happen that $\gamma(t)$ is negative we could replace the first of (1.19) by $\gamma = -\dot{f}^{-2}$ to keep $\dot{f}$ real. If $\gamma(t)$ has zeros, the transformation (1.17) would strictly apply to the interval between successive zeros and matching of the results for successive intervals would have to be undertaken.

One may wonder what relevance this has to the original Emden equation (1.1). This equation comes from the equation of equilibrium[1]

$$\frac{1}{r^2}\frac{d}{dr}\left(\frac{r^2}{\rho}\frac{dP}{dr}\right) = -4\pi G\rho, \tag{1.20}$$

where r is the radial variable in the gaseous sphere, ρ is the density of the gas, P is its pressure, and G is the universal gravitational constant, by means of the substitutions

$$\rho = \lambda\theta^n, \quad P = K\lambda^{(n+1)/n}\theta^{n+1} \tag{1.21}$$

and suitable rescaling. The density and pressure are related by

$$P = K\rho^{(n+1)/n}. \tag{1.22}$$

If we imagine that the constant K is replaced by a function of r, we obtain an equation of type (1.10) in terms of a variable σ, where

$$P = K_1\sigma^{(n+1)/n}, \quad \sigma = f(r)\rho. \tag{1.23}$$

II. APPLICATION OF NOETHER'S THEOREM

Following Sarlet and Cantrijn,[8] if the functional

$$J = \int_{t_0}^{t_1} L(q,\dot{q},t)\,dt, \tag{2.1}$$

where L is regular in $\dot{q}$, admits a gauge-variant symmetry generated by

$$G = \tau\frac{\partial}{\partial t} + \xi\frac{\partial}{\partial q}, \tag{2.2}$$

where $\tau = \tau(q,\dot{q},t)$ and $\xi = \xi(q,\dot{q},t)$, there exists a first integral

$$I = \left[L\tau + \frac{\partial L}{\partial\dot{q}}(\xi - \dot{q}\tau)\right] - f(q,\dot{q},t), \tag{2.3}$$

where f, τ, and ξ are determined by the equations

$$\begin{aligned} &L\frac{\partial\tau}{\partial\dot{q}} + \frac{\partial L}{\partial\dot{q}}\left(\frac{\partial\xi}{\partial\dot{q}} - \dot{q}\frac{\partial\tau}{\partial\dot{q}}\right) = \frac{\partial f}{\partial\dot{q}}, \\ &\tau\frac{\partial L}{\partial t} + \xi\frac{\partial L}{\partial q} + L\left(\frac{\partial\tau}{\partial t} + \dot{q}\frac{\partial\tau}{\partial q}\right) \\ &\quad + \frac{\partial L}{\partial\dot{q}}\left\{\frac{\partial\xi}{\partial t} + \dot{q}\frac{\partial\xi}{\partial q} - \dot{q}\left(\frac{\partial\tau}{\partial t} + \dot{q}\frac{\partial\tau}{\partial q}\right)\right\} \\ &\quad = \frac{\partial f}{\partial t} + \dot{q}\frac{\partial f}{\partial\dot{q}}. \end{aligned} \tag{2.4, 2.5}$$

Since, further,

$$\xi - \dot{q}\tau = -g\frac{\partial I}{\partial\dot{q}}, \tag{2.6}$$

where

$$\frac{\partial^2 L}{\partial\dot{q}^2}g = 1, \tag{2.7}$$

we may without loss of generality set $\tau = 0$ and do so.

The equation of motion we study is

$$\ddot{q} + \alpha(t)\dot{q} + q^n = 0, \quad n \neq 1,0,1. \tag{2.8}$$

(The cases $n = 0,1$ are excluded since they are linear systems and have already been treated elsewhere.[9] For the case $n = -1$ see the Appendix.) A Lagrangian for this equation is

$$L(q,\dot{q},t) = \tfrac{1}{2}A(t)\dot{q}^2 - [A(t)/(n+1)]\,q^{n+1}, \tag{2.9}$$

where

$$A(t) = \exp\left(\int^t \alpha(t')\,dt'\right). \tag{2.10}$$

As the first integrals obtained in Refs. 2, 3, 4, and 6 were quadratic in $\dot{q}$, we look for one of the same type. From (2.6) it is evident that

$$\xi = a(q,t)\dot{q} + b(q,t). \tag{2.11}$$

From (2.3) we see that the first integral is

$$I = A\dot{q}(a\dot{q} + b) - f. \tag{2.12}$$

Substituting for L, ξ, and τ into (2.4) and integrating with respect to $\dot{q}$ we see that

$$f = \tfrac{1}{2}aA\dot{q}^2 + c(q,t) \tag{2.13}$$

and (2.12) becomes

$$I = \tfrac{1}{2}aA\dot{q}^2 + bA\dot{q} - c. \tag{2.14}$$

Substituting for L, ξ, τ, and f into (2.5) and equating the coefficients of like powers of $\dot{q}$ to zero, we have

$$(\dot{q}^3) \quad A\frac{\partial a}{\partial q} = \frac{1}{2}\frac{\partial}{\partial q}(Aa), \tag{2.15}$$

$$(\dot{q}^2) \quad A\frac{\partial a}{\partial t} + A\frac{\partial b}{\partial q} = \frac{1}{2}\frac{\partial}{\partial t}(aA), \tag{2.16}$$

$$(\dot{q}^1) \quad -aAq^n + A\frac{\partial b}{\partial t} = \frac{\partial c}{\partial q}, \tag{2.17}$$

$$(\dot{q}^0) \quad -Abq^n = \frac{\partial c}{\partial t}. \tag{2.18}$$

From (2.15), (2.16), and (2.17) in turn we find that

$$a = \sigma(t), \tag{2.19}$$

$$b = -\frac{1}{2A}(\dot{\sigma}A - \sigma\dot{A})\dot{q} - \gamma(t), \tag{2.20}$$

$$\begin{aligned} c = &-\frac{\sigma Aq^{n+1}}{n+1} - \frac{1}{4}A\frac{d}{dt}\left[\frac{1}{A}(\dot{\sigma}A - \sigma\dot{A})\right]q^2 \\ &- A\dot{\gamma}q - \delta(t). \end{aligned} \tag{2.21}$$

Substitution for b and c into (2.18) gives the consistency condition

$$\frac{q^{n+1}}{n+1}\frac{d}{dt}(\sigma A)+\frac{1}{4}\frac{d}{dt}\left\{A\frac{d}{dt}\left[\frac{1}{a}(\dot\sigma A-\sigma\dot A)\right]\right\}q^2 + \frac{d}{dt}(A\dot\gamma)q+\frac{1}{2}(\dot\sigma A-\sigma\dot A)q^{n+1}+A\gamma q^n+\dot\delta=0. \tag{2.22}$$

Since $n\neq -1,0,1$, we distinguish two cases.

(1) $n=2$: Equating the coefficients of like powers of q to zero, we have

$$(q^3)\quad \frac{1}{3}\frac{d}{dt}(\sigma A)+\frac{1}{2}(\dot\sigma A-\dot A)=0, \tag{2.23}$$

$$(q^2)\quad \frac{1}{4}\frac{d}{dt}\left\{A\frac{d}{dt}\left[\frac{1}{A}(\dot\sigma A-\sigma\dot A)\right]\right\}+A\gamma=0, \tag{2.24}$$

$$(q^1)\quad \frac{d}{dt}(A\dot\gamma)=0, \tag{2.25}$$

$$(q^0)\quad \dot\delta=0. \tag{2.26}$$

(2) $n\neq 2$: Equating the coefficient of like powers of q to zero, we have

$$(q^{n+1})\quad \frac{1}{n+1}\frac{d}{dt}(\sigma A)+\frac{1}{2}(\dot\sigma A-\sigma\dot A)=0, \tag{2.27}$$

$$(q^n)\quad A\gamma=0, \tag{2.28}$$

$$(q^2)\quad \frac{d}{dt}\left\{A\frac{d}{dt}\left[\frac{1}{A}(\dot\sigma A-\sigma\dot A)\right]\right\}=0, \tag{2.29}$$

$$(q^1)\quad \frac{d}{dt}(A\dot\gamma)=0, \tag{2.30}$$

$$(q^0)\quad \dot\delta=0. \tag{2.31}$$

III. THE CASE $n\neq 2$

It is evident that $\gamma=0$ and δ is an ignorable constant. From (2.27) we obtain

$$(n+3)(\dot\sigma/\sigma)=(n-1)(\dot A/A),\quad n\neq -3, \tag{3.1}$$

$$\sigma\dot A=0,\quad n=-3. \tag{3.2}$$

From (3.2) it follows that either $\sigma=0$, in which case the first integral is trivially a constant, or $\dot A=0$, in which case $\alpha=0$. As neither is of interest to us we take $n\neq -3$ henceforth. Integrating (3.1), we have

$$\sigma^{n+3}=K_1A^{n-1}, \tag{3.3}$$

where K_1 is some constant. If we make the substitution

$$\sigma=\rho^{n-1},\quad A=K_2\rho^{n+3}, \tag{3.4}$$

(2.29) becomes

$$\rho^2\ddot\rho+(4n-3)\rho\dot\rho\ddot\rho+2n(n-2)\dot\rho^3=0. \tag{3.5}$$

Since

$$A(t)=\exp\left(\int^t\alpha(t')\,dt'\right),$$

we have

$$\alpha=(n+3)\dot\rho\rho^{-1}, \tag{3.6}$$

and the function $\alpha(t)$, for which a first integral exists, satisfies

$$\ddot\alpha+[4n/(n+3)]\alpha\dot\alpha+[2(n^2-1)/(n+3)^2]\alpha^3=0, \tag{3.7}$$

which comes from (3.5). In terms of ρ, the first integral (2.14) is

$$I=\tfrac{1}{2}\rho^{2n+2}\dot q^2+2\dot\rho\rho^{2n+1}\dot qq+\rho^{2n+2}q^{n+1}/(n+1) - \{(n-2)\rho^{2n}\dot\rho^2+\rho^{2n+1}\ddot\rho\}q^2. \tag{3.8}$$

The invariance of I is easily checked by direct differentiation. An explicit form for the integral requires a knowledge of the functional expression for ρ. As we also require α, we look for the solution of (3.7) and then obtain ρ by quadrature of (3.6).

It is relatively straightforward to perform the first integration of (3.7). We obtain

$$C_1\left(\dot\alpha+\frac{n+1}{n+3}\alpha^2\right)^{n+1}=C_2\left(\dot\alpha+\frac{n-1}{n+3}\alpha^2\right)^{n-1},\quad |C_1|+|C_2|\neq 0. \tag{3.9}$$

(This form is suggested by the similar result found in Kamke,[10] p. 329, No. 1.204.) Before considering the next integration of (3.9) in general, we note two particular cases. If $C_1=0$,

$$\dot\alpha+[(n-1)/(n+3)]\alpha^2=0, \tag{3.10}$$

which may be integrated to give

$$\alpha(t)=(K_1+[(n-1)/(n+3)]t)^{-1}, \tag{3.11}$$

and from (3.6)

$$\rho(t)=K_2(K_1+[(n-1)/(n+3)]t)^{1/(n-1)}. \tag{3.12}$$

Likewise, if $C_2=0$,

$$\alpha(t)=(K_1+[(n+1)/(n+3)]t)^{-1}, \tag{3.13}$$

$$\rho(t)=K_2(K_1+[(n+1)/(n+3)]t)^{1/(n+1)}, \tag{3.14}$$

where K_1 and K_2 are arbitrary constants and K_2 may be set at unity since it has only a scaling effect on the first integral. These results are in accordance with those of Moreira[4] if K_1 is set at zero. For ρ as given in (3.12) the first integral is

$$I_1=\frac{1}{2}\left(K+\frac{n-1}{n+3}t\right)^{(2n+2)/(n-1)}\dot q^2 + \frac{2}{n+3}\left(K+\frac{n-1}{n+3}t\right)^{(n+3)/(n-1)}q\dot q + \left(K+\frac{n-1}{n+3}t\right)^{(2n+2)/(n-1)}\frac{q^{n+1}}{n+1}, \tag{3.15}$$

and for ρ as given in (3.14) the first integral is

$$I_2=\frac{1}{2}\left(K+\frac{n+1}{n+3}t\right)^2\dot q^2+\frac{2}{n+3}\left(K+\frac{n+1}{n+3}t\right)q\dot q + \left(K+\frac{n+1}{n+3}t\right)^2\frac{q^{n+1}}{n+1}+\frac{2}{(n+3)^2}q^2. \tag{3.16}$$

The integration of (3.9) is not simple when both C_1 and C_2 are nonzero. If we make the substitution

$$u(t)=-\alpha^{-1}(t), \tag{3.17}$$

(3.9) becomes, after suitable rearrangement,

$$u=k\left[\frac{(\dot u+(n+1)/(n+3))^{n+1}}{(\dot u+(n-1)/(n+3))^{n-1}}\right]^{1/4}. \tag{3.18}$$

The solution of (3.18) may be written in parametric form (Kamke,[10] p. 30, 4.17) as

$$u = k\left[\frac{(\eta + (n+1)/(n+3))^{n+1}}{(\eta + (n-1)/(n+3))^{n-1}}\right]^{1/4}, \tag{3.19}$$

$$t = \frac{1}{2}k\int^{\eta}\left[\frac{(\eta' + (n+1)/(n+3))^{n-3}}{(\eta' + (n-1)/(n+3))^{n+3}}\right]^{1/4} d\eta' . \tag{3.20}$$

In the particular case $n = 3$, we obtain

$$t = \frac{1}{2}k\int^{\eta}\left(s + \frac{1}{3}\right)^{-3/2} ds \tag{3.21}$$

and so

$$\alpha(t) = 3(t - K_1)/[K_2 + (t - K_1)^2] , \tag{3.22}$$

$$\rho(t) = \{ K_2 + (t - K_1)^2\}^{1/4} . \tag{3.23}$$

Hence the differential equation

$$\ddot{q} + 3(t - K_1)\dot{q}/[K_2 + (t - K_1)^2] + q^3 = 0 \tag{3.24}$$

possesses the first integral

$$I = [K_2 + (t - K_1)^2]^2\dot{q}^2 + 2(t - K_1)[K_2 + (t - K_1)^2]q\dot{q} + \tfrac{1}{2}[K_2 + (t - K_1)^2]^2q^4 - K_2 q^2 . \tag{3.25}$$

For general n, we make the substitution

$$\eta + \frac{n-1}{n+2} = \frac{2}{n+3}\frac{1}{x^2}, \quad x \neq 0 , \tag{3.26}$$

so that (3.20) is now

$$t = k\left(\frac{n+3}{2}\right)^{1/2}\int^{x}(1 + x'^2)^{(n-3)/4} dx' . \tag{3.27}$$

This integral may be evaluated in the case where n is an odd integer, but, with the exception of $n = 3$, it is not possible to invert the result to obtain x and so u as an explicit function of t.

Note that, if in (3.22) K_1 and K_2 are set equal to zero, $\alpha(t) = 3/t$ and, if $K_2 = -K_1^2$ and the limit $K_1 \to \infty$ is taken, $\alpha(t) = 3/2t$. These are the results obtained by Moreira for the case $n = 3$.

IV. THE CASE $n = 2$

We recall the equations for this case [(2.23) – (2.25)], viz.,

$$\frac{1}{3}\frac{d}{dt}(\sigma A) + \frac{1}{2}(\dot{\sigma}A - \sigma\dot{A}) = 0 , \tag{4.1}$$

$$\frac{1}{4}\frac{d}{dt}\left\{A\frac{d}{dt}\left[\frac{1}{A}(\dot{\sigma}A - \sigma\dot{A})\right]\right\} + A\gamma = 0 , \tag{4.2}$$

$$\frac{d}{dt}(A\dot{\gamma}) = 0 . \tag{4.3}$$

We observe that, when $\gamma \equiv 0$, we obtain the same parametric forms, (4.3), (3.19), and (3.20), defining $\alpha(t)$ as for the general case. However, we proceed to consider the possibilities of richer results in the case $\gamma \not\equiv 0$.

From (4.1),

$$\sigma^5 = KA , \tag{4.4}$$

so that (4.2) and (4.3) become

$$\frac{d}{dt}(\sigma^5\ddot{\sigma}) - \sigma^5\gamma = 0 , \tag{4.5}$$

$$\frac{d}{dt}(\sigma^5\dot{\gamma}) = 0 . \tag{4.6}$$

Substituting for γ from (4.5) into (4.6) and integrating once, we have

$$\frac{d^2}{dt^2}(\sigma^5\ddot{\sigma}) - 5\frac{\dot{\sigma}}{\sigma}\frac{d}{dt}(\sigma^5\ddot{\sigma}) = M . \tag{4.7}$$

It has not been possible to make any progress on the general solution of (4.7). However, a number of particular solutions are available. These are, for $M = 0$,

$$\sigma(t) = K_1 t + K_2 , \tag{4.8}$$

$$\sigma(t) = (K_1 t + K_2)^3 , \tag{4.9}$$

$$\sigma(t) = (K_1 t + K_2)^{1/3} , \tag{4.10}$$

and for $M \neq 0$,

$$\sigma(t) = (K_1 t + K_2)^{2/3} , \tag{4.11}$$

for which the value of M is given by

$$M = \tfrac{28}{27}K_1^4 . \tag{4.12}$$

The corresponding solutions for α and γ are

$$\alpha(t) = 5K_1/(K_1 t + K_2), \quad \gamma(t) = 0 , \tag{4.13}$$

$$\alpha(t) = 15K_1/(K_1 t + K_2), \quad \gamma(t) = 96K_1^3 , \tag{4.14}$$

$$\alpha(t) = 5K_1/3(K_1 t + K_2), \quad \gamma(t) = 0 , \tag{4.15}$$

$$\alpha(t) = \frac{10K_1}{3(K_1 t + K_2)}, \quad \gamma(t) = -\frac{4K_1^3}{9(K_1 t + K_2)^{7/3}} . \tag{4.16}$$

The solutions of (4.13) and (4.15) for $\alpha(t)$ are the special solutions obtained from the integration of (3.9) with C_1 and C_2 zero, respectively. The corresponding first integrals are

$$I = \tfrac{1}{2}(t + K)^6\dot{q}^2 + 2(t + K)^5 q\dot{q} + \tfrac{1}{3}(t + K)^6 q^3 , \tag{4.17}$$

for

$$\ddot{q} + 5\dot{q}/(t + K) + q^2 = 0 , \tag{4.18}$$

$$I = \tfrac{1}{2}(t + K)^{18}\dot{q}^2 + 6(t + K)^{17}q\dot{q} + \tfrac{1}{3}(t + K)^{18}q^3 - 6(t + K)^{16}q^2 - 96(t + K)^{15}\dot{q} , \tag{4.19}$$

for

$$\ddot{q} + 15\dot{q}/(t + K) + q^2 = 0 , \tag{4.20}$$

$$I = \tfrac{1}{2}(t + K)^2\dot{q}^2 + \tfrac{2}{3}(t + K)q\dot{q} + \tfrac{1}{3}(t + K)^2 q^3 + \tfrac{2}{9}q^2 , \tag{4.21}$$

for

$$\ddot{q} + 5\dot{q}/3(t + K) + q^2 = 0 \tag{4.22}$$

and

$$I = \tfrac{1}{2}(t + K)^4\dot{q}^2 + \tfrac{4}{3}(t + K)^3 q\dot{q} + \tfrac{1}{3}(t + K)^4 q^3 + \tfrac{2}{9}(t + K)^2 q^2 + \tfrac{4}{9}(t + K)\dot{q} + \tfrac{28}{27}q , \tag{4.23}$$

for

$$\ddot{q} + 10\dot{q}/3(t + K) + q^2 = 0 . \tag{4.24}$$

The results given in (4.17) and (4.22) are the same as those obtained by Moreira[4] when n is replaced by 2 in his general formula.

V. CONCLUSION

We have seen that the differential equation

$$\ddot{q} + \alpha(t)\dot{q} + q^n = 0$$

possesses first integrals for more general $\alpha(t)$ than has previously been reported. However, it is only in the case $n = 3$ that it is possible to write the most general $\alpha(t)$ as an explicit function of time. For other $n(\neq 2)$ it is possible only to define $\alpha(t)$ parametrically. The case $n = 2$ is especially difficult and only four particular functions $\alpha(t)$ have been recognized.

In an attempt to avoid the impasse covered by the general nonintegrability of (3.27) we examined the differential equation for $\alpha(t)$, (3.7), for Lie symmetries. For $n \neq 3$ two generators of symmetry were found. Unfortunately they yielded the same information as is already contained in (3.9) and (3.20) and so added nothing to what was already known. The integrable case, $n = 3$, was found to have eight generators of symmetry. Inasmuch as (3.7) is rather nonlinear and the maximum number of generators for a second-order ordinary differential equation is eight, this result was unexpected, although a similar result has been observed before.[11]

ACKNOWLEDGMENTS

We express our appreciation to I.de C. Moreira and H. R. Lewis for supplying the preprints referenced as 4 and 6, respectively. It is a pleasure to thank the referee for making constructive suggestions and for supplying the contents of the Appendix.

APPENDIX: THE CASE $n = 1$

The referee has kindly supplied the results for the case $n = -1$, i.e., for the equation

$$\ddot{q} + \alpha(t)\dot{q} + 1/q = 0 . \tag{A1}$$

These cases emerge. For

$$\alpha(t) = K , \tag{A2}$$

where K is constant,

$$I = \tfrac{1}{2}\dot{q}^2 + Kq\dot{q} + \log q + \tfrac{1}{2}K^2q^2 + Kt . \tag{A3}$$

For

$$\alpha(t) = K\tan(Kt + M) , \tag{A4}$$

where K and M are constants (either both real or both purely imaginary),

$$I = \tfrac{1}{2}\dot{q}^2 + K\tan(Kt + M)q\dot{q} - \tfrac{1}{2}K^2q^2 + \log q - \log\cos(K + M) . \tag{A5}$$

Finally, for

$$\alpha(t) = -1/(t + M) , \tag{A6}$$

where, again, M is a constant,

$$I = \tfrac{1}{2}\dot{q} - q\dot{q}/(t + M) + \log q - \log(t + M) . \tag{A7}$$

[1]S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (Dover, New York, 1957), Chap. IV.
[2]J. D. Logan, *Invariant Variational Principles* (Academic, New York, 1977), p. 52.
[3]W. Sarlet and L. Y. Bahar, "A direct construction of first integrals for certain non-linear dynamical systems," Int. J. Non-Linear Mech. **15**, 133 (1980).
[4]I. de C. Moreira, Comments on "A direct approach to finding exact invariants for one-dimensional time-dependent classical Hamiltonians," Research report IF/UFRJ/83/25, Universidade Federal do Rio de Janeiro, Instituto de Física, Cidade Universitária, Ilha do Fundão, Rio de Janeiro, Brasil.
[5]H. R. Lewis and P. G. L. Leach, "A direct approach to finding exact invariants for one-dimensional time-dependent classical Hamiltonians," J. Math. Phys. **23**, 2371 (1982).
[6]M. Feix and H. R. Lewis, "Invariants for dissipative non-linear systems by using rescaling," Research report LA-UR-83-3660, Los Alamos National Laboratory, Los Alamos, New Mexico.
[7]A. Munier, J. R. Burgan, M. R. Feix, and E. Fijalkow, "Schrödinger equation with time-dependent boundary conditions," J. Math. Phys. **22**, 1219 (1981).
[8]W. Sarlet and F. Cantrijn, "Generalizations of Noether's theorem in classical mechanics," SIAM Rev. **23**, 467 (1981).
[9]For example: P. G. L. Leach, "Note on the time-dependent damped and forced Harmonic oscillator," Am. J. Phys. **46**, 1247 (1979).
[10]E. Kamke, *Differentialgleichungen Lösungsmethoden und Lösungen* (Chelsea, New York, 1971).
[11]J. L. Reid, on the symmetries of the differential equation
$$y'' + 4y' + 2Ky^{-5/2}\exp\left\{-\frac{1}{2}k\int^x y^{-1}(x')\,dx'\right\} \times (M\sin x + N\cos x) = 0$$
(private communication dated 20 August 1981).

Compact quantum systems: Internal geometry of relativistic systems

A. O. Barut and A. J. Bracken[a)]
Department of Physics, University of Colorado, Boulder, Colorado 80309

(Received 21 February 1985; accepted for publication 31 May 1985)

A generalization is presented of the kinematical algebra so(5), shown previously to be relevant for the description of the internal dynamics (*Zitterbewegung*) of Dirac's electron. The algebra so($n+2$) is proposed for the case of a compact quantum system with n degrees of freedom. Associated wave equations follow from boosting these compact quantum systems. There exists a contraction to the kinematical algebra of a system with n degrees of freedom of the usual type, by which the commutation relations between n coordinate operators Q_i and corresponding momentum operators P_i, occurring within the so($n+2$) algebra, go over into the usual canonical commutation relations. The so($n+2$) algebra is contrasted with the sl(l,n) superalgebra introduced recently by Palev in a similar context: because so($n+2$) has spinor representations, its use allows the possibility of interpreting the half-integral spin in terms of the angular momentum of internal finite quantum systems. Connection is made with the ideas of Weyl on the possible use in quantum mechanics of ray representation of finite Abelian groups, and so also with other recent works on finite quantum systems. Possible directions of future research are indicated.

I. INTRODUCTION

Many years ago, Weyl[1] considered the unitary representation of the Lie group defined by Heisenberg's canonical commutation relations, and noted that it may also be considered as a ray representation of an infinite Abelian group. He speculated that unitary ray representations of finite Abelian groups might also prove important in quantum mechanics. Indeed, he gave the example of the unitary ray representation

$$g_1 \to i\sigma_1, \quad g_2 \to i\sigma_2, \quad g_3 \to i\sigma_3, \quad e \to I_2,$$

of the four-element Abelian group (Klein four-group), whose elements satisfy

$$\begin{aligned} &(g_1)^2 = (g_2)^2 = (g_3)^2 = e \text{ (identity)},\\ &g_2g_3 = g_3g_2 = g_1, \quad g_3g_1 = g_1g_3 = g_2,\\ &g_1g_2 = g_2g_1 = g_3, \end{aligned} \tag{1.1}$$

in connection with the description of the electron's spin. (Here the σ_i are Pauli matrices.)

Recent interest in "finite quantum systems" has approached the subject in three essentially different ways.

Santhanam and co-workers[2] have proceeded directly from Weyl's position, writing the unitary ray representatives of finite Abelian groups in exponential form in order to define finite-dimensional Hermitian analogs of Heisenberg's position and momentum variables, satisfying modified commutation relations. A related approach has been adopted by Gudder and Naroditsky,[3] and also by Stovicek and Tolar.[4]

Palev[5] has considered a simple dynamical system, the isotropic harmonic oscillator in n dimensions, and adopted a noncanonical quantization (in the spirit of Wigner's[6] well-known work, but along different lines) in order to arrive at noncanonical position and momentum variables with finite-dimensional representations.

Our own work[7] and continuing interest in this area has stemmed from the observation that Dirac's equation for the electron may be regarded as providing the covariant description of a finite quantum oscillator—the *Zitterbewegung*. Associated with this equation, in the rest frame of the electron's center of mass (or in any fixed frame with definite center of mass momentum), are internal coordinates Q_i and momenta $P_i (i = 1,2,3)$, which satisfy noncanonical commutation relations and have a finite (four-) dimensional Hermitian representation. The kinematical algebra generated by these three Q's and P's under commutation is isomorphic to the Lie algebra so(5).

The authors mentioned above, together with many others (see Jagannathan[8] and Saavedra and Utreras[9] for references), have speculated on the possible utility of novel kinematics in the description of the *internal* dynamics of real systems, and in particular, of some relativistic "particles." However, the so(5) algebra has the important distinguishing feature that it is *known* to be relevant to an important, real relativistic physical system, because of its association with Dirac's equation.[7]

Therefore, the structure of this particular kinematical algebra, its relation to the Heisenberg algebra and to Weyl's ideas, and its generalization to the case of n degrees of freedom (that is, n Q's and n P's) are of particular interest. This interest is heightened by the thought that the heavy leptons μ and τ may represent excited states of an internal electron dynamics. Furthermore, we show elsewhere that the cases $n = 2$ and $n = 4$, respectively, arise in the description of the internal dynamics of the neutrino,[10] and of the electron in a proper time formalism.[11]

II. THE KINEMATICAL ALGEBRA SO($n+2$)

In the description of the *Zitterbewegung* of the Dirac electron in the rest frame of its center of mass,[7] the three Hermitian operators Q_i appear as the coordinate of the charge relative to the center of mass. The three Hermitian operators P_i have been introduced as the corresponding rela-

[a)] On leave from Department of Mathematics, University of Queensland, Australia.

 0022-2488/85/102515-05$02.50

tive momentum variables. Together they generate the so(5) kinematical algebra, with commutation relations

$$[Q_i,Q_j] = (i\lambda^2/\hbar)S_{ij}, \tag{2.1a}$$
$$[P_i,P_j] = (4i\hbar/\lambda^2)S_{ij}, \tag{2.1b}$$
$$[Q_i,P_j] = i\hbar\delta_{ij}J, \tag{2.1c}$$
$$[Q_i,S_{jk}] = i\hbar(\delta_{ik}Q_j - \delta_{ij}Q_k), \tag{2.1d}$$
$$[P_i,S_{jk}] = i\hbar(\delta_{ik}P_j - \delta_{ij}P_k), \tag{2.1e}$$
$$[Q_i,J] = (i\lambda^2/\hbar)P_i, \tag{2.1f}$$
$$[P_i,J] = (4i\hbar/\lambda^2)Q_i, \tag{2.1g}$$
$$[J,S_{ij}] = 0, \tag{2.1h}$$
$$[S_{ij},S_{kl}] = i\hbar(\delta_{ik}S_{jl} + \delta_{jl}S_{ik} - \delta_{jk}S_{il} - \delta_{il}S_{jk}). \tag{2.1i}$$

Here λ is a constant with the dimension of length. As has been emphasized before,[7] the appearance of at least one such constant is inevitable in any finite quantum system incorporating Hermitian coordinate variables, whose eigenvalues are necessarily discrete, with dimensions of length. In the application of the so(5) algebra to the internal dynamics of the electron, λ equals the Compton wavelength of that particle. Furthermore, in that application the operators of the algebra (2.1) can be expressed in terms of the more familiar Dirac matrices as

$$Q_i = \tfrac{1}{2}i\lambda\alpha_i\beta, \tag{2.2a}$$
$$P_i = (\hbar/\lambda)\alpha_i, \tag{2.2b}$$
$$J = -\beta, \tag{2.2c}$$

while S_{ij} is the usual spin tensor

$$S_{ij} = -\tfrac{1}{4}i\hbar[\alpha_i,\alpha_j] = \epsilon_{ijk}S_k. \tag{2.2d}$$

The relevant representation of so(5) is then the four-dimensional spinor representation, in which $J(=-\beta)$ is a traceless operator with unit square.

There is an obvious generalization of the algebra (2.1) to the case of n degrees of freedom: simply allow the indices there to run over 1,2,...,n instead of 1,2,3. Then the Lie algebra so($n+2$) is obtained. If one defines $J_{AB}(=-J_{BA}$, $A,B = 1,2,...,n+2)$ by setting $J_{ij} = S_{ij}/\hbar$, $J_{i,n+1} = \lambda^{-1}Q_i$, $J_{i,n+2} = (\lambda/2\hbar)P_i$, and $J_{n+1,n+2} = \frac{1}{2}J$, then the J_{AB} satisfy the so($n+2$) commutation relations in standard form

$$[J_{AB},J_{CD}] = i(\delta_{AC}J_{BD} + \delta_{BD}J_{AC} - \delta_{BC}J_{AD} - \delta_{AD}J_{BC}). \tag{2.3}$$

The fundamental spinor representations of so($n+2$), of dimension 2^p, are of particular interest. [Here $p = \frac{1}{2}(n+1)$ if n is odd, and $p = \frac{1}{2}n$ if n is even. In the latter case there are two inequivalent representations.] The relations of such representations to Clifford algebras, and associated anticommutation relations, are well known. Only in these representations does the operator J, which is traceless in every representation, have unit square, so that its eigenvalues are ± 1. Inspection of (2.1c) suggests that one is then, in an intuitive sense, as close as possible to the canonical commutation relations

$$[q_i,p_j] = i\hbar\delta_{ij}I, \tag{2.4}$$

where I is the unit operator. (Note that the commutator of any Q_i and P_j represented by finite matrices must be traceless.)

Various dynamics are possible within the framework of the so($n+2$) algebra, corresponding to various choices of Hamiltonian operator H in the enveloping algebra of the particular representation at hand. In the case $n=3$, when the fundamental (Dirac) spinor representation is chosen, the only true so(3) scalars available (as distinct from pseudoscalars) are $J(=-\beta)$ and I (identity). With H of the form $cI + d\beta$, where c and d are numbers with dimensions of energy, the commutation relations (2.1f) and (2.1g), together with Heisenberg's equation of motion

$$ih\dot{A} = [A,H], \quad \dot{A} = \frac{dA}{dt}, \tag{2.5}$$

imply

$$\dot{Q}_i = d(\lambda^2/\hbar^2)P_i, \quad \dot{P}_i = -(4d/\lambda^2)Q_i, \tag{2.6}$$

so that

$$\ddot{Q}_i = -(4d^2/\hbar^2)Q_i, \quad \ddot{P}_i = -(4d^2/\hbar^2)P_i. \tag{2.7}$$

Thus harmonic oscillator dynamics is singled out in this case.[7] This would not be true for other representations of so(5), nor for larger values of n, even in the fundamental spinor representations.

Nevertheless, because the constants h and λ are available, dimensionless creation and annihilation operators can *always* be defined, whatever the representation and whatever the dynamics, as

$$A_i = Q_i/\lambda + i(\lambda/2\hbar)P_i,$$
$$A_i^\dagger = Q_i/\lambda - i(\lambda/2\hbar)P_i, \quad i = 1,2,...,n. \tag{2.8}$$

The $A_i^\dagger$ is Hermitian conjugate to A_i, and relations (2.1) become

$$[A_i,A_j] = 0 = [A_i^\dagger,A_j^\dagger],$$
$$[A_i,A_j^\dagger] = \delta_{ij}J + (2i/\hbar)S_{ij}, \tag{2.9}$$
$$[A_i,J] = -2A_i, \quad [A_i^\dagger,J] = +2A_i^\dagger,$$

together with (2.1h) and (2.1i) and relations like (2.1d) and (2.1e), which express the n-vector nature of A_i and $A_i^\dagger$.

The relations (2.1) are also equivalent to

$$[[A_i,A_j^\dagger],A_k] = 2(\delta_{ij}A_k + \delta_{jk}A_i - \delta_{ik}A_j),$$
$$[[A_i,A_j^\dagger],A_k^\dagger] = 2(-\delta_{ij}A_k^\dagger + \delta_{jk}A_i^\dagger - \delta_{ik}A_j^\dagger),$$
$$[[A_i,A_j^\dagger],[A_k,A_l^\dagger]] \tag{2.10}$$
$$= 2(\delta_{ik}[A_l,A_j^\dagger] - \delta_{lj}[A_i,A_k^\dagger]$$
$$+ \delta_{jk}[A_i,A_l^\dagger] - \delta_{il}[A_k,A_j^\dagger])$$
$$[A_i,A_j] = 0 = [A_i^\dagger,A_j^\dagger],$$

in which form they show most clearly how these operators differ from the ones introduced for a finite quantum oscillator by Palev.[5] His operators satisfy

$$[\{A_i^\dagger,A_j\},A_k] = -\delta_{ik}A_j + \delta_{ij}A_k,$$
$$[\{A_I^\dagger,A_j\},A_k^\dagger] = \delta_{jk}A_i^\dagger - \delta_{ij}A_k^\dagger,$$
$$[\{A_i^\dagger,A_j\},\{A_k^\dagger,A_l\} = \delta_{jk}\{A_i^\dagger,A_l\} - \delta_{il}\}\{A_k^\dagger,A_j\}, \tag{2.11}$$
$$\{A_i,A_j\} = 0 = \{A_i^\dagger,A_j^\dagger\},$$

and define the Lie superalgebra sl(l,n). Like so($n+2$), this has infinitely many inequivalent irreducible Hermitian representations.

Palev considered his algebra as a dynamical algebra associated with a particular Hamiltonian

$$H = [\hbar\omega/(n-1)]\{A_i^\dagger, A_i\}, \tag{2.12}$$

for an isotropic oscillator. [Here the constant ω is introduced, with dimensions of (time)$^{-1}$, but Palev also needs to introduce a constant with dimensions of a length in order to define coordinate and momentum operators.] In contrast, we view the so($n+2$) algebra as kinematical. It always admits as a *particular* dynamics, that associated with the Hamiltonian

$$H = (\hbar\omega/2n)[A_i^\dagger, A_i], \tag{2.13}$$

which leads to the harmonic oscillator equations

$$\dot{A}_i = -i\omega A_i, \quad \dot{A}_i^\dagger = +i\omega A_i^\dagger, \tag{2.14}$$

or equivalently, to Eqs. (2.7) with

$$d = \tfrac{1}{2}\hbar\omega. \tag{2.15}$$

As already remarked, this is the only dynamics permitted in the case of the fundamental spinor representation of so(5) ($n=3$), when it is directly relevant to the description of the *Zitterbewegung* of the electron as a finite quantum oscillator.[7] No doubt Palev's algebra (without reference to ω) could also be viewed more widely as a kinematical algebra admitting a variety of representations, and a variety of dynamics in most representation.

Another important distinction between the so($n+2$) and sl(l,n) algebras relates to the representations of the so(n) subalgebra that can appear. This subalgebra is associated in both cases with the "angular momentum" of the finite quantum system. Since spinor representation of so($n+2$) are allowed (as for the electron), then spinor representations of the so(n) subalgebra can be accomodated. However, the so(n) subalgebra of sl(l,n) appears in the chain

$$\mathrm{so}(n) < \mathrm{sl}(n) < \mathrm{sl}(l,n), \tag{2.16}$$

and only tensor representations of so(n) appear in the representations of sl(n). Thus Palev's algebra can only describe finite quantum systems with integral angular momentum or spin.

We comment at the end about the noncompact versions of the so($n+2$) algebras.

III. RELATIONSHIP TO THE HEISENBERG ALGEBRA AND TO WEYL'S IDEA

The so($n+2$) algebra, which is of dimension $\frac{1}{2}(n+1)(n+2)$, is generated by the n Q's and n P's under commutation, as Eqs. (2.1) show. In contrast, n canonical Q's and P's generate the Heisenberg algebra, which is of the smaller dimension $(2n+1)$:

$$\begin{aligned} &[q_i q_j] = 0 = [p_i, p_i], \quad [q_i p_i] = i\hbar\,\delta_{ij} I, \\ &[I, q_i] = 0 = [I, p_i]. \end{aligned} \tag{3.1}$$

These may be compared with Eqs. (2.1a)–(2.1c), (2.1e), and (2.1f). However, it is more appropriate to compare the so($n+2$) algebra with the kinematical Lie algebra k_n, also of dimension $\frac{1}{2}(n+1)(n+2)$, obtained by extending the Heisenberg algebra by the algebra so(n) of rotations; introduce the $\frac{1}{2}n(n-1)$ so(n) (angular momentum) operators $1_{ij}(=-l_{ji}, i,j=1,2,\ldots,n)$ satisfying

$$\begin{aligned} &[q_i, l_{jk}] = i\hbar(\delta_{ik} q_j - \delta_{ij} q_k), \\ &[p_i, l_{jk}] = i\hbar(\delta_{ik} p_j - \delta_{ij} p_k), \quad [I, l_{ij}] = 0, \\ &[l_{ij}, l_{kl}] = ih\,(\delta_{ik} l_{jl} + \delta_{jl} l_{ik} - \delta_{jk} l_{il} - \delta_{il} l_{jk}), \end{aligned} \tag{3.2}$$

which may be compared with Eqs. (2.1d), (2.1e), (2.1h), and (2.1i). Any representation of the Heisenberg algebra can be extended to a representation of k_n by setting

$$l_{jk} = q_j p_k - q_k p_j. \tag{3.3}$$

However, there are also representations of k_n in which the relation (3.3) does not hold. We may always add one or more "spin terms" to the right-hand side of Eq. (3.3), thus ensuring in particular that spinor representations of so(n) can occur.

It is noteworthy that, although there is (up to equivalence) only one unitary representation of the (Weyl) group associated with the Heisenberg Lie algebra, by von Neumann's theorem, there are evidently infinitely many inequivalent unitary representations (with various spin content) of the group K_n whose Lie algebra is k_n. Corresponding to this in our case is the fact that there are infinitely many inequivalent unitary representations of the group SO($n+2$).

There is a contraction[12] from the algebra so($n+2$) to k_n; this emphasizes the naturalness of the choice of so($n+2$) as an appropriate kinematical algebra for finite quantum systems. To see this without going into details, define

$$\tilde{q}_i = \epsilon_1 Q_i, \quad \tilde{p}_i = \epsilon_2 P_i, \quad \tilde{I} = \epsilon_1\epsilon_2 J, \quad \tilde{l}_{ij} = S_{ij}, \tag{3.4}$$

with Q_i, p_i, etc., as in (2.1) and ϵ_1, ϵ_2 real parameters. Then

$$\begin{aligned} &[\tilde{q}_i, \tilde{q}_j] = i(\lambda^2/\hbar)(\epsilon_1)^2 l_{ij}, \\ &[\tilde{p}_i \tilde{p}_j] = (4i\hbar/\lambda^2)(\epsilon_2)^2 l_{ij}, \\ &[\tilde{q}_i, \tilde{p}_j] = i\hbar\delta_{ij}\tilde{I}, \\ &[\tilde{q}_i, \tilde{I}] = -i(\lambda^2/\hbar)(\epsilon_1)^2 \tilde{p}_i, \\ &[\tilde{p}_i, \tilde{I}] = (4i\hbar/\lambda^2)(\epsilon_2)^2 \tilde{q}_i, \end{aligned} \tag{3.5}$$

while the remaining relations are as in Eqs. (3.2), with q_i replacing q_i, etc. When ϵ_1 and ϵ_2 are set to zero, Eqs. (3.5) reduce to Eqs. (3.1). If ϵ_1 is set to zero but not ϵ_2 (or vice versa), the Lie algebra obtained can be seen to be that of the Euclidean group $E(n+1)$. (These cases correspond physically to an oscillator or free particle.) This indicates that the contraction from so($n+2$) to k_n can proceed in two stages, via $e(n+1)$ (and that there are two distinct routes along which this may be accomplished).

There is also a close relationship between the fundamental spinor representations of the so($n+2$) algebra, and unitary ray representations of finite Abelian groups, so that contact can be made with Weyl's idea,[1] and also the work of Santhanam,[2] mentioned in the Introduction. Consider, for example, the case $n=1$ (one Q and one P) and the fundamental spinor representation of so(3), which is two dimensional. We may take in this case

$$Q = \tfrac{1}{2}\lambda\sigma_1, \quad P = (\hbar/\lambda)\sigma_2, \tag{3.6}$$

where σ_1 and σ_2 are Pauli matrices. Then Eqs. (2.1) show

$$J = \sigma_3, \tag{3.7}$$

while there are no so(n) operators in this case. Define the unitary operators

$$A(\theta) = \exp(i(\theta/\lambda)Q), \quad B(\phi) = \exp(i(\phi\lambda P/2\hbar)). \tag{3.8}$$

Then

$$A(\pi) = i\sigma_1, \quad B(\pi) = i\sigma_2, \tag{3.9}$$

and it can be seen that $A(\pi)$ and $B(\pi)$ generate under multiplication the unitary ray representation of the four-element Abelian group defined by Eqs. (1.1). In contrast, the set of all unitary operators $A(\theta)$, $B(\phi)$, with $\theta,\phi\in[0,2\pi]$, generate under multiplication a two-valued representation of SO(3) [that is, a true representation of SU(2)].

Note that if we started with the unitary ray representation of the Abelian group, and hence with $A(\pi)$ and $B(\pi)$, we could define

$$Q = -(i\lambda/\pi)\log A(\pi), \quad P = -(2i\hbar/\lambda)\log B(\pi), \tag{3.10}$$

and recover the so(3) algebra generated under commutation by Q and P. On the other hand, if we started with a unitary representation of su(2), we would more naturally identify Q and P by setting

$$Q = -i\lambda \left.\frac{dA(\theta)}{d\theta}\right|_{\theta=0}, \quad P = -\frac{i\lambda}{2\hbar}\left.\frac{dB(\phi)}{d\phi}\right|_{\phi=0}. \tag{3.11}$$

IV. CONCLUDING REMARKS

Of various approaches to the description of a finite quantum system with n degrees of freedom, the one using the so($n+2$) kinematical algebra is distinguished primarily by the fact that it is known to be relevant to real relativistic systems.[7,10,11] Furthermore, it has been shown that there is a well-defined relationship between the so($n+2$) algebra and the kinematical algebra k_n of a system with n degrees of freedom of the usual (noncompact) type. This relationship is defined by a group contraction.

Of course, we do not claim that so($n+2$) is the only algebra which could have such a relationship with k_n. However, the existence of this relationship suggests the possibility of studying a class of finite quantum systems which are well-defined analogs of infinite quantum systems, and also the connection between the two, through the contraction process. One could start with the finite quantum oscillator, as in Eqs. (2.7), for example, but it would be interesting also to construct finite analogs of other well-known dynamical systems, such as the Kepler system, and to investigate their symmetry and dynamical algebras.

Another important distinguishing feature of the so($n+2$) algebra which has been emphasized above is the existence of spinor representations. This makes possible the "explanation" of the half-integral spin of "elementary" particles as the angular momentum of internal finite quantum systems. Such an idea dates back to Schrödinger's work on Dirac's electron,[13] and has been further brought out in our own recent efforts.[7]

Finite systems can be accommodated naturally in the vector space setting of quantum mechanics—we merely need to consider finite-dimensional subspaces of Hilbert space. On the other hand, one might suppose that they have no classical counterparts. That this is not necessarily the case is shown, for example, by the recent construction of a classical analog of Dirac's spinning electron.[14] (In this connection, we mention also the earlier work by Grossmann and Peres.[15])

There is clearly more to be done towards understanding the relationship of finite quantum systems to the more familiar dynamical systems of classical and quantum mechanics. The use of the so($n+2$) kinematical algebra defines a class of finite systems for which some possible directions of future research seem reasonably well defined.

Once the commutation relations of the internal dynamical variables have been recognized, we can also take the infinite-dimensional representations of the internal algebra so($n+2$). These then represent many-body systems with n degrees of freedom in the center of mass frame. Relativistic theories of composite atoms or hadrons,[16] or relativistic oscillator and rotator,[17] belong to this category. The boosting of such a system (i.e., induced representations of the Poincaré group) gives relativistic finite-component wave equations in the case of finite-dimensional representations, and infinite-component wave equations in the case of composite systems.

In the infinite-dimensional case one can use perhaps more appropriately the unitary representations of the noncompact form of the algebras so(p,q). The exact form of the noncompact form depends on the physical interpretation of the generators as Hermitian operators. For example, the so(3,2) form of so(5) has been used extensively.[11,16–18]

[1] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1950), pp. 272–280.
[2] T. S. Santhanam and A. R. Tekumalla, Found. Phys. **6**, 583 (1976); T. S. Santhanam, Found. Phys. **7**, 121 (1977); in *Uncertainty Principle and Foundations of Quantum Mechanics* edited by W. Price and S. S. Chissick (Wiley, New York, 1977), p. 227; in *Proceedings of the 1978 International Meeting on Frontiers of Physics*, edited by K. K. Phua, C. K. Chew, and Y. K. Lim (Singapore National Academy of Sciences, Singapore, 1978), p. 1167; R. Jagannathan, T. S. Santhanam, and R. Vasudevan, Int. J. Theor. Phys. **20**, 755 (1981); R. Jagannathan and T. S. Santhanam, Int. J. Theor. Phys. **21**, 351 (1982).
[3] S. Gudder and V. Naroditsky, Int. J. Theor. Phys. **20**, 619 (1981).
[4] P. Stovicek and J. Tolar, preprint IC/79/147, Trieste, International Centre for Theoretical Physics, 1979.
[5] T. D. Palev, J. Math. Phys. **23**, 1778 (1982); Czech. J. Phys. B **32**, 680 (1982).
[6] E. P. Wigner, Phys. Rev. **77**, 711 (1950).
[7] A. O. Barut and A. J. Bracken, in *Group Theoretical Methods in Physics, Lecture Notes in Physics*, Vol. 135, edited by K. B. Wolf (Springer, New York, 1980), p. 206; Phys. Rev. D **23**, 2454 (1981); Aust. J. Phys. **35**, 353 (1982); Phys. Rev. D **24**, 3333 (1981).
[8] R. Jagannathan, Int. J. Theor. Phys. **22**, 1105 (1983).
[9] I. Saavedra and C. Utreras, Phys. Lett. B **98**, 74 (1981).
[10] A. O. Barut, A. J. Bracken, and W. D. Thacker, Lett. Math. Phys. **8**, 472 (1984).
[11] A. O. Barut, and W. D. Thacker, Phys. Rev. D **31**, 1386, 2076 (1985).
[12] E. Inönü and E. P. Wigner, Proc. Natl. Acad. Sci. USA **39**, 510 (1953).
[13] E. Schrödinger, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl. **24**, 418 (1930).
[14] A. O. Barut and N. Zanghi, Phys. Rev. Lett. **52**, 2009 (1984).
[15] Z. Grossmann and A. Peres, Phys. Rev. **132**, 2346 (1963).
[16] A. O. Barut, "Dynamical groups for the motion of relativistic composite systems," in *Groups, Systems and Many-body Physics*, edited by P.

Kramer and M. Dal Cin (Vieweg, Braunschweig, 1980), pp. 285–313.
[17]R. R. Aldinger, A. Bohm, P. Kielanowski, M. Loewe, P. Magnollay, N. Mukunda, W. Drechsler, and S. R. Komy, Phys. Rev. D **28**, 3020 (1983); A. Bohm, M. Loewe, L. C. Biedenharn, and H. van Dam, Phys. Rev. D **28**, 3030 (1983).
[18]E. Celeghini and M. Tarlini, Nuovo Cimento **68** B, 133 (1982); A. Bohm, M. Loewe, and P. Magnollay, "The quantum relativistic osscillator I, II," University of Texas preprint.

Deficiency indices and singular boundary conditions in quantum mechanics

W. Bulla
Institut für Theoretische Physik, Technische Universität Graz, A-8010 Graz, Austria

F. Gesztesy[a)]
Zentrum für Interdisziplinäre Forschung, Universität Bielefeld, D-4800 Bielefeld 1, Federal Republic of Germany and Centre de Physique Théorique, Centre National de la Recherche Scientifique, Centre de Luminy, F-13288 Marseille, Cedex 9, France and Université de Provence, U. E. R. de Physique, F-13288 Marseille, Cedex 9, France

(Received 15 January 1985; accepted for publication 12 April 1985)

We consider Schrödinger operators H in $L^2(\mathbb{R}^n)$, $n \in \mathbb{N}$, with countably infinitely many local singularities of the potential which are separated from each other by a positive distance. It is proved that due to locality each singularity yields a separate contribution to the deficiency index of H. In the special case where the singularities are pointlike and the potential exhibits certain symmetries near these points we give an explicit construction of self-adjoint boundary conditions.

I. INTRODUCTION

Our interest in the computation of deficiency indices and in the construction of self-adjoint boundary conditions for singular Schrödinger operators stems from several investigations of certain idealized model Hamiltonians, so-called point interactions.[1-10]

These analytically solvable models have a long history and play an important role in nuclear and solid state physics (cf., e.g., Ref. 9 and the literature therein). In this paper we particularly study the mathematical structure behind point interactions and some of their generalizations (interactions concentrated on submanifolds).

In Sec II we consider Schrödinger operators H in $L^2(\mathbb{R}^n)$, $n \in \mathbb{N}$ with countably infinitely many local singularities of the potential V which are uniformly separated from each other by a distance $\epsilon > 0$. Our main result (Theorem 2.5) concerning the deficiency index of H confirms the intuitive statement that due to locality each singularity should separately yield a contribution to the total deficiency index of H. Our proof is patterned after a result of Behncke[11] (c.f. also Ref. 12) where the corresponding problem is solved for strongly singular Dirac operators. Theorem 2.5 is general enough not only to include the case of point interactions in addition to V but also to allow additional interactions concentrated on submanifolds (like δ-shell interactions[8(b)]).

Section II represents the first step in the analysis, namely to reduce the computation of the deficiency indices of a Schrödinger operator H with several singularities to that of several Schrödinger operators H_j with a single singularity. The second step, the explicit construction of self-adjoint boundary conditions for H_j, is studied in Sec. III. In the special case where the singularity in H_j is pointlike and H_j exhibits certain symmetries around this point such that H_j reduces to a direct sum of ordinary Schrödinger operators in $L^2((0,\infty))$ (a case particularly important in applications) a general treatment of singular boundary conditions at the origin is presented. In particular, we study systems of the type

$$\frac{-d^2}{dr^2} + \frac{\lambda(\lambda-1)}{r^2} + \frac{\gamma}{r} + \frac{\alpha}{r^a} + W(r), \quad r > 0, \tag{1.1}$$

with $W \in L^\infty((0,\infty))$ real valued, $\frac{1}{2} \leq \lambda < \frac{3}{2}$, $\alpha, \gamma \in \mathbb{R}$, $0 < a < 2$.

Our methods rely heavily on the use of (ir)regular solutions associated with (1.1) and on corresponding Volterra integral equations. This yields a generalization of previous results of Rellich,[13] where the case $\alpha = 0$ in (1.1) has been considered.

II. DEFICIENCY INDICES OF SINGULAR SCHRÖDINGER OPERATORS

In this section we show that countably infinitely many local singularities of the potential which are uniformly separated from each other by a distance $\epsilon > 0$ do not interfere when considering the total deficiency index of the corresponding Schrödinger operator.

We introduce the following.

Hypothesis H: Let $J \subset \mathbb{Z} \setminus \{0\}$ be a finite or countably infinite index set, $J_0 = J \cup \{0\}$.

(i) $\Sigma_j \subset \mathbb{R}^n$, $n \in \mathbb{N}$ is a compact set of Lebesgue measure zero for all $j \in J$, $\Sigma_0 = \emptyset$.

(ii) $V_j \in L^2_{\text{loc}}(\mathbb{R}^n \setminus \Sigma_j)$ is real valued, $j \in J_0$, and (a) supp (V_j) is compact for all $j \in J$, or (b) V_j are bounded from below on every compact subset of $\mathbb{R}^n \setminus \Sigma_j$ for all $j \in J_0$.

(iii) For some $\epsilon > 0$: $\operatorname{dist}(\{\operatorname{supp}(V_j) \cup \Sigma_j\}, \{\operatorname{supp}(V_{j'}) \cup \Sigma_{j'}\}) \geq \epsilon$ for all $j, j' \in J_0$, $j \neq j'$.

(iv) $W \in L^\infty(\mathbb{R}^n)$ is real-valued.

For notational convenience we will also use the abbreviations

$$A_j = \begin{cases} \operatorname{supp}(V_j) \cup \Sigma_j \text{ if condition H(ii)(a) holds,} \\ \Sigma_j \text{ if condition H(ii)(b) or conditions} \\ \quad \text{H(ii)(a) and H(ii)(b) hold,} \quad j \in J_0, \end{cases} \tag{2.1}$$

$$A = \bigcup_{j \in J_0} A_j, \quad \Sigma = \bigcup_{j \in J_0} \Sigma_j, \quad V(x) = \sum_{j \in J_0} V_j(x),$$

and note that Σ is closed and of Lebesgue measure zero by hypotheses H(i) and H(iii).

As our first technical result we state the following.

Lemma 2.1: Assume conditions H(i) and H(iii). Then there exist $\phi_j, \tilde{\phi}_j \in C^\infty(\mathbb{R}^n)$, $j \in J_0$ such that we have the following.

(i) $\partial^\alpha \phi_j \in L^\infty(\mathbb{R}^n)$, $0 \leq |\alpha| \leq 2$, $\phi_j|_{A_j} = 1$, $j \in J_0$.

(ii) $\operatorname{supp}(\phi_j) \cap \operatorname{supp}(\tilde{\phi}_{j'}) = \emptyset$, $j, j' \in J_0$, $j \neq j'$.

[a)] On leave of absence from Institut für Theoretische Physik, Universität Graz, Austria.

 0022-2488/85/102520-09$02.50

(iii) For some $0<\delta<\epsilon/2$: $\operatorname{dist}(\operatorname{supp}(1-\phi_j),A_j)\geqslant\delta$, $j\in J_0$.

(iv) $\partial^\alpha\tilde{\phi}_j\in L^\infty(\mathbb{R}^n)$, $0\leqslant|\alpha|\leqslant 2$, $\tilde{\phi}_j|_{\operatorname{supp}(\phi_j)}=1$, $j\in J_0$.

(v) $\operatorname{supp}(\tilde{\phi}_j)\cap\operatorname{supp}(\tilde{\phi}_{j'})=\emptyset$, $j,j'\in J_0$, $j\neq j'$.

(vi) $\phi_0(x)=1$ for $x\in\{y\in\mathbb{R}^n\mid |y|\geqslant R\}$ for some $R>0$ if condition H(ii)(a) holds, and $\phi_j\in C_0^\infty(\mathbb{R}^n)$ and $\operatorname{dist}(\operatorname{supp}(\phi_j),\operatorname{supp}(V_{j'}))\geqslant\delta$ for all $j,j'\in J_0, j\neq j'$ if condition H(ii)(b) holds.

Proof: Fix $l\in J_0$ and define

$$U_{l,\eta}=\bigcup_{a\in A_l}S(a;\eta),\quad \eta>0. \tag{2.2}$$

[$S(x_0;R)$ is the open ball of radius R centered at x_0.] Then $U_{l,\epsilon/32}$ is an open neighborhood of A_l. If we introduce

$$E_{l,\epsilon/32}=\overline{U_{l,\epsilon/32}},\quad F_{l,\epsilon/16}=\mathbb{R}^n\backslash U_{l,\epsilon/16}, \tag{2.3}$$

then $E_{l,\epsilon/32}$ and $F_{l,\epsilon/16}$ are closed and disjoint and we can apply Corollary 1.4.11 of Ref. 14 to get the existence of $\phi_l\in C^\infty(\mathbb{R}^n)$ such that

$$\begin{aligned}&\phi_l|_{E_{l,\epsilon/32}}=1,\quad \phi_l|_{F_{l,\epsilon/16}}=0,\\ &\partial^\alpha\phi_l\in L^\infty(\mathbb{R}^n),\quad 0\leqslant|\alpha|\leqslant 2.\end{aligned} \tag{2.4}$$

The collection of all such ϕ_l, $l\in J_0$ obviously fulfills assertions (i)–(vi) (with $\delta\geqslant\epsilon/32$). For the construction of $\tilde{\phi}_j, j\in J_0$ one simply replaces ϵ by 3ϵ.

For the rest of this section ϕ_j (resp. $\tilde{\phi}_j$) always denote the $C^\infty(\mathbb{R}^n)$ functions of Lemma 2.1. Next we introduce the minimal Schrödinger operators

$$\dot{H}_j=-\Delta+V_j \text{ on } \mathcal{D}(\dot{H}_j)=C_0^\infty(\mathbb{R}^n\backslash\Sigma_j),\quad j\in J_0, \tag{2.5}$$

$$\dot{H}=-\Delta+V+W \text{ on } \mathcal{D}(\dot{H})=C_0^\infty(\mathbb{R}^n\backslash\Sigma), \tag{2.6}$$

and denote their closures by

$$H_j=\overline{\dot{H}_j},\quad H=\overline{\dot{H}}. \tag{2.7}$$

Due to hypotheses H(i), H(ii), and H(iv) the corresponding adjoint operators read

$$\begin{aligned}&H_j^* g=-\Delta g+V_j g \text{ in } C_0^\infty(\mathbb{R}^n\backslash\Sigma_j)',\\ &\text{for } g\in\mathcal{D}(H_j^*)=\{f\in L^2(\mathbb{R}^n)\mid -\Delta f\\ &\qquad +V_j f\in L^2(\mathbb{R}^n)\},\quad j\in J_0,\end{aligned} \tag{2.8}$$

$$\begin{aligned}&H^* g=-\Delta g+(V+W)g \text{ in } C_0^\infty(\mathbb{R}^n\backslash\Sigma)',\\ &\text{for } g\in\mathcal{D}(H^*)=\{f\in L^2(\mathbb{R}^n)\mid -\Delta f\\ &\qquad +(V+W)f\in L^2(\mathbb{R}^n)\}.\end{aligned} \tag{2.9}$$

We start our analysis with the following.

Lemma 2.2: Assume hypotheses H(i)–H(iv). Then, for all $j\in J_0$,

(i) $g\in\mathcal{D}(H_j^*)$ implies $\phi_j g\in\mathcal{D}(H_j^*)\cap\mathcal{D}(H^*)$, (2.10)

(ii) $g\in\mathcal{D}(H^*)$ implies $\phi_j g\in\mathcal{D}(H^*)\cap\mathcal{D}(H_j^*)$, (2.11)

and $H^*(\phi_j g)=H_j^*(\phi_j g)+W(\phi_j g)$ in both cases.

Proof: (a) Suppose condition H(ii)(a) to be valid. Then we first prove

$$\mathcal{D}(H_j^*)\subset H^{2,2}_{\mathrm{loc}}(\mathbb{R}^n\backslash\{\operatorname{supp}(V_j)\cup\Sigma_j\}),\quad j\in J_0, \tag{2.12}$$

$$\mathcal{D}(H^*)\subset H^{2,2}_{\mathrm{loc}}(\mathbb{R}^n\backslash\{\operatorname{supp}(V)\cup\Sigma\}). \tag{2.13}$$

Let $g\in\mathcal{D}(H_j^*)$, $\psi_j\in C_0^\infty(\mathbb{R}^n\backslash A_j)$. Then $H_j^* g=-\Delta g+V_j g\in L^2(\mathbb{R}^n)$ implies $\psi_j(H_j^* g)=-\psi_j(\Delta g)\in L^2(\mathbb{R}^n)$ and by the arbitrariness of ψ_j we infer $\Delta g\in L^2_{\mathrm{loc}}(\mathbb{R}^n\backslash A_j)$. By Theorem 1 of Ref. 15 we get $\nabla g\in L^2_{\mathrm{loc}}(\mathbb{R}^n\backslash A_j)$ and hence relation (2.12) results. Similarly $g\in\mathcal{D}(H^*)$, $\psi\in C_0^\infty(\mathbb{R}^n\backslash A)$ implies $\psi(H^* g)=-\psi(\Delta g)+\psi W g\in L^2(\mathbb{R}^n)$ and hence relation (2.13) follows. If condition H(ii)(b) holds then again $\nabla g\in L^2_{\mathrm{loc}}(\mathbb{R}^n\backslash A_j)$ for $g\in\mathcal{D}(H_j^*)$ and $\nabla g\in L^2_{\mathrm{loc}}(\mathbb{R}^n\backslash A)$ for $g\in\mathcal{D}(H^*)$ by Theorem 1 of Ref. 15.

(b) Let $g\in\mathcal{D}(H_j^*)$. Then

$$\begin{aligned}&(-\Delta+V_j+W)(\phi_j g)\\ &\quad=(-\Delta+V+W)(\phi_j g)\\ &\quad=\phi_j(-\Delta+V_j+W)g-2(\nabla\phi_j)(\nabla g)\\ &\qquad-(\Delta\phi_j)g\in L^2(\mathbb{R}^n),\end{aligned}$$

since ϕ_j, $\Delta\phi_j\in L^\infty(\mathbb{R}^n)$, $\nabla\phi_j\in C_0^\infty(\mathbb{R}^n\backslash A)$, and $\nabla g\in L^2_{\mathrm{loc}}(\mathbb{R}^n\backslash A_j)$ by (a).

(c) Let $g\in\mathcal{D}(H^*)$. Then

$$\begin{aligned}&(-\Delta+V+W)(\phi_j g)\\ &\quad=(-\Delta+V_j+W)(\phi_j g)\\ &\quad=\phi_j(-\Delta+V+W)g-2(\nabla\phi_j)(\nabla g)\\ &\qquad-(\Delta\phi_j)g\in L^2(\mathbb{R}^n),\end{aligned}$$

since ϕ_j, $\Delta\phi_j\in L^\infty(\mathbb{R}^n)$, $\nabla\phi_j\in C_0^\infty(\mathbb{R}^n\backslash A)$, and $\nabla g\in L^2_{\mathrm{loc}}(\mathbb{R}^n\backslash A)$ by (a). ∎

Lemma 2.3: Assume conditions H(i)–H(iv). Then, for all $j\in J_0$,

$$g\in\mathcal{D}(H) \text{ implies } \phi_j g\in\mathcal{D}(H)\cap\mathcal{D}(H_j) \tag{2.14}$$

and $H(\phi_j g)=H_j(\phi_j g)+W(\phi_j g)$.

Proof: Let $g\in\mathcal{D}(H)$. Then there exists a sequence

$$\{g_m\}_{m\in\mathbb{N}}\subset C_0^\infty(\mathbb{R}^n\backslash\Sigma),$$

such that $g_m\underset{m\to\infty}{\overset{s}{\to}}g$, $\dot{H}_{g_m}\underset{m\to\infty}{\overset{s}{\to}}Hg$.

Consequently, $\phi_j g_m\in C_0^\infty(\mathbb{R}^n\backslash\Sigma)$, $\phi_j g_m\underset{m\to\infty}{\overset{s}{\to}}\phi_j g$, and

$$\begin{aligned}\dot{H}(\phi_j g_m)&=(-\Delta+V_j+W)(\phi_j g_m)\\ &=\phi_j(-\Delta+V+W)g_m-2(\nabla\phi_j)(\nabla g_m)\\ &\quad-(\Delta\phi_j)g_m.\end{aligned} \tag{2.15}$$

Since

$$\phi_j(-\Delta+V+W)g_m\underset{m\to\infty}{\overset{s}{\to}}\phi_j Hg,\quad (\Delta\phi_j)g_m\underset{m\to\infty}{\overset{s}{\to}}(\Delta\phi_j)g, \tag{2.16}$$

it remains to consider the second term on the right-hand side of (2.15).

Let $\psi\in C_0^\infty(\mathbb{R}^n\backslash A)$ be real valued. Then $g\in H^{2,2}_{\mathrm{loc}}(\mathbb{R}^n\backslash A)$ implies

$$\begin{aligned}&\int_{\mathbb{R}^n}d^n x\,\psi^2|\nabla(g_m-g)|^2\\ &\quad=-2\int_{\mathbb{R}^n}d^n x\,\overline{(g_m-g)}(\nabla\psi)\psi\nabla(g_m-g)\\ &\qquad-\int_{\mathbb{R}^n}d^n x\,\overline{(g_m-g)}\psi^2\Delta(g_m-g),\end{aligned} \tag{2.17}$$

and thus

$$\|\psi|\nabla(g_m-g)|\,\|_2 \leqslant 2\|\,|\nabla\psi|\,\|_\infty \|g_m-g\|_2 \|\psi|\nabla(g_m-g)|\,\|_2 + \|g_m-g\|\,\|\psi^2\Delta(g_m-g)\|_2. \tag{2.18}$$

Inequality (2.18) proves $\|\psi|\nabla(g_m-g)|\,\|_2 \underset{m\to\infty}{\to} 0$. By taking $\psi_p^2 = |\partial_p\phi_j|^2$, $p=1,\ldots,n$ we infer $\|(\partial_p\phi_j)|\nabla(g_m-g)|\,\|_2 \underset{m\to\infty}{\to} 0$ and hence also

$$\|(\nabla\phi_j)\nabla(g_m-g)\|_2 \underset{m\to\infty}{\to} 0, \quad j\in J_0. \tag{2.19}$$

Consequently,

$$\dot H(\phi_j g_m) = \dot H_j(\phi_j g_m) + W(\phi_j g_m) \underset{m\to\infty}{\overset{s}{\to}} \phi_j H g - 2(\nabla\phi_j)\nabla g - (\Delta\phi_j)g. \tag{2.20}$$

Since H and H_j are closed we get from Eqs. (2.15), (2.16), and (2.20) $\phi_j g\in\mathscr{D}(H)\cap\mathscr{D}(H_j)$,

$$H(\phi_j g) = H_j(\phi_j g) + W(\phi_j g) = \phi_j H g - 2(\nabla\phi_j)(\nabla g) - (\Delta\phi_j)g. \quad \blacksquare$$

Lemma 2.4: Assume hypotheses H(i)–H(iv). Then, for all $j\in J_0$,

(i) $g\in\mathscr{D}(H_j^*)$ implies $(1-\phi_j)g\in\mathscr{D}(H_j)$, (2.21)

(ii) $g\in\mathscr{D}(H^*)$ implies $\left(1-\sum_{j\in J_0}\phi_j\right)g\in\mathscr{D}(H)$. (2.22)

Proof: (a) Let $g\in\mathscr{D}(H_j^*)$. Then $(1-\phi_j)g\in\mathscr{D}(H_j^*)$ by Lemma 2.2 (i). Denote $\psi_j = 1-\phi_j$: then $H_j^*(\psi_j g) = -\Delta(\psi_j g)$ in $C_0^\infty(\mathbb{R}^n\backslash\Sigma_j)'$ implies $-\Delta(\psi_j g)\in L^2(\mathbb{R}^n)$ and hence $\psi_j g\in H^{2,2}(\mathbb{R}^n)$. Let

$$0\leqslant\chi\in C_0^\infty(\mathbb{R}^n),\quad |\partial^\alpha\chi|\leqslant M<\infty,\quad 0\leqslant|\alpha|\leqslant 2,\quad \chi(x)=\begin{cases}1, & |x|\leqslant 1,\\ 0, & |x|\geqslant 2,\end{cases} \tag{2.23}$$

$$\chi_R(x)=\chi(xR^{-1}),\quad R>0,$$

and

$$0\leqslant j\in C_0^\infty(\mathbb{R}^n),\quad j(x)=0,\quad |x|\geqslant 1,\quad \int_{\mathbb{R}^n} d^n x\, j(x)=1, \tag{2.24}$$

$$j_\epsilon(x)=\epsilon^{-n}j(x/\epsilon),\quad \epsilon>0.$$

Then

$$\psi_j g\chi_R * j_{R^{-1}}\in C_0^\infty(\mathbb{R}^n\backslash A_j),\quad \text{for } R>0 \text{ large enough} \tag{2.25}$$

and

$$\psi_j g\chi_R * j_{R^{-1}} \underset{R\to\infty}{\overset{H^{2,2}(\mathbb{R}^n\backslash A_j)}{\to}} \psi_j g. \tag{2.26}$$

Thus, $\psi_j g\in H_0^{2,2}(\mathbb{R}^n\backslash A_j)$. Consequently there exists a sequence

$\{f_m\}_{m\in\mathbb{N}}\subset C_0^\infty(\mathbb{R}^n\backslash A_j)$, such that

$f_m \underset{m\to\infty}{\to} \psi_j g$ in $H^{2,2}(\mathbb{R}^n\backslash A_j)$ norm. This proves

$$\dot H_j f_m = -\Delta f_m \underset{m\to\infty}{\overset{s}{\to}} -\Delta(\psi_j g),$$

and hence

$$\psi_j g\in\mathscr{D}(H_j) \quad\text{and}\quad H_j(\psi_j g) = -\Delta(\psi_j g), \tag{2.27}$$

since H_j is closed.

(b) Let $g\in\mathscr{D}(H^*)$, $\Psi = 1-\Sigma_{j\in J_0}\phi_j$. Then $\Psi g\in H_0^{2,2}(\mathbb{R}^n\backslash A)$ as above and (2.22) follows. $\blacksquare$

Given Lemmas 2.2–2.4, we are able to state the main result of this section.

Theorem 2.5: Assume conditions H(i)–H(iv). Then

$$\operatorname{def}(H) = \sum_{j\in J_0}\operatorname{def}(H_j). \tag{2.28}$$

Proof: We first assume that $\operatorname{def}(H_j)<\infty$ for all $j\in J_0$.

(a) Let $\{\Phi_{jl}\in\mathscr{D}(H_j^*),\ 1\leqslant l\leqslant 2\operatorname{def}(H_j)\}$ be linearly independent modulo $\mathscr{D}(H_j)$, $j\in J_0$. (We note that the deficiency indices of H_j coincide since V_j is real valued.) By Lemma 2.2 (i) we get

$$\phi_j\Phi_{jl}\in\mathscr{D}(H_j^*)\cap\mathscr{D}(H^*),\quad 1\leqslant l\leqslant 2\operatorname{def}(H_j), j\in J_0. \tag{2.29}$$

Suppose there exist $\beta_{jl}\in\mathbb{C}$ such that

$$\sum_{j\in J_0}\sum_{l=1}^{2\operatorname{def}(H_j)}\beta_{jl}\,\phi_j\,\Phi_{jl}\in\mathscr{D}(H). \tag{2.30}$$

Then, by the locality of H,

$$\phi_j\sum_{l=1}^{2\operatorname{def}(H_j)}\beta_{jl}\,\Phi_{jl}\in\mathscr{D}(H),\quad j\in J_0, \tag{2.31}$$

since $\operatorname{supp}(\phi_j)\cap\operatorname{supp}(\phi_{j'})=\varnothing$ for $j,j'\in J_0, j\neq j'$. Now choose $\tilde\phi_j$ as in Lemma 2.1. Then Lemma 2.3 implies

$$\tilde\phi_j\left(\phi_j\sum_{l=1}^{2\operatorname{def}(H_j)}\beta_{jl}\,\Phi_{jl}\right) = \phi_j\sum_{l=1}^{2\operatorname{def}(H_j)}\beta_{jl}\,\Phi_{jl}\in\mathscr{D}(H_j),\quad j\in J_0. \tag{2.32}$$

On the other hand, from Lemma 2.4(i) we infer

$$(1-\phi_j)\sum_{l=1}^{2\operatorname{def}(H_j)}\beta_{jl}\,\Phi_{jl}\in\mathscr{D}(H_j),\quad j\in J_0, \tag{2.33}$$

and hence

$$\sum_{l=1}^{2\operatorname{def}(H_j)}\beta_{jl}\,\Phi_{jl}\in\mathscr{D}(H_j),\quad j\in J_0, \tag{2.34}$$

implying

$$\beta_{jl}=0,\quad 1\leqslant l\leqslant 2\operatorname{def}(H_j),\quad j\in J_0. \tag{2.35}$$

Thus, because of (2.31), $\{\phi_j\,\Phi_{jl},\ 1\leqslant l\leqslant 2\operatorname{def}(H_j), j\in J_0\}$ are linearly independent modulo $\mathscr{D}(H)$. Consequently,

$$2\operatorname{def}(H) = \dim\mathscr{D}(H^*)/\mathscr{D}(H)\geqslant 2\sum_{j\in J_0}\operatorname{def}(H_j). \tag{2.36}$$

(b) Conversely, let $\{\Psi_p\in\mathscr{D}(H^*),\ 1\leqslant p\leqslant 2\operatorname{def}(H)\}$ be linearly independent modulo $\mathscr{D}(H)$. By Lemma 2.4(ii) we get

$$\left(1-\sum_{j\in J_0}\phi_j\right)\Psi_p\in\mathscr{D}(H), \tag{2.37}$$

i.e.,

$$\Psi_p = \left(\sum_{j\in J_0}\phi_j\right)\Psi_p \quad\text{modulo } \mathscr{D}(H). \tag{2.38}$$

By Lemma 2.2(ii)

$$\phi_j \Psi_p \in \mathscr{D}(H^*) \cap \mathscr{D}(H_j^*), \quad j \in J_0, \tag{2.39}$$

and thus $\phi_j \Psi_p$ can be written

$$\phi_j \Psi_p = \Phi_{j0} + \sum_{l=1}^{2\,\mathrm{def}(H_j)} C_{pjl} \Phi_{jl}, \quad \Phi_{j0} \in \mathscr{D}(H_j),$$
$$C_{pjl} \in \mathbb{C}, \quad 1 \leqslant p \leqslant 2\,\mathrm{def}(H), \quad j \in J_0 \tag{2.40}$$

[i.e., $\phi_j \Psi_p = \Sigma_{l=1}^{2\,\mathrm{def}(H_j)} C_{pjl} \Phi_{jl}$ modulo $\mathscr{D}(H_j)$]. Since $1 \leqslant p \leqslant 2\,\mathrm{def}(H)$, we get from Eq. (2.38)

$$2\,\mathrm{def}(H) \leqslant 2 \sum_{j \in J_0} \mathrm{def}(H_j). \tag{2.41}$$

If $\mathrm{def}(H_j) = \infty$ for some $j \in J_0$ then it suffices to follow part (a) in order to conclude $\mathrm{def}(H) = \infty$. ■

For the rest of this section we discuss hypotheses H(i)–H(iv) and sketch possible generalizations. We start with two examples which clearly demonstrate the range of applicability of Theorem 2.5.

Example 2.6: Let

$$V(x) + W(x) = \sum_{j=1}^{N} C_j^{(1)} |x - y_j|^{-\alpha_j} + \sum_{j=1}^{N} C_j^{(2)} |x - z_j|^{-\beta_j} \times (x - z_j) \cdot e_j + \sum_{j=1}^{N} C_j^{(3)} \| x - x_j| - R_j|^{-\gamma_j} + \sum_{j=1}^{N} C_j^{(4)} x^j, \tag{2.42}$$

where

$$C_j^{(l)} \in \mathbb{R}, \quad l = 1,...,4, \quad e_j \in \mathbb{R}^n, \quad |e_j| = 1,$$
$$x_j, y_j, z_j \in \mathbb{R}^n, \quad R_j > 0,$$
$$\{x \in \mathbb{R}^n | \ |x - x_j| \leqslant R_j\} \cap \{x \in \mathbb{R}^n | \ |x - x_{j'}| \leqslant R_{j'}\} = \varnothing, \quad j \neq j', \tag{2.43}$$
$$\{y_j, z_j, j = 1,...,N\} \cap \bigcup_{j'=1}^{N} \{x \in \mathbb{R}^n | \ |x - x_{j'}| = R_{j'}\} = \varnothing,$$
$$\alpha_j \geqslant 0, \quad \beta_j \geqslant 0, \quad \gamma_j \geqslant 0, \quad j, j' = 1,...,N, \quad N \in \mathbb{N}.$$

Thus one may choose

$$\Sigma = \bigcup_{j=1}^{N} (\{y_j\} \cup \{z_j\} \cup \{x \in \mathbb{R}^n | \ |x - x_j| = R_j\}). \tag{2.44}$$

Example 2.7: Let

$$V(x) + W(x) = \prod_{m=1}^{n} |\sin(\mu_m x_{(m)})|^{-\nu_m}, \tag{2.45}$$

where

$$x = (x_{(1)},...,x_{(n)}) \in \mathbb{R}^n, \quad \mu_m > 0, \quad \nu_m \geqslant 0, \quad m = 1,...,n, \tag{2.46}$$

and hence we choose

$$\Sigma = \left\{ \left(\frac{\pi p_1}{\mu_1},..., \frac{\pi p_n}{\mu_n} \right) \in \mathbb{R}^n | p_m \in \mathbb{Z}, m = 1,...,n \right\}. \tag{2.47}$$

Remark 2.8: The strategy in the proof of Theorem 2.5 is taken from that of Theorem 1(1) in Behncke[11] (cf. also Ref. 12), where the case of strongly singular Dirac operators (including, e.g., the anomalous magnetic moment term) is treated. In particular, Behncke discussed the case where the Σ_j are disjoint finite sets and also derived the invariance of essential spectra. For related results in the context of distinguished self-adjoint extensions for the Dirac operator with a potential dominated by multicenter Coulomb potentials we refer to Refs. 16.

The main ingredients for Theorem 2.5 are obviously relations (2.10), (2.11), (2.14), (2.21), and (2.22). [We also note that the existence of $\phi_j, \tilde{\phi}_j$ in Lemma 2.1 is clearly independent of the fact whether $A_j, j \in J$ are compact or not; only the fact that $\mathrm{dist}(A_j, A_{j'}) \geqslant \epsilon, j \neq j'$ has been used.[14]] In particular, the main technical difference between the Schrödinger case presented above and the Dirac case in Ref. 11 now concerns the necessity to control ∇g for $g \in \mathscr{D}(H_j^*)$ or $g \in \mathscr{D}(H^*)$ [i.e., to prove $\nabla g \in L^2_{\mathrm{loc}}(\mathbb{R}^n \backslash A_j)$ or $\nabla g \in L^2_{\mathrm{loc}}(\mathbb{R}^n \backslash A)$]. For example, if $V_j \in \mathscr{Q}_{\alpha_j, \mathrm{loc}}(\mathbb{R}^n \backslash \{x_j\})$, $0 < \alpha_j \leqslant 1$ then $\mathscr{D}(H_j^*) \subset H^{2,2}_{\mathrm{loc}}(\mathbb{R}^n \backslash \{x_j\}) \cap L^2(\mathbb{R}^n)$ (Ref. 17) is obviously sufficient to prove Lemma 2.2–Theorem 2.5 (with $A_j = \Sigma_j = \{x_j\}$).

For related results on local properties of elements in $\mathscr{D}(H_j^*)$ we also refer to Pearson,[18] Combescure and Ginibre,[19] and Amrein.[20] Another particular important case where $\mathrm{def}(H_j) = \mathrm{def}(H) = 0$ appeared in Simader,[21] Brezis,[22] and Cycon.[23] Similarly, Morgan[24] used the idea of local partitions to prove stability of operator bounds and form bounds in the context of Schrödinger operators whose potentials have separated singularities. Finally, Svendsen[25] discussed the case where $A: C_0^\infty(\Omega, \mathbb{C}^s) \to L^2(\Omega, \mathbb{C}^s)$ is a linear symmetric differential operator with C^∞ coefficients, $\Omega \subset \mathbb{R}^n$ open, $n, s \in \mathbb{N}$. If M is a C^∞ manifold of Ω which is closed in Ω and has codimension greater than zero he studied the relation between the deficiency indices of A and $A|_{C_0^\infty(\Omega \backslash M, \mathbb{C}^s)}$.

We also mention the possibility of replacing $\mathbb{R}^n$ by $\Omega \subset \mathbb{R}^n$ open in the above treatment. The corresponding minimal operators are then given by $H_j = -\Delta + V_j$ on $C_0^\infty(\Omega \backslash \Sigma_j)$ and results on $\mathscr{D}(H_j^*)$ in this case may be found in Jörgens[26] and Kalf[15] and the references therein.

Theorem 2.5 relates the computation of $\mathrm{def}(H)$ to that of $\mathrm{def}(H_j), j \in J_0$. For the determination of deficiency indices of singular Schrödinger operators we refer to Piepenbrink and Rejto[27] and Behncke and Focke.[28] In the special case where $V_j(x) = V_j(|x - x_j|)$ is spherically symmetric [or H_j can be decomposed into a direct sum of ordinary differential operators like in the case of

$$V_j(x) = c_j |x - x_j|^{-3} (x - x_j) \cdot e_j, \quad x_j, e_j \in \mathbb{R}^n, \quad |e_j| = 1, \quad c_j \in \mathbb{R};$$

cf. Ref. 29], numerous methods to calculate the deficiency indices of the underlying ordinary differential operators are known.[30–32] In the special case

$$\Sigma = \{x_1,...,x_N\}, \quad x_j \in \mathbb{R}^n, \quad j = 1,...,N, \quad N \in \mathbb{N},$$

the corresponding deficiency subspaces have been obtained by Zorbas[33] with the help of suitable Green's functions.

III. SINGULAR BOUNDARY CONDITIONS FOR ORDINARY DIFFERENTIAL OPERATORS

In Sec. II we indicated how to reduce the computation of $\mathrm{def}(H)$ to that of $\mathrm{def}(H_j)$. In the special case where H_j can be decomposed into a direct sum of ordinary differential operators on $(0, \infty)$ [e.g.,

$$V_j(x) = v_j(|x - x_j|)$$

or

$$V_j(x) = c_j|x - x_j|^{-3}(x - x_j)\cdot e_j,$$
$$x_j, e_j \in \mathbb{R}^n, \quad |e_j| = 1, \quad c_j \in \mathbb{R}, \quad n \in N\,],$$

we now discuss possible self-adjoint boundary conditions at the origin. More precisely, we consider in $L^2((0,\infty))$ the minimal operator

$$\dot{h} = -\frac{d^2}{dr^2} + \frac{\lambda(\lambda-1)}{r^2} + \frac{\gamma}{r} + \frac{\alpha}{r^a} + W$$
$$\text{on } \mathscr{D}(\dot{h}) = C_0^\infty((0,\infty)),$$
$$W \in L^\infty((0,\infty)) \text{ real valued}, \quad \alpha,\gamma \in \mathbb{R},$$
$$0 < a < 2, \quad \tfrac{1}{2} \leqslant \lambda < \tfrac{3}{2}. \tag{3.1}$$

As has been discussed in Refs. 2–10 and 34, self-adjoint extensions of $\dot{h}$ different from its Friedrichs extension correspond to a $\lambda(\lambda-1)r^{-2} + \gamma r^{-1} + \alpha r^{-a} + W +$ "point interaction." It is the purpose of this section to construct all self-adjoint extensions of $\dot{h}$

Due to our conditions on λ, the closure of $\dot{h}$, denoted by h, is bounded from below and has $\operatorname{def}(h) = 1$. In order to determine explicitly the one-parameter family of self-adjoint extensions of h we shall study solutions of the equation

$$-\psi''(r) + [\lambda(\lambda-1)r^{-2} + V(r)]\psi(r) = 0, \quad r > 0, \tag{3.2}$$

where

$$V(r) = \gamma r^{-1} + \alpha r^{-a} + W(r). \tag{3.3}$$

Let $F_\lambda(r)$ be the regular solution of Eq. (3.2), i.e.,

$$F_\lambda(r) = F_\lambda^{(0)}(r) - \int_0^r dr'\, g_\lambda^{(0)}(r,r')V(r')F_\lambda(r'), \tag{3.4}$$

where

$$g_\lambda^{(0)}(r,r') = G_\lambda^{(0)}(r)F_\lambda^{(0)}(r') - G_\lambda^{(0)}(r')F_\lambda^{(0)}(r), \tag{3.5}$$
$$F_\lambda^{(0)}(r) = r^\lambda,$$
$$G_\lambda^{(0)}(r) = \begin{cases} -r^{1/2}\ln r, & \lambda = \frac{1}{2}, \\ (2\lambda-1)^{-1}r^{1-\lambda}, & \frac{1}{2} < \lambda < \frac{3}{2}. \end{cases} \tag{3.6}$$

We also note that $[G_\lambda^{(0)}, F_\lambda^{(0)}]_r = 1$, where $[g,f]_r = (\bar{g}, f' - \bar{g}'f)(r)$ denotes the Wronskian of g and f. Since $\int_0^R dr\, r|V(r)| < \infty$ for any $0 < R < \infty$ (needed in the case $\frac{1}{2} < \lambda < \frac{3}{2}$) as well as $\int_0^{r_0} dr\, r|\ln r|\,|V(r)| < \infty$ for all $0 < r_0 < 1$ (needed for $\lambda = \frac{1}{2}$), we may iterate Eq. (3.4) to get[35,36]

$$|F_\lambda(r)| \leqslant r^\lambda \begin{cases} \exp\left[\int_0^r dr'\, r'|V(r')|\right], & \frac{1}{2} < \lambda < \frac{3}{2}, \\ \exp\left[\int_0^r dr'\, r'|\ln r'|\,|V(r')|\right], & \lambda = \frac{1}{2}, \quad r \leqslant r_0, \\ \leqslant cr^\lambda. \end{cases} \tag{3.7}$$

Similarly, if

$$F_\lambda(r) = \sum_{m=0}^\infty F_\lambda^{(m)}(r) \tag{3.8}$$

denotes the absolute convergent series obtained by iterating Eq. (3.4), then

$$\left|F_\lambda(r) - \sum_{m=0}^N F_\lambda^{(m)}(r)\right| = \left|\int_0^r dr_1\, g_\lambda^{(0)}(r,r_1)V(r_1)\int_0^{r_1} dr_2\, g_\lambda^{(0)}(r_1,r_2)V(r_2)\cdots\int_0^{r_N} dr_{N+1}\, g_\lambda^{(0)}(r_N,r_{N+1})V(r_{N+1})F_\lambda(r_{N+1})\right|$$
$$\leqslant c[(N+1)!]^{-1}r^\lambda \begin{cases} \left[\int_0^r dr'\, r'|\ln r'|\,|V(r')|\right]^{N+1}, & \lambda = \frac{1}{2}, \quad r \leqslant r_0, \\ \left(\frac{2}{2\lambda-1}\right)^{N+1}\left[\int_0^r dr'\, r'|V(r')|\right]^{N+1}, & \frac{1}{2} < \lambda < \frac{3}{2}. \end{cases} \tag{3.9}$$

Here the estimate

$$|g_\lambda^{(0)}(r,r')| \leqslant \begin{cases} (rr')^{1/2}|\ln r'|, & \lambda = \frac{1}{2}, \quad r' \leqslant r \leqslant r_0 < 1, \\ [2/(2\lambda-1)]r^\lambda r'^{1-\lambda}, & \frac{1}{2} < \lambda < \frac{3}{2}, \end{cases} \tag{3.10}$$

has been used. Moreover,

$$\left|\int_0^r dr'\, g_\lambda^{(0)}(r,r')V(r')F_\lambda(r')\right|$$
$$\leqslant cr^\lambda \begin{cases} \int_0^r dr'\, r'|\ln r'|\,|V(r')|, & \lambda = \frac{1}{2}, \quad r \leqslant r_0, \\ \int_0^r dr'\, r'|V(r')|, & \frac{1}{2} < \lambda < \frac{3}{2}, \end{cases} \tag{3.11}$$

using (3.4) and (3.6), shows that

$$F_\lambda(r) = F_\lambda^{(0)}(r)[1 + \widehat{F}_\lambda(r)], \tag{3.12}$$

where

$$|\widehat{F}_\lambda(r)| \leqslant \begin{cases} c\int_0^r dr'\, r'|\ln r'|\,|V(r')|, & \lambda = \frac{1}{2}, \quad r \leqslant r_0, \\ c\int_0^r dr'\, r'|V(r')|, & \frac{1}{2} < \lambda < \frac{3}{2}, \end{cases} \tag{3.13}$$

and thus

$$\widehat{F}_\lambda(r) \underset{r\to 0_+}{=} o(1), \quad \tfrac{1}{2} \leqslant \lambda < \tfrac{3}{2}. \tag{3.14}$$

Consequently, we infer

$$|F_\lambda(r)| \leqslant c_1(r_0)|F_\lambda^{(0)}(r)| = c_1(r_0)r^\lambda,$$
$$|F_\lambda(r)| \geqslant c_2(r_0)|F_\lambda^{(0)}(r)| = c_2(r_0)r^\lambda,$$
$$\tfrac{1}{2} \leqslant \lambda < \tfrac{3}{2}, \quad r \leqslant r_0. \tag{3.15}$$

Introducing the irregular solution $G_\lambda(r)$ associated with Eq. (3.2) by[37]

$$G_\lambda(r) = F_\lambda(r) \int_r^{r_0} dr' [F_\lambda(r')]^{-2}, \tag{3.16}$$

we obtain the bound

$$|G_\lambda(r)| \leqslant c_2(r_0)^{-2} c_1(r_0) r^\lambda \int_r^{r_0} dr'\, r'^{-2\lambda}$$

$$\leqslant c_3(r_0) \begin{cases} r^{1/2} |\ln r|, & \lambda = \frac{1}{2}, \\ r^{1-\lambda}, & \frac{1}{2} < \lambda < \frac{3}{2}; \quad r \leqslant r_0. \end{cases} \tag{3.17}$$

We also note that $[G_\lambda, F_\lambda]_r = 1$.

Given the above preliminaries we now derive all self-adjoint extensions of h. The adjoint operator of h reads

$$h^* = -\frac{d^2}{dr^2} + \lambda(\lambda-1) r^{-2} + V,$$

$$\mathscr{D}(h^*) = \{ g \in L^2((0,\infty)) | g, g' AC_{\text{loc}}((0,\infty)), \tag{3.18}$$

$$-g'' + \lambda(\lambda-1) r^{-2} g + Vg \in L^2((0,\infty))\},$$

and by the general theory of second-order ordinary differential operators all self-adjoint extensions h_ν of h are given by[30-32,38]

$$h_\nu = -\frac{d^2}{dr^2} + \lambda(\lambda-1) r^{-2} + V,$$

$$\mathscr{D}(h_\nu) = \{ g \in \mathscr{D}(h^*) | \lim_{r \to 0_+} [\phi_{\nu,\lambda}, g]_r = 0 \}, \tag{3.19}$$

$$-\infty < \nu \leqslant \infty,$$

where

$$\phi_{\nu,\lambda}(r) = G_\lambda(r) + \nu F_\lambda(r), \quad -\infty < \nu \leqslant \infty \tag{3.20}$$

[i.e., $\phi_{\infty,\lambda}(r) = F_\lambda(r)$ for $\nu = \infty$]. Since any $g \in \mathscr{D}(h^*)$ can be written as[38]

$$g(r) = c_1 F_\lambda(r) + c_2 G_\lambda(r) - F_\lambda(r) \int_\rho^r dr'\, G_\lambda(r')(h^* g)(r')$$

$$+ G_\lambda(r) \int_\rho^r dr'\, F_\lambda(r')(h^* g)(r'), \tag{3.21}$$

for some $\rho > 0$,

a straightforward computation shows

$$[\phi_{\nu,\lambda}, g]_r = c_1 - c_2 \nu - \nu \int_\rho^r dr'\, F_\lambda(r')(h^* g)(r')$$

$$- \int_\rho^r dr'\, G_\lambda(r')(h^* g)(r'). \tag{3.22}$$

Thus we obtain for $g \in \mathscr{D}(h_\nu)$

$$c_1 - \int_\rho^r dr'\, G_\lambda(r')(h^* g)(r')$$

$$= \nu \left[c_2 + \int_\rho^r dr'\, F_\lambda(r')(h^* g)(r') \right]. \tag{3.23}$$

From Eqs. (3.12) and (3.14) we get

$$\lim_{r \to 0_+} F_\lambda(r)/G_\lambda^{(0)}(r) = 0, \quad \lim_{r \to 0_+} F_\lambda(r)/F_\lambda^{(0)}(r) = 1,$$

$$\widetilde{F}_\lambda(r) \underset{r \to 0_+}{=} o(1), \tag{3.24}$$

where

$$\widetilde{F}_\lambda(r) = -1 + [1 + \widehat{F}_\lambda(r)]^{-1}. \tag{3.25}$$

Thus Eq. (3.16) implies

$$G_\lambda(r) = F_\lambda^{(0)}(r)(1 + \widehat{F}_\lambda(r)) \int_r^{r_0} dr' (F_\lambda^{(0)}(r'))^{-2} (1 + \widetilde{F}_\lambda(r'))^2$$

$$\underset{r \to 0_+}{=} G_\lambda^{(0)}(r) + o(G_\lambda^{(0)}(r)) \tag{3.26}$$

and hence

$$\lim_{r \to 0_+} G_\lambda(r)/G_\lambda^{(0)}(r) = 1. \tag{3.27}$$

Equations (3.24) and (3.27) together with Eq. (3.21) then prove

$$g_{0,\lambda} := \lim_{r \to 0_+} g(r)/G_\lambda^{(0)}(r)$$

$$= c_2 - \int_0^\rho dr'\, F_\lambda(r')(h^* g)(r'), \quad g \in \mathscr{D}(h^*). \tag{3.28}$$

If we insert the asymptotic expansion of $F_\lambda(r)$ as $r \to 0_+$ into Eq. (3.16) we get the corresponding expansion for $G_\lambda(r)$. Let $G_\lambda^B(r)$ denote the asymptotic expansion of $G_\lambda(r)$ up to the smallest order such that

$$\lim_{r \to 0_+} [G_\lambda(r) - G_\lambda^B(r)]/F_\lambda^{(0)}(r) = 0. \tag{3.29}$$

[A constructive approach to calculate $G_\lambda^B(r)$ will be given later on.] Then Eq. (3.21) implies

$$\frac{g(r)}{F_\lambda^{(0)}(r)} - g_{0,\lambda} \frac{G_\lambda^B(r)}{F_\lambda^{(0)}(r)}$$

$$= c_1 \frac{F_\lambda(r)}{F_\lambda^{(0)}(r)} + c_2 \frac{G_\lambda(r) - G_\lambda^B(r)}{F_\lambda^{(0)}(r)}$$

$$+ \frac{F_\lambda(r)}{F_\lambda^{(0)}(r)} \int_r^\rho dr'\, G_\lambda(r')(h^* g)(r')$$

$$- \frac{G_\lambda(r) - G_\lambda^B(r)}{F_\lambda^{(0)}(r)} \int_r^\rho dr'\, F_\lambda(r')(h^* g)(r')$$

$$+ \frac{G_\lambda^B(r)}{F_\lambda^{(0)}(r)} \int_0^r dr'\, F_\lambda(r')(h^* g)(r'). \tag{3.30}$$

Using Eqs. (3.15) and (3.17), one shows that

$$\left| \frac{G_\lambda^B(r)}{F_\lambda^{(0)}(r)} \int_0^r dr'\, F_\lambda(r')(h^* g)(r') \right|$$

$$\leqslant c_2(r_0) \left\{ \begin{matrix} r^{1/2} |\ln r|, & \lambda = \frac{1}{2} \\ r^{1-\lambda}, & \frac{1}{2} < \lambda < \frac{3}{2} \end{matrix} \right\} r^{-\lambda}$$

$$\times \left(\int_0^r dr' |F_\lambda(r')|^2 \right)^{1/2} \|h^* g\|_2$$

$$\leqslant \text{const} \begin{cases} r |\ln r|, & \lambda = \frac{1}{2}, \\ r^{3/2-\lambda}, & \frac{1}{2} < \lambda < \frac{3}{2}; \quad r \leqslant r_0. \end{cases} \tag{3.31}$$

Thus, we define

$$g_{1,\lambda} := \lim_{r \to 0_+} [\, g(r) - g_{0,\lambda} G_\lambda^B(r)]/F_\lambda^{(0)}(r), \quad g \in \mathscr{D}(h^*), \tag{3.32}$$

and obtain from Eqs. (3.30), (3.31), and the fact that h is in the limit circle case at the origin [i.e., $F_\lambda, G_\lambda \in L^2((0,R))$ for any $0<R<\infty$]

$$g_{1,\lambda} = c_1 + \int_0^\rho dr'\, G_\lambda(r')(h*g)(r'), \quad g \in \mathscr{D}(h^*). \tag{3.33}$$

The self-adjoint boundary condition (3.23) thus reads

$$\nu g_{0,\lambda} = g_{1,\lambda}, \quad -\infty < \nu \leqslant \infty. \tag{3.34}$$

It is not hard to see that $\nu = \infty$, i.e., $g_{0,\lambda} = 0$ corresponds to the Friedrichs extension of h. For $|\nu| < \infty$, h_ν describes a $\lambda(\lambda-1)r^{-2} + \gamma r^{-1} + \alpha r^{-a} + W +$ "point interaction."

It remains to determine $G_\lambda^B(r)$ explicitly. Since W has no influence on boundary conditions we simply put $W = 0$ from now on. As $G_\lambda^B(r)$ will be constructed with the help of Eq. (3.16), we first derive the asymptotic expansion of $F_\lambda(r)$ as $r \to 0_+$.

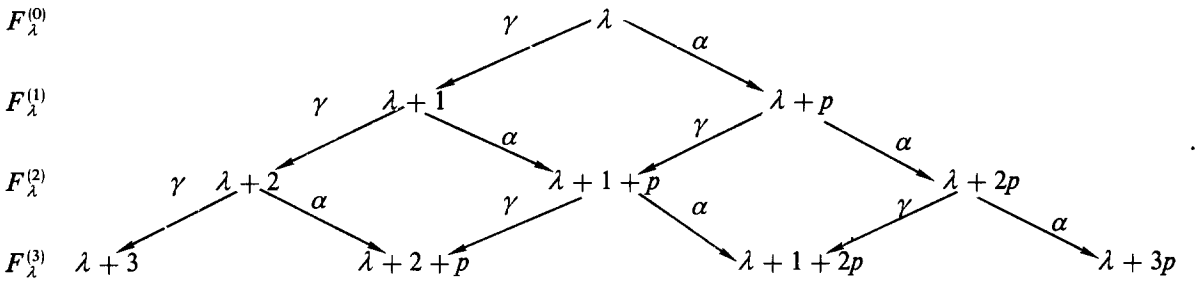
Suppose first that $\alpha = 0$ [i.e., $V(r) = \gamma r^{-1}$]. Then the mth iteration of Eq. (3.4) yields

$$F_{\lambda,\gamma}^{(m)}(r) = \left[m! \prod_{j=0}^{m-1} (2\lambda + j)\right]^{-1} \gamma^m r^{\lambda+m}, \quad m = 0,1,\ldots. \tag{3.35}$$

Similarly, if $\gamma = 0$ [i.e., $V(r) = \alpha r^{-a}$] the mth iteration of Eq. (3.4) yields

$$F_{\lambda,\alpha}^{(m)}(r) = \left[m!\, p^m \prod_{j=1}^{m} (2\lambda - 1 + jp)\right]^{-1} \alpha^m r^{\lambda+mp}, \quad p = 2-a, \quad m = 0,1,\ldots. \tag{3.36}$$

For the total potential $V(r) = \gamma r^{-1} + \alpha r^{-a}$ we get of course additional mixed exponents which are exhibited in the following diagram:



(3.37)

After inserting the above asymptotic expansion of $F_\lambda(r)$ as $r \to 0_+$ into Eq. (3.16) a closer look at condition (3.29) and (3.16) then shows that in order to obtain $G_\lambda^B(r)$ we need an asymptotic expansion for $F_\lambda(r)$ up to exponents of the type r^s, $s \leqslant 3\lambda - 1$. We denote the corresponding asymptotic expansion of $F_\lambda(r)$ by $F_\lambda^B(r)$, i.e.,

$$F_\lambda(r) - F_\lambda^B(r) \underset{r\to 0_+}{=} o(r^s), \quad s \leqslant 3\lambda - 1. \tag{3.38}$$

Using Eq. (3.16) this implies that $G_\lambda^B(r)$ represents the asymptotic expansion of $G_\lambda(r)$ as $r \to 0_+$ up to the order r^t, $t \leqslant 2\lambda - 1$, i.e.,

$$G_\lambda(r) - G_\lambda^B(r) \underset{r\to 0_+}{=} o(r^t), \quad t \leqslant 2\lambda - 1. \tag{3.39}$$

The above-described mechanism works for all $\frac{1}{2} \leqslant \lambda < \frac{3}{2}$, $\gamma \in \mathbb{R}$ and $0 < a < 2$. But clearly the number of terms in $F_\lambda^B(r)$ and $G_\lambda^B(r)$ drastically increases as $a \to 2_-$ as long as λ runs through the whole interval $-\frac{1}{2} \leqslant \lambda < \frac{3}{2}$. In order to keep the treatment reasonably short we give a complete discussion in the case $p \geqslant \frac{1}{2}$ (i.e., $0 < a \leqslant \frac{3}{2}$). From diagram (3.37) and (3.38) we infer that $F_\lambda^B(r)$ must consist of the terms r^λ, $r^{\lambda+p}$, $r^{\lambda+1}$, $r^{\lambda+2p}$, $r^{\lambda+1+p}$, and $r^{\lambda+3p}$. More precisely

$$F_\lambda^B(r) = r^\lambda [1 + A_1 r^p + (\gamma/2\lambda) r + A_2 r^{2p} + A_3 r^{3p} + A_4 r^{1+p}], \tag{3.40}$$

$$\begin{aligned} A_1 &= [p(2\lambda + p - 1)]^{-1}\alpha, \\ A_2 &= [2p^2(2\lambda + p - 1)(2\lambda + 2p - 1)]^{-1}\alpha^2, \\ A_3 &= [6p^3(2\lambda + p - 1)(2\lambda + 2p - 1)(2\lambda + 3p - 1)]^{-1}\alpha^3, \\ A_4 &= (p+1)^{-1}(2\lambda + p)^{-1}\{[p(2\lambda + p - 1)]^{-1} \\ &\quad - (2\lambda)^{-1}\}\gamma\alpha. \end{aligned} \tag{3.41}$$

Next we compute $G_\lambda^B(r)$. By Eq. (3.39) we have to take into account terms up to order r^{3p}. According to Eq. (3.16) we expand

$$F_\lambda^B(r) \int_r^{r_0} dr'\, r'^{-2\lambda} [1 + f^B(r')]^{-1}, \tag{3.42}$$

where

$$\begin{aligned} [F_\lambda^B(r)]^2 &= r^{2\lambda}[1 + f^B(r)], \\ f^B(r) &= 2A_1 r^p + (\gamma/\lambda) r + (2A_2 + A_1^2) r^{2p} \\ &\quad + [2A_4 + (A_1\gamma/\lambda)] r^{1+p} + (2A_3 + 2A_1A_2) r^{3p}. \end{aligned} \tag{3.43}$$

Taking r_0 small enough ($0 < r \leqslant r_0$) we get $|f^B(r)| < 1$ for $r \in [0, r_0]$ and hence

$$\begin{aligned} [1 + f^B(r)]^{-1} \underset{r\to 0_+}{=}\ & 1 - 2A_1 r^p + (3A_1^2 - 2A_2) r^{2p} \\ & - (\gamma/\lambda) r + (6A_1A_2 - 4A_1^3 - 2A_3) r^{3p} \\ & + [(3A_1\gamma/\lambda) - 2A_4] r^{1+p} + O(r^{1+2p}). \end{aligned} \tag{3.44}$$

A formal integration then yields

$$\begin{aligned} &\int_r^{r_0} dr'\, r'^{-2\lambda} [1 + f^B(r')]^{-1} \\ &= r^{1-2\lambda}\left\{\frac{1}{2\lambda - 1} + 2A_1 \frac{r^p}{p + 1 - 2\lambda} + (2A_2 - 3A_1^2) \right. \\ &\quad \times \frac{r^{2p}}{2p + 1 - 2\lambda} + \frac{\gamma}{\lambda} \frac{r}{2(1-\lambda)} \\ &\quad \left. + (4A_1^3 + 2A_4 - 6A_1A_2) \right. \end{aligned}$$

$$\times\frac{r^{3p}}{3p+1-2\lambda}+[2A_3-(3A_1\gamma/\lambda)]\frac{r^{1+p}}{2(1-\lambda)+p}$$
$$+C(r_0)+O(r^{1+2p})\Big\}. \tag{3.45}$$

But Eq. (3.45) has to be supplemented by the following exceptions:

If $\lambda=\frac{1}{2}$,

$\frac{r^{1-2\lambda}}{2\lambda-1}$ should be replaced by $(-\ln r)$;

if $\lambda=1$,

$\frac{r^{2(1-\lambda)}}{2(1-\lambda)}$ should be replaced by $\ln r$;

if $p=2\lambda-1$,

$\frac{r^{p+1-2\lambda}}{p+1-2\lambda}$ should be replaced by $\ln r$; (3.46)

if $p=(2\lambda-1)/2$,

$\frac{r^{2p+1-2\lambda}}{2p+1-2\lambda}$ should be replaced by $\ln r$;

if $p=(2\lambda-1)/3$,

$\frac{r^{3p+1-2\lambda}}{3p+1-2\lambda}$ should be replaced by $\ln r$

if $p=2(\lambda-1)$,

$\frac{r^{2(1-\lambda)+p}}{2(1-\lambda)+p}$ should be replaced by $\ln r$.

Without loss of generality we take $C(r_0)=0$ and obtain from Eqs. (3.16) and (3.45)

$$G_\lambda^B(r)=G_\lambda^{(0)}(r)+r^{1-\lambda}\{B_1r^p+B_2r^{2p}$$
$$-\gamma[2(1-2\lambda)(1-\lambda)]^{-1}r+B_3r^{3p}+B_4r^{1+p}\}, \tag{3.47}$$

$$B_1=[2(p+1-2\lambda)^{-1}+(2\lambda-1)^{-1}]A_1,$$
$$B_2=(p+1-2\lambda)^{-1}2A_1^2+(2p+1-2\lambda)^{-1}$$
$$\times[2A_2-3A_1^2]+(2\lambda-1)^{-1}A_2, \tag{3.48}$$
$$B_3=(3p+1-2\lambda)^{-1}(4A_1^3+2A_3-6A_1A_2)$$
$$+(2p+1-2\lambda)^{-1}(2A_1A_2-3A_1^3)$$
$$+(p+1-2\lambda)^{-1}2A_1A_2+(2\lambda-1)^{-1}A_3,$$
$$B_4=\{[\lambda(p+1-2\lambda)]^{-1}+[2\lambda(1-\lambda)]^{-1}\}\gamma A_1$$
$$+[2(1-\lambda)+p]^{-1}[2A_4-(3\gamma A_1/\lambda)]+(2\lambda-1)^{-1}A_4.$$

Of course Eqs. (3.47) and (3.48) do not hold for the exceptional values of λ and p [listed in (3.46)]. If λ takes on some of the values described in (3.46) then $G_\lambda^B(r)$ results after inserting the corresponding logarithmic term in Eq. (3.45) and multiplying with $F_\lambda^B(r)$. These logarithmic cases are familiar from the theory of Fuchsian differential equations.

Finally, we note two special cases.

(A) $\lambda=\frac{1}{2}$ (the s-wave Schrödinger operator in two dimensions): In this case the construction of $F_{1/2}^B(r)$ and $G_{1/2}^B(r)$ is particularly simple since Eqs. (3.38) and (3.39) imply

$$F_{1/2}^B(r)=F_{1/2}^{(0)}(r)=r^{1/2},$$
$$G_{1/2}^B(r)=G_{1/2}^{(0)}(r)=-r^{1/2}\ln r, \tag{3.49}$$

and thus

$$g_{0,1/2}=\lim_{r\to0_+}[-(r^{1/2}\ln r)^{-1}g(r)], \tag{3.50}$$
$$g_{1,1/2}=\lim_{r\to0_+}r^{-1/2}[g(r)+g_{0,1/2}r^{1/2}\ln r],\quad g\in\mathscr{D}(h^*).$$

(B) $\lambda=1$ (the s-wave Schrödinger operator in three dimensions): For $p>\frac{1}{4}$ (i.e., $0<a<\frac{7}{4}$) one obtains

$$G_1^B(r)=1+B_1'r^p+B_2'r^{2p}+\gamma r\ln r$$
$$+(\gamma/2)r+B_3'r^{3p}, \tag{3.51}$$
$$B_1'=[1+(p-1)^{-1}2]A_1',$$
$$B_2'=A_2'+(p-1)^{-1}2A_1'^2+(2p-1)^{-1}(2A_2'-3A_1'^2), \tag{3.52}$$
$$B_3'=A_3'+(3p+1-2)^{-1}(4A_1'^3+2A_3'-6A_1'A_2')$$
$$+(p-1)^{-1}2A_1'A_2'$$
$$+(2p-1)^{-1}(2A_1'A_2'-3A_1'^2),$$

and

$$A_1'=[p(p+1)]^{-1}\alpha,$$
$$A_2'=[2p^2(p+1)(2p+1)]^{-1}\alpha^2, \tag{3.53}$$
$$A_3'=[6p^3(p+1)(2p+1)(3p+1)]^{-1}\alpha^3.$$

For $p>\frac{1}{3}$ one can delete the r^{3p} term and for $p>\frac{1}{2}$ one can in addition delete the r^{2p} term in Eq. (3.51).

Summarizing the whole section, we have proved the following.

Theorem 3.1: Assume the conditions in (3.1). Then all self-adjoint extensions h_ν of h can be characterized by

$$h_\nu=-\frac{d^2}{dr^2}+\lambda(\lambda-1)r^{-2}+\gamma r^{-1}+\alpha r^{-a}+W, \tag{3.54}$$
$$\mathscr{D}(h_\nu)=\{g\in L^2((0,\infty))\,|\,g,g'\in AC_{\rm loc}((0,\infty));$$
$$\nu g_{0,\lambda}=g_{1,\lambda};$$
$$-g''+\lambda(\lambda-1)r^{-2}g+\gamma r^{-1}g+\alpha r^{-a}g$$
$$\in L^2((0,\infty))\},$$
$$-\infty<\nu\leqslant\infty,\quad \tfrac{1}{2}\leqslant\lambda<\tfrac{3}{2},\quad \alpha,\gamma\in\mathbb{R},\quad 0<a<2.$$

Here the boundary values $g_{0,\lambda}$ and $g_{1,\lambda}$ are defined as

$$g_{0,\lambda}=\lim_{r\to0_+}g(r)/G_\lambda^{(0)}(r), \tag{3.55}$$
$$g_{1,\lambda}=\lim_{r\to0_+}[g(r)-g_{0,\lambda}G_\lambda^B(r)]/F_\lambda^{(0)}(r),\quad g\in\mathscr{D}(h^*),$$

where $F_\lambda^{(0)}(r)$ and $G_\lambda^{(0)}(r)$ are given by Eq. (3.6) and $G_\lambda^B(r)$ denotes the asymptotic expansion of $G_\lambda(r)$ as $r\to0_+$ up to order r^t, $t\leqslant2\lambda-1$. The boundary condition $g_{0,\lambda}=0$ (i.e., $\nu=\infty$) represents the Friedrichs extension of h.

In the trivial case $\lambda=1$, $\alpha=\gamma=0$, the boundary values take on the familiar form

$$g_{0,1} = g(0_+), \quad g_{1,1} = g'(0_+). \tag{3.56}$$

Remark 3.2: In the special case where $V(r)$ has a Laurent expansion of the type $V(r) = \Sigma_{m=-2}^{\infty} a_m r^m$ near the origin, the above result has been derived by Rellich.[13] His proof relies entirely on the meromorphic structure of V whereas ours seems to be more direct and covers the general case $V(r) = \alpha r^{-a}, 0 < a < 2$. It is obvious from the arguments presented above that our method extends to potentials of the type

$$V(r) = \sum_{j=1}^{N} \alpha_j r^{-a_j} + W,$$
$$W \in L^{\infty}((0,\infty)) \text{ real valued}, \quad \alpha_j \in \mathbb{R},$$
$$0 < a_j < 2, \quad N \in \mathbb{N}. \tag{3.57}$$

In addition, our analysis extends in a straightforward manner to $\lambda(\lambda - 1) \in \mathbb{R}$ since semiboundedness of h (i.e., $\lambda \geqslant \frac{1}{2}$) turns out to be inessential.

ACKNOWLEDGMENTS

We would like to thank Professor S. Albeverio, Professor H. Behncke, Professor R. Høegh-Krohn, Professor H. Holden, Professor L. Streit, and Professor B. Thaller for numerous illuminating discussions and helpful suggestions.

We are indebted to Professor L. Streit for his kind invitation to participate in Project Nr. 2: Mathematics + Physics at the ZiF, Universität Bielefeld, Federal Republic of Germany. One of us (F.G.) would like to thank Professor M. Sirugue-Collin and Professor M. Sirugue for their kind hospitality extended to him at the C. P. T., Centre National de lá Recherche Scientifique, de Luminy, Marseille and Université de Provence, U. E. R. de Physique, where this work has been completed.

Financial support by the above-mentioned institutions is gratefully acknowledged.

[1]S. Albeverio and R. Høegh-Krohn, J. Operator Th. **6**, 313 (1981).
[2]S. Albeverio, F. Gesztesy, and R. Høegh-Krohn, Ann. Inst. H. Poincaré A **37**, 1 (1982).
[3]S. Albeverio, F. Gesztesy, R. Høegh-Krohn, and L. Streit, Ann. Inst. H. Poincaré A **38**, 263 (1983).
[4]S. Alberverio, F. Gesztesy, R. Høegh-Krohn, and W. Kirsch, J. Operator Th. **12**, 101 (1984).
[5]S. Albeverio, F. Gesztesy, R. Høegh-Krohn, and H. Holden, "Some exactly solvable models in quantum mechanics and the low energy expansion," in *Proceedings of the Second International Conference on Operator Algebras, Ideals, and Their Applications in Theoretical Physics, Leipzig, 1983*, edited by H. Baumgärtel, G. Lassner, A. Pietsch, and A. Uhlmann (Teubner, Leipzig, 1984), p. 12.
[6]F. Gesztesy and W. Kirsch, "One-dimensional Schrödinger operators with interactions singular on a discrete set," preprint, ZiF, Universität Bielefeld, Federal Republic of Germany 1984; F. Gesztesy, "Perturbation theory for resonances in terms of Fredholm determinants," in *Resonances–Models and Phenomena, Lecture Notes in Physics*, Vol. 211, edited by S. Albeverio, L. S. Ferreira, and L. Streit (Springer, Berlin, 1984), p. 78.
[7]L. Dabrowski and H. Grosse, "On nonlocal point interactions in one, two and three dimension," preprint, University of Vienna, UWThPh-1983-15.
[8](a) F. Gesztesy, G. Karner, and L. Streit, "Charged particles with a short-range force: Perturbation theory with respect to the range and to additional effects," preprint, ZiF, Universität Bielefeld, Federal Republic of Germany, 1984; (b) J. P. Antoine, F. Gesztesy, and J. Shabani (in preparation).
[9]S. Albeverio, F. Gesztesy, R. Høegh-Krohn, and H. Holden *Exactly Solvable Models in Quantum Mechanics* (book in preparation).
[10]A. Grossman and T. T. Wu, J. Math. Phys. **25**, 1742 (1984).
[11]H. Behnke, "Spectral properties of the Dirac equation with anomalous magnetic moment," preprint, Universität Osnabrück, Federal Republic of Germany 1983.
[12]B. Karnarski, "Generalized Dirac operators with several singularities," J. Operator Th. (in print).
[13]F. Rellich, Math. Z. **49**, 702 (1943/44).
[14]L. Hörmander, *The Analysis of Linear Partial Differential Operators I* (Springer, Berlin, 1983).
[15]H. Kalf, Ark. Mat. **18**, 19 (1980).
[16]G. Nenciu, Helv. Phys. Acta **50**, 1 (1977); M. Klaus, Helv. Phys. Acta **53**, 463 (1980).
[17]H. Kalf, J. Funct. Anal. **10**, 230 (1972).
[18]D. B. Pearson, Helv. Phys. Acta. **48**, 639 (1975).
[19]M. Combescure and J. Ginibre, J. Funct. Anal. **29**, 54 (1978).
[20]W. O. Amrein, *Non-Relativistic Quantum Dynamics* (Reidel, Dordrecht, 1981), Chap. 2.5.
[21]C. G. Simader, Math. Z. **159**, 47 (1978).
[22]H. Brézis, J. Operator Th. **1**, 287 (1979).
[23]H. L. Cycon, Int. J. Math. Math. Sci. **5**, 545 (1982).
[24]J. D. Morgan, J. Operator Th. **1**, 109 (1979).
[25]E. C. Svendsen, J. Math. Anal. Appl. **80**, 551 (1981).
[26]K. Jörgens, Math. Scand. **15**, 5 (1964).
[27]J. Piepenbrink and P. Rejto, Duke Math. J. **41**, 593 (1974).
[28]H. Behncke and H. Focke, Proc. Soc. Edinburgh Sec. A **78**, 119 (1977); Math. Z. **158**, 87 (1978).
[29]W. Hunziker and C. Günther, Helv. Phys. Acta **53**, 201 (1980).
[30]N. Dunford and J. T. Schwartz, *Linear Operators II. Spectral Theory* (Interscience, New York, 1964).
[31]M. A. Naimark, *Linear Differential Operators II* (Ungar, New York, 1968).
[32]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. II: Fourier Analysis, Self-Adjointness* (Academic, New York, 1975).
[33]J. Zorbas, J. Math. Phys. **21**, 840 (1980).
[34]F. Gesztesy and L. Pittner, Rep. Math. Phys. **19**, 143 (1984).
[35]F. Gesztesy and B. Thaller, J. Phys. A **14**, 639 (1981).
[36]D. Bollé and F. Gesztesy, Phys. Rev. A **30**, 1279 (1984).
[37]E. L. Ince, *Ordinary Differential Equations* (Longmans, London, 1927).
[38]J. Weidmann, *Linear Operators in Hilbert Spaces* (Springer, New York, 1980).

On the semiclassical description of N-level systems interacting with radiation fields

G. A. Raggio and H. S. Zivi
Laboratorium für Physikalische Chemie, ETH Zürich, CH-8092 Zürich, Switzerland

(Received 18 February 1985; accepted for publication 7 June 1985)

It is shown that the dynamics of an N-level quantum system driven by a classical radiation field can be derived rigorously from an associated fully quantum-mechanical model with the help of an asymptotic limit. By formally expanding the full quantum dynamics around the semiclassical limit, quantum corrections to the semiclassical description are systematically constructed. The observables describing these corrections are shown to exist in any order, and explicit expressions for their time dependence are given.

I. INTRODUCTION

Most experimental investigations of quantum-mechanical systems subject these systems to controlled external disturbances and observe their response. In order for these disturbances to be controlled experimentally they must be specifiable without reference to the system under investigation. Hence they must be describable in classical terms.[1]

The quantum-mechanical formalism, at the outset, does not include classical external forces. Such forces have to be introduced in an *ad hoc* manner for each specific case, so as to describe the particular experimental situation. This is done by the use of Hamiltonians containing time-dependent parameters that characterize the external forces.

In the present study, which has been inspired by a paper of Hepp[2] on the classical limit of quantum-mechanical correlation functions, it is shown that the dynamics of a quantum-mechanical system driven by an classical external force may be derived, rigorously, from an associated fully quantum-mechanical model with the help of an asymptotic limit. By expanding the full quantum dynamics around the semiclassical limit, quantum corrections to the semiclassical dynamics can be systematically constructed.

In the simplest possible case, our result can be paraphrased as follows: Assume that the Hamiltonian for a two-level system interacting with one mode of the radiation field is given by

$$H = \omega_0\sigma_3 + \lambda\sigma_1(\rho a^* + \bar{\rho}a) + \omega a^* a,$$

where σ_1,σ_3 are the usual Pauli matrices for the two-level system, and a is the usual annihilation operator for one field mode. Suppose that the initial state is given by a product

$$D\,|z\rangle\langle z|$$

of an arbitrary density operator D for the two-level system, and the projection $|z\rangle\langle z|$ associated with the coherent-state vector $|z\rangle$, z a complex number, for the field mode. Then, in the limit

$\lambda \to 0$ (weak coupling),
$|z| \to \infty$ (high field),

$\lambda z = \text{const}$ (constant coupling energy),

the dynamics of the two-level system is given by the time-dependent Hamiltonian

$$H(z,\lambda,t) = \omega_0\sigma_3 + \lambda(\rho e^{i\omega t}\bar{z} + \bar{\rho}e^{-i\omega t}z)\sigma_1.$$

Our paper is organized as follows: In Sec. II, we motivate the Hamiltonian which will specify our quantum-mechanical model. This model, and the results concerning its semiclassical limit (i.e., weak coupling, high field, but constant coupling energy) and the existence of fluctuation observables in any order, are given in Sec. III. The proofs are provided in Sec. IV. Section V contains our conclusions.

II. MOTIVATING THE HAMILTONIAN

Let us consider a system of K particles with masses m_j, charges z_j, and charge distributions $z_j\,\rho(\mathbf{x}_j)$, $j \in \{1,...,K\}$, coupled to an electromagnetic field. Here, ρ is a real-valued spherically symmetric generalized function satisfying

$$\int_{\mathbf{R}^3} d^3x\,\rho(\mathbf{x}) = 1, \tag{2.1}$$

i.e., $\rho(\mathbf{x}) = \delta(\mathbf{x})$ for point particles. The Hamiltonian, in the Coulomb gauge, is given by[3]

$$H = \sum_{j=1}^{K} (2m_j)^{-1}\,[\not{\partial}_j - z_j\mathbf{A}(x_j)]^2 + V(x) + \frac{1}{2}\int_{\mathbf{R}^3} d^3x{:}\mathbf{E}_{\mathrm{tr}}(\mathbf{x})^2 + \mathbf{B}(\mathbf{x})^2{:}, \tag{2.2}$$

where

$$V(\mathbf{x}) := \sum_{\substack{j=1\\ j<m}}^{K}\sum_{m=1}^{K} z_j\,z_m \int_{\mathbf{R}^3} d^3r \int_{\mathbf{R}^3} d^3r'\,\rho(\mathbf{x}_j - \mathbf{r}) \times \rho(\mathbf{x}_m - \mathbf{r}')\cdot|\mathbf{r} - \mathbf{r}'|^{-1}, \tag{2.3}$$

$$\mathbf{A}(\mathbf{x}) := \int_{\mathbf{R}^3} d^3y\,\mathbf{A}_{\mathrm{tr}}(\mathbf{x}+\mathbf{y})\,\rho(\mathbf{y}), \tag{2.4}$$

$$\mathbf{A}_{\mathrm{tr}}(\mathbf{x}) := (2\pi)^{-3/2}\sum_{\alpha=1}^{2}\int_{\mathbf{R}^3} d^3k\,(|\mathbf{k}|)^{-1/2}\mathbf{e}_\alpha(\mathbf{k}) \times [a_\alpha(\mathbf{k})e^{i\mathbf{k}\mathbf{x}} + a_\alpha^*(\mathbf{k})\,e^{-i\mathbf{k}\mathbf{x}}]. \tag{2.5}$$

We have the usual commutation relations

$$[x_{j\mu}, \not{\partial}_{m\nu}]_- = i\delta_{jm}\delta_{\mu\nu}, \quad \mu,\nu \in \{1,2,3\},\ j,m \in \{1,...,K\}, \tag{2.6}$$

 0022-2488/85/102529-11$02.50

$$[a_\alpha(\mathbf{k}),a_\beta^*(\mathbf{k}')]_- = \delta_{\alpha\beta}\delta(\mathbf{k}-\mathbf{k}'),$$
$$\alpha,\beta\in\{1,2\},\quad \mathbf{k},\mathbf{k}'\in\mathbb{R}^3. \tag{2.7}$$

The $\mathbf{e}_\alpha(\mathbf{k})$ are the polarization vectors satisfying

$$\mathbf{k}\mathbf{e}_\alpha(\mathbf{k})=0,\quad \mathbf{e}_\alpha(\mathbf{k})\,\mathbf{e}_\beta(\mathbf{k})=\delta_{\alpha\beta},\quad \alpha,\beta\in\{1,2\}. \tag{2.8}$$

If we neglect the $\mathbf{A}(x_j)^2$ terms in (2.2),[4] we obtain

$$H = H_0 + \sum_{\alpha=1}^{2}\int_{\mathbb{R}^3} d^3k\,|\mathbf{k}|a_\alpha^*(\mathbf{k})a_\alpha(\mathbf{k}) + H', \tag{2.9}$$

where

$$H_0 := \sum_{j=1}^{K}(2m_j)^{-1}\not{p}_j^2 + V(x), \tag{2.10}$$

$$H' := \sum_{\alpha=1}^{2}\int_{\mathbb{R}^3} d^3k\,(2|\mathbf{k}|)^{-1/2}\tilde{\rho}(\mathbf{k})\,\mathbf{e}_\alpha(\mathbf{k}) \times[a_\alpha(\mathbf{k})\mathbf{G}(\not{p},x,\mathbf{k}) + a_\alpha^*(\mathbf{k})\mathbf{G}(\not{p},x,-\mathbf{k})]. \tag{2.11}$$

Here, $\tilde{\rho}$ is the Fourier transform of ρ,

$$\tilde{\rho}() = (2\pi)^{-3/2}\int_{\mathbb{R}^3} d^3x\,e^{-i\mathbf{k}\mathbf{x}}\rho(\mathbf{x}) \tag{2.12}$$

[notice that, by the assumptions on ρ, $\tilde{\rho}$ is real valued, spherically symmetric, and $\tilde{\rho}(\mathbf{O})=1$], and

$$\mathbf{G}(\not{p},x,\mathbf{k}) := -\sum_{j=1}^{K} z_j(2m_j)^{-1}(\not{p}_j\,e^{i\mathbf{k}x_j} + e^{i\mathbf{k}x_j}\not{p}_j) = G(\not{p},x,-\mathbf{k})^*. \tag{2.13}$$

Assume that $\mathscr{H}_N$ is an N-dimensional subspace of $L^2(\mathbb{R}^{3K})$, contained in the domains of definition of H_0 and $\not{p}_j$, $j\in\{1,\ldots,K\}$. Let $\mathscr{P}$ be the projection operator onto this subspace. Then the projection $\mathscr{P}H\mathscr{P}$ of H onto this subspace is given by

$$\mathscr{P}H\mathscr{P} = F + \sum_{\alpha=1}^{2}\int_{\mathbb{R}^3} d^3k\,|\mathbf{k}|a_\alpha^*(\mathbf{k})a_\alpha(\mathbf{k}) + \sum_{\alpha=1}^{2}\int_{\mathbb{R}^3} d^3k\,(2|\mathbf{k}|)^{-1/2}\mathbf{e}_\alpha(\mathbf{k})\,\tilde{\rho}(\mathbf{k}) \times[a_\alpha(\mathbf{k})\mathbf{G}(\mathbf{k}) + a_\alpha^*(\mathbf{k})\mathbf{G}(-\mathbf{k})], \tag{2.14}$$

where

$$\mathbf{G}(\mathbf{k}) := \mathscr{P}\mathbf{G}(\not{p},x,\mathbf{k})\mathscr{P} = \mathbf{G}(-\mathbf{k})^*, \tag{2.15}$$

$$F := \mathscr{P}H_0\mathscr{P}. \tag{2.16}$$

Note that $\mathbf{G}(\not{p},x,\mathbf{k})$ is an unbounded linear operator defined on $L^2(\mathbb{R}^{3K})$ satisfying $\mathbf{G}(\not{p},x,\mathbf{k})^* = \mathbf{G}(\not{p},x,-\mathbf{k})$, and that $\mathbf{G}(\mathbf{k})$ is a bounded linear operator on $\mathscr{H}_N$ satisfying $\mathbf{G}(\mathbf{k})^* = \mathbf{G}(-\mathbf{k})$. In what follows, we will discuss two special cases in which

$$\mathbf{e}_\alpha(\mathbf{k})\mathbf{G}(\pm\mathbf{k}) = \tilde{g}_\alpha(\mathbf{k})V_\alpha,\quad \alpha\in\{1,2\}, \tag{2.17}$$

for self-adjoint linear operators V_α on $\mathscr{H}_N$, and real-valued functions $\tilde{g}_\alpha$. We thus have

$$\mathscr{P}H\mathscr{P} = F + \sum_{\alpha=1}^{2}\int_{\mathbb{R}^3} d^3k\,|\mathbf{k}|a_\alpha^*(\mathbf{k})a_\alpha(\mathbf{k}) + \sum_{\alpha=1}^{2}\int_{\mathbb{R}^3} d^3k \times(2|\mathbf{k}|)^{-1/2}\tilde{\rho}(\mathbf{k})\,\tilde{g}_\alpha(\mathbf{k})[a_\alpha(\mathbf{k}) + a_\alpha^*(\mathbf{k})]\,V_\alpha. \tag{2.18}$$

A. First case: Two-level approximation[5]

Let $\varphi_1,\varphi_2\in L^2(\mathbb{R}^{3K})$ be eigenfunctions of H_0 to the respective eigenvalues E_1 and E_2, $E_1\neq E_2$. Assume further that (i) φ_1 and φ_2 are real valued, and (ii) $\varphi_1(-\mathbf{x}) = \varphi_1(\mathbf{x})$, $\varphi_2(-\mathbf{x}) = -\varphi_2(\mathbf{x})$.

Let $\mathscr{P}$ be the projection operator onto the two-dimensional subspace of $L^2(\mathbb{R}^{3K})$ spanned by φ_1 and φ_2, i.e., $\mathscr{P} = |\varphi_1)(\varphi_1| + |\varphi_2)(\varphi_2|$. We have, for $\gamma\in\{1,2\}$

$$\begin{aligned}(\varphi_\gamma,\mathbf{G}(\not{p},x,\mathbf{k})\varphi_\gamma) &= (\mathbf{G}(\not{p},x,\mathbf{k})^*\varphi_\gamma,\varphi_\gamma)\\ &= (\mathbf{G}(\not{p},x,-\mathbf{k})\,\varphi_\gamma,\varphi_\gamma)\\ &= \overline{(\varphi_\gamma,\mathbf{G}(\not{p},x,-\mathbf{k})\,\varphi_\gamma)}.\end{aligned} \tag{2.19}$$

Using (2.13), (2.6), (i), and (ii) above, we find

$$(\varphi_\gamma,\mathbf{G}(\not{p},x,\mathbf{k})\varphi_\gamma) = \mathbf{O}, \tag{2.20}$$

and hence

$$\begin{aligned}\mathbf{G}(\mathbf{k}) = &-\sum_{j=1}^{K} z_j(m_j)^{-1}(\varphi_1,\cos(\mathbf{k}x_j)\not{p}_j\varphi_2)\\ &\times[|\varphi_1)(\varphi_2| - |\varphi_2)(\varphi_1|]\\ &-\mathbf{k}\sum_{j=1}^{K} z_j(m_j)^{-1}(\varphi_1,e^{i\mathbf{k}x_j}\varphi_2)|\varphi_1)(\varphi_2|\\ &+\mathbf{k}\sum_{j=1}^{K} z_j(m_j)^{-1}(\varphi_2,e^{i\mathbf{k}x_j}\varphi_1)|\varphi_2)(\varphi_1|.\end{aligned} \tag{2.21}$$

Then, using the transversality condition (2.8)

$$\begin{aligned}&\mathbf{e}_\alpha(\mathbf{k})\mathbf{G}(\pm\mathbf{k})\\ &= i\sum_{j=1}^{K} z_j(m_j)^{-1}\mathbf{e}_\alpha(\mathbf{k})(\varphi_1,\cos(\mathbf{k}x_j)\not{p}_j\varphi_2)\\ &\quad\times[i|\varphi_1)(\varphi_2| - i|\varphi_2)(\varphi_1|] = \tilde{g}_\alpha(\mathbf{k})V,\end{aligned} \tag{2.22}$$

where

$$\tilde{g}_\alpha(\mathbf{k}) := i\sum_{j=1}^{K} z_j(m_j)^{-1}\mathbf{e}_\alpha(\mathbf{k})(\varphi_1,\cos(\mathbf{k}x_j)\not{p}_j\varphi_2), \tag{2.23}$$

$$V := i|\varphi_1)(\varphi_2| - i|\varphi_2)(\varphi_1|. \tag{2.24}$$

B. Second case: N-level approximation in the dipole approximation for linearly polarized fields

In the dipole approximation, $\mathbf{A}(x_j)$ in (2.2) is replaced by $A(\mathbf{O})$. Then $\mathbf{G}(\not{p},x,\mathbf{k})$ is replaced by $\mathbf{G}(\not{p})$, i.e., by

$$\mathbf{G}(\not{p}) := -\sum_{j=1}^{K} z_j(m_j)^{-1}\not{p}_j, \tag{2.25}$$

and $\mathbf{G}(\mathbf{k})$ by $\mathbf{G}$, where $\mathbf{G}$ is given by

$$\mathbf{G} := \mathscr{P}\mathbf{G}(\not{p})\,\mathscr{P}. \tag{2.26}$$

If the field is linearly polarized, say

$$\mathbf{k} = (0,0,|\mathbf{k}|),\quad \mathbf{e}_1(\mathbf{k}) = (1,0,0),\quad \mathbf{e}_2(\mathbf{k}) = (0,1,0), \tag{2.27}$$

then

$$\mathbf{e}_\alpha(\mathbf{k})\mathbf{G} = :G_\alpha = -\sum_{j=1}^{K} z_j(m_j)^{-1}\,\mathscr{P}\not{p}_{j\alpha}\,\mathscr{P} = \tilde{g}_\alpha(\mathbf{k})V_\alpha, \tag{2.28}$$

$$\tilde{g}_\alpha(\mathbf{k}) = 1,\quad V_\alpha = G_\alpha. \tag{2.29}$$

Let us now return to the Hamiltonian (2.18), writing only one polarization component, and introducing a coupling constant $\lambda>0$

$$\mathcal{P}H\mathcal{P} = F + \int_{\mathbb{R}^3} d^3k\, |\mathbf{k}| a^*(\mathbf{k})a(\mathbf{k}) + \lambda \int_{\mathbb{R}^3} d^3k\, (2|\mathbf{k}|)^{-1/2} \times \tilde{\rho}(\mathbf{k})\, \tilde{g}(\mathbf{k})[a(\mathbf{k}) + a^*(\mathbf{k})]\, V. \tag{2.30}$$

In Sec. III, this Hamiltonian will be given a precise mathematical meaning as a self-adjoint operator on $\mathfrak{F} \otimes \mathbb{C}^N$, where $\mathfrak{F}$ is the symmetric Fock space over the one-particle Hilbert space $\mathfrak{H}$

$$\mathfrak{H} := L^2(\mathbb{R}^3, |\mathbf{k}|^{-1} d^3k), \tag{2.31}$$

and $\mathbb{C}^N$ is an N-dimensional complex Hilbert space.

Semiclassical radiation theory is considered a suitable approximation to quantum electrodynamics, if the field is in a highly excited coherent state $\Omega(f)$,[6]

$$\Omega(f) := W(f)\varphi,$$

$$[\Omega(f)]_n = \exp\{-\tfrac{1}{4}\|f\|^2\}(n!)^{-1/2} 2^{1/2n} f \otimes \underset{n \text{ times}}{f \otimes \cdots \otimes f}, \quad f \in \mathfrak{H}. \tag{2.32}$$

Here, $[\Omega(f)]_n$ denotes the component of $\Omega(f)$ in the n-particle subspace of $\mathfrak{F}$, φ is the Fock vacuum,

$$\varphi = 1 \oplus 0 \oplus 0 \oplus \cdots, \tag{2.33}$$

and the Weyl operator $W(f)$ is defined by

$$W(f) := \exp\{2^{-1/2}[a^*(f) - a(f)]\}. \tag{2.34}$$

Here, $a(f)$ and $a^*(f)$ are the tested field operators satisfying the commutation relation

$$[a(f), a^*(g)]_- = (f,g)\mathbb{1}, \tag{2.35}$$

where $(\cdot,\cdot)$ denotes the scalar product in $\mathfrak{H}$.

We are thus led to consider the time-evolved observables of the N-level system, reduced with the state $\Omega(f)$, and the expectation values of the time-evolved field observables with respect to the state $|\Omega(f)\rangle\langle\Omega(f)| \otimes D$, D being the initial state of the N-level system, in the limit $\|f\| \to \infty$. It is convenient to replace $\|f\|$ by $\epsilon^{-1/2}\|f\|$, where $\epsilon := \|f\|^{-2}$, and to take the limit $\epsilon \to 0$. In order that the expectation values of the observables of our model system remain finite in the limit $\epsilon \to 0$, we have to introduce scaled field operators $a^\epsilon(f)$ and $a^{\epsilon*}(f)$,

$$a^\epsilon(f) := \epsilon^{1/2} a(f) = a(\epsilon^{1/2} f),$$
$$a^{\epsilon*}(f) := \epsilon^{1/2} a^*(f) = a^*(\epsilon^{1/2} f). \tag{2.36}$$

Since $a(f)$ and $a^*(f)$ satisfy the commutation relation (2.35), it follows that the commutation relation for the scaled field operators is of order ϵ, so that, heuristically, $a^\epsilon(f)$ and $a^{\epsilon*}(f)$ are expected to become classical, i.e., commutative variables, in the limit $\epsilon \to 0$. It will be seen in Sec. III that the coupling constant λ has to be rescaled as $\lambda \to \epsilon^{1/2}\lambda$ in order that the coupling energy remains at a finite prescribed value in the limit $\epsilon \to 0$.

III. THE QUANTUM-MECHANICAL MODEL, ITS SEMICLASSICAL LIMIT, AND FLUCTUATIONS

We write $\mathfrak{B}(\mathfrak{K})$ for the bounded, linear operators on a Hilbert space $\mathfrak{K}$. We specify the one-particle Hilbert space $\mathfrak{H}$ to be $L^2(\mathbb{R}^3, d^3k\, |\mathbf{k}|^{-1})$ with scalar product $(f,g) := \int_{\mathbb{R}^3} d^3k\, |\mathbf{k}|^{-1}\, \overline{f(\mathbf{k})}\, g(\mathbf{k})$. The one-particle Hamiltonian (self-adjoint and positive) on $\mathfrak{H}$ is given by

$$\{\mathcal{h} f\}(\mathbf{k}) := |\mathbf{k}| f(\mathbf{k}), \quad \text{with} \int_{\mathbb{R}^3} d^3k\, |\mathbf{k}|\, |f(\mathbf{k})|^2 < \infty. \tag{3.1}$$

For $t \in \mathbb{R}$, let

$$\omega_t := \exp\{-it\mathcal{h}\}, \quad \text{i.e.,} \quad \{\omega_t f\}(\mathbf{k}) = \exp\{-it|\mathbf{k}|\} f(\mathbf{k}), \quad f \in \mathfrak{H} \tag{3.2}$$

be the generated unitary operator in $\mathfrak{B}(\mathfrak{H})$. On the boson Fock space over $\mathfrak{H}$, denoted by $\mathfrak{F}$ (with scalar product $\langle\cdot,\cdot\rangle$, linear in the second component), the unitary operators $W(f)$, $f \in \mathfrak{H}$, defined by (2.34), satisfy the commutation relation

$$W(f)W(g) = \exp\{-i\,\mathrm{Im}(f,g)/2\}\, W(f+g) \tag{3.3}$$

and the translation relation

$$W(-f)a(g)W(f) = a(g) + (1/\sqrt{2})(g,f)\mathbb{1}. \tag{3.4}$$

Here, and in the following, $\mathbb{1}$ will denote the identity operator on the appropriate Hilbert space. We write Ω for the second-quantization map,[7] and

$$U_0(t) := \exp\{-it\Omega(\mathcal{h})\}, \quad t \in \mathbb{R}, \tag{3.5}$$

for the unitary group in $\mathfrak{B}(\mathfrak{F})$ generated by the second quantization of the one-particle Hamiltonian $\mathcal{h}$. One has

$$U_0(-t)\, a(f) U_0(t) = a(\omega_{-t} f),$$
$$U_0(-t) W(f) U_0(t) = W(\omega_{-t} f). \tag{3.6}$$

Furthermore,

$$\langle\varphi, W(f)\varphi\rangle = \exp\{-\|f\|^2/4\},$$
$$\langle\Omega(f), W(g)\Omega(f)\rangle = \exp\{i\,\mathrm{Im}(f,g) - \|g\|^2/4\}. \tag{3.7}$$

The dynamics of the composite system field/N-level system are specified on $\mathfrak{F} \otimes \mathbb{C}^N$ by the Hamiltonian

$$H^\lambda := H_0^\lambda + \mathbb{1} \otimes F, \tag{3.8}$$

where

$$H_0^\lambda := \Omega(\mathcal{h}) \otimes \mathbb{1} + (\lambda/\sqrt{2})[a(\rho) + a^*(\rho)] \otimes V.$$

Here, F and V are self-adjoint elements of $\mathfrak{B}(\mathbb{C}^N)$, and ρ satisfies

$$\rho \in \mathfrak{H}, \quad \text{i.e.,} \quad \int_{\mathbb{R}^3} d^3k\, |\mathbf{k}|^{-2} |\rho(\mathbf{k})|^2 < \infty, \tag{3.9}$$

$$\rho \in \mathrm{Dom}(\mathcal{h}^{-1/2}), \quad \text{i.e.,} \quad \int_{\mathbb{R}^3} d^3k\, |\mathbf{k}|^{-2} |\rho(\mathbf{k})|^2 < \infty.$$

Abusing notation, we have set $\rho = \tilde{\rho}\,\tilde{g}$ in (2.30). Conditions (3.9) are satisfied if ρ belongs to the Schwartz space over $\mathbb{R}^3$, and they are compatible with the interpretation of ρ as the Fourier transform of a charge distribution. With (3.9), one can show (see, e.g., Ref. 8) that (3.8) defines a self-adjoint operator on $\mathfrak{F} \otimes \mathbb{C}^N$. We write

$$U^\lambda(t) := \exp\{-it H^\lambda\}, \quad t \in \mathbb{R}, \tag{3.10}$$

for the generated unitary operator group in $\mathfrak{B}(\mathfrak{F} \otimes \mathbb{C}^N)$.

The expectation value of the field part of the interaction in (3.8), namely $(\lambda/\sqrt{2})[a(\rho) + a^*(\rho)] \otimes V$, in the coherent state $\Omega(f)$ is given [using (2.32), (3.4), and $a(g)\varphi = 0$] by

$\lambda\,\mathrm{Re}\,(f,\rho)$. (3.11)

Whence, if we scale f by $f\to\epsilon^{-1}f$, and consider the limit $\epsilon\to 0$, (3.11) remains constant only if λ is scaled as $\lambda\to\epsilon\lambda$.

If we assume that the initial state of the composite system is given by

$$|\Omega(f)\rangle\langle\Omega(f)|\otimes D, \tag{3.12}$$

where $D\in\mathfrak{B}(\mathbb{C}^N)$ is any density operator for the N-level system and $|\Omega(f)\rangle\langle\Omega(f)|$ is the projection operator in $\mathfrak{B}(\mathfrak{F})$ projecting onto the ray spanned by $\Omega(f)$ in $\mathfrak{F}$, then the expectation value of the operator $W(g)\otimes A\in\mathfrak{B}(\mathfrak{F}\otimes\mathbb{C}^N)$, $g\in\mathfrak{H}$, $A\in\mathfrak{B}(\mathbb{C}^N)$, at time t is

$$\mathrm{Tr}_{\mathfrak{F}\otimes\mathbb{C}^N}\{(|\Omega(f)\rangle\langle\Omega(f)|\otimes D)\,U^\lambda(-t)\times(W(g)\otimes A)\,U^\lambda(t)\}. \tag{3.13}$$

Performing first the partial trace over $\mathfrak{F}$, we set

$$\Phi_{A,g}(f,\lambda,t):=\mathrm{Tr}_{\mathfrak{F}}\{(|\Omega(f)\rangle\langle\Omega(f)|\otimes 1)U^\lambda(-t)\times(W(g)\otimes A)\,U^\lambda(t)\}, \tag{3.14}$$

and have that (3.13) is given by

$$\mathrm{Tr}_{\mathbb{C}^N}\{D\Phi_{A,g}(f,\lambda,t)\}.$$

Here, $\Phi_{A,g}(f,\lambda,t)\in\mathfrak{B}(\mathbb{C}^N)$, and contains essentially all dynamical information for initial states (3.12) since the linear span of $\{W(g): g\in\mathfrak{H}\}$ is dense in $\mathfrak{B}(\mathfrak{F})$ in the strong operator topology.[9] Particularly,

$$A(f,\lambda,t):=\mathrm{Tr}_{\mathfrak{F}}\{(|\Omega(f)\rangle\langle\Omega(f)|\otimes 1)\,U^\lambda(-t)\times(1\otimes A)U^\lambda(t)\}=\Phi_{A,0}(f,\lambda,t) \tag{3.15}$$

is the reduced dynamics of the operator $A\in\mathfrak{B}(\mathbb{C}^N)$ for initial states (3.12), and

$$\omega(g;f,\lambda,t):=\mathrm{Tr}_{\mathfrak{F}\otimes\mathbb{C}^N}\{(|\Omega(f)\rangle\langle\Omega(f)|\otimes D)\,U^\lambda(-t)\times(W(g)\otimes 1)\,U^\lambda(t)\}=\mathrm{Tr}_{\mathbb{C}^N}\{D\Phi_{1,g}(f,\lambda,t)\} \tag{3.16}$$

is the state-generating functional of the state (3.12).[10]

We are interested in the asymptotic situation described in Sec. II, so that we study

$$\Phi_{A,\epsilon\cdot g}(\epsilon^{-1}\cdot f,\epsilon\cdot\lambda,t),$$

for $\epsilon\to 0$. Our result, formulated in terms of the directly relevant quantities, is the following.

Theorem 1: (a) For the semiclassical limit, we have the following.

(1) $A_{(0)}(f,\lambda,t):=\lim_{\epsilon\to 0}A(\epsilon^{-1}f,\epsilon\lambda,t)$ exists and is the solution of

$$i\dot{A}_{(0)}(f,\lambda,t)=[A_{(0)}(f,\lambda,t),H(f,\lambda,t)]_-,$$
$$A_{(0)}(f,\lambda,0)=A,$$

where the semiclassical Hamiltonian $H(f,\lambda,t)$ is given by

$$H(f,\lambda,t)=F+\lambda\,\mathrm{Re}(\alpha_t f,\rho)V.$$

(2) $\omega_{(0)}(g;f,\lambda,t):=\lim_{\epsilon\to 0}\omega(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t)$ exists, and

$$\omega_{(0)}(g;f,\lambda,t)=\exp\{i\,\mathrm{Im}(\alpha_t f,g)\}.$$

(b) For fluctuations, we have the following.

(1) The fluctuation of order σ, $\sigma=1,2,\ldots$, of $A(f,\lambda,t)$, defined recursively by

$$A_{(\sigma)}(f,\lambda,t):=\lim_{\epsilon\to 0}A^\epsilon_{(\sigma)}(f,\lambda,t),$$

where

$$A^\epsilon_{(\sigma)}(f,\lambda,t):=\epsilon^{-2\sigma}\Big\{A(\epsilon^{-1}f,\epsilon\lambda,t)-\sum_{\nu=0}^{\sigma-1}\epsilon^{2\nu}A_{(\nu)}(f,\lambda,t)\Big\},$$

exists as an operator in $\mathfrak{B}(\mathbb{C}^N)$.

(2) The fluctuation of order σ, $\sigma=1,2,\ldots$, of $\omega(g;f,\lambda,t)$, defined recursively by

$$\omega_{(\sigma)}(g;f,\lambda,t):=\lim_{\epsilon\to 0}\omega^\epsilon_{(\sigma)}(g;f,\lambda,t),$$

where

$$\omega^\epsilon_{(\sigma)}(g;f,\lambda,t)=\epsilon^{-2\sigma}\Big\{\omega(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t)-\sum_{\nu=0}^{\sigma-1}\epsilon^{2\nu}\omega_{(\nu)}(g;f,\lambda,t)\Big\},$$

exists.

In the proof, to be given in the next section, we shall obtain explicit infinite series expressions for all the quantities involved in Theorem 1. Let us now make a number of remarks and comments pertaining to this result: Firstly, in terms of the Schrödinger picture, we would have for a density operator $D\in\mathfrak{B}(\mathbb{C}^N)$

$$D(f,\lambda,t):=\mathrm{Tr}_{\mathfrak{F}}\{U^\lambda(t)(|\Omega(f)\rangle\langle\Omega(f)|\otimes D)\,U^\lambda(-t)\},$$

and then

$$D_{(0)}(f,\lambda,t):=\lim_{\epsilon\to 0}D(\epsilon^{-1}f,\epsilon,\lambda,t)$$

exists and satisfies

$$i\dot{D}_{(1)}(f,\lambda,t)=[H(f,\lambda,t),D_{(0)}(f,\lambda,t)]_-,$$
$$D_{(0)}(f,\lambda,0)=D.$$

Secondly, in the semiclassical limit the dynamics of the N-level system is governed by the time-dependent Hamiltonian $H(f,\lambda,t)$. The time dependence is given by

$$\lambda\,\mathrm{Re}(\alpha_t f,\rho),$$

which equals

$$(\lambda/\sqrt{2})\langle U_0(t)\Omega(f),[a(\rho)+a^*(\rho)]\,U_0(t)\Omega(f)\rangle,$$

the expectation value of the field part of the interaction in the coherent state $\Omega(f)$ evolved according to the free-field dynamics $U_0(t)$. Thirdly, as is to be expected, in the semiclassical limit the dynamics of the field is independent of the coupling. Thus, the N-level system does not influence the field in this limit. Fourthly, as was suggested in the Introduction the semiclassical limit of the field is in fact a classical limit; one can construct a classical dynamical system with the very same dynamical behavior. The phase space is given by the Hilbert space $\mathfrak{H}$ and the state at time t by $f_t:=\alpha_t f\in\mathfrak{H}$. If we write $R(f)$ [resp. $I(f)$] for the real-valued (resp. imaginary-valued) part of $f\in\mathfrak{H}$, we have $f_t=R(f_t)+iI(f_t)$, and

$$\dot{R}(f_t)=\mathcal{H}I(f_t),\quad \dot{I}(f_t)=-\mathcal{H}\mathrm{R}(f_t).$$

If X_g, $g\in\mathfrak{H}$, is the "characteristic function" given by

$$X_g(f):=\exp\{i\,\mathrm{Im}(f,g)\},\quad f\in\mathfrak{H},$$

then

$$\omega_{(0)}(g;f,\lambda,t)=X_g(f_t),$$

the (dispersion-free) value of X_g at time t. In this limit, the field is equivalent to a system of uncoupled, classical harmonic oscillators. Finally, the fluctuations (around the semiclassical solution) appear as the terms of asymptotic series in ϵ (only even powers of ϵ arise) in the sense of Poincaré

$$A(\epsilon^{-1}f,\epsilon\lambda,t)\sim\sum_{\sigma=0}^{\infty}\epsilon^{2\sigma}A_{(\sigma)}(f,\lambda,t),$$

$$\omega(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t)\sim\sum_{\sigma=0}^{\infty}\epsilon^{2\sigma}\omega_{(\sigma)}(g;f,\lambda,t).$$

IV. PROOFS

We chose once and for all an orthonormal basis $\{e_n:n=1,2,...,N\}$ of $\mathbb{C}^N$ such that V (see 3.8) is diagonal, i.e., $Ve_n=v_ne_n$, where $\{v_n:n=1,2,...,N\}$ is the set of eigenvalues of V numbered according to their multiplicities. It will often be convenient to identify $\mathfrak{F}\otimes\mathbb{C}^N$ with $\overset{N}{\underset{n=1}{\oplus}}\mathfrak{F}$; under this identification, every linear operator C acting on $\mathfrak{F}\otimes\mathbb{C}^N$ corresponds to an $N\times N$ matrix $[C]$ whose matrix elements $[C]_{jk}$ are linear operators acting on $\mathfrak{F}$. In particular,

$$[B\otimes A]_{jk}=A_{jk}B,\quad\text{where } A_{jk}=(e_j,Ae_k)_{\mathbb{C}^N}. \tag{4.1}$$

We introduce some cumbersome but unavoidable notation.

(a) $\lambda_n:=\lambda v_n\in\mathbb{R},\quad n=1,2,...,N.$

(b) $\gamma(f,t):=\int_0^t ds\int_0^s dr\,\mathrm{Im}(\omega_r f,f)\in\mathbb{R},\quad t\in\mathbb{R},\quad f\in\mathfrak{H}.$

(c) $\zeta(f,t):=-i\int_0^t ds\,\omega_s f\in\mathfrak{H},\quad t\in\mathbb{R},\quad f\in\mathfrak{H}.$

(d) If $n=0,1,2,...$, $\{m_{-1},m_0,m,m_2,...,m_n\}$ is a set of $(n+2)$ indices with values in $\{1,2,...,N\}$, and $\{t_0,t_1,t_2,...,t_n\}$ are $(n+1)$ reals, and if $\kappa=0$ or $\kappa=-1$, then let

$$\begin{aligned}\alpha(m_\kappa,m_{\kappa+1},...,m_n;t_{\kappa+1},t_{\kappa+2},...,t_n;\lambda;g)&:=\frac{1}{2}\sum_{\nu=\kappa+1}^{n}\{(\lambda^2_{m_\nu}-\lambda^2_{m_{\nu-1}})\gamma(g,t_\nu)\}\\&+\frac{1}{2}\sum_{\nu=\kappa+1}^{n}\sum_{r=0}^{n-1-\nu}\{(\lambda_{m_\nu}-\lambda_{m_{\nu-1}})\\&\times(\lambda_{m_{\nu+r+1}}-\lambda_{m_{\nu+r}})(\gamma(g,t_{\nu+r+1}-t_\nu)\\&+\gamma(g,t_\nu)-\gamma(g,t_{\nu+r+1}))\}\in\mathbb{R},\\&\lambda\in\mathbb{R},\quad g\in\mathfrak{H},\end{aligned}$$

$$\begin{aligned}\beta(m_\kappa,m_{\kappa+1},...,m_n;t_{\kappa+1},t_{\kappa+2},...,t_n;f,\lambda;g)&:=\sum_{\nu=\kappa+1}^{n}\{(\lambda_{m_{\nu-1}}-\lambda_{m_\nu})\,\mathrm{Im}(f,\zeta(g,-t_\nu))\}\in\mathbb{R},\\&\lambda\in\mathbb{R},\quad f,g\in\mathfrak{H},\end{aligned}$$

$$\begin{aligned}\xi(m_\kappa,m_{\kappa+1},...,m_n;t_{\kappa+1},t_{\kappa+2},...,t_n;\lambda;g)&:=\sum_{\nu=\kappa+1}^{n}\{(\lambda_{m_{\nu-1}}-\lambda_{m_\nu})\zeta(g,-t_\nu)\}\in\mathfrak{H},\\&\lambda\in\mathbb{R},\quad g\in\mathfrak{H}.\end{aligned}$$

(e) *If* $n=1,2,...$, $\{l_1,l_2,...,l_{n-1}\}$ are $(n-1)$ indices with values in $\{1,2,...,N\}$, and $j,m,k\in\{1,2,...,N\}$, and $\{t,t_1,t_2,...,t_n\}$ are $(n+1)$ reals, then we use the following shorthand notation for the quantities introduced in (d):

$$\begin{aligned}&\alpha(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;g)\\&\quad\equiv\alpha(j,m,l_1,l_2,...,l_{n-1},k;t,t_1,...,t_n;\lambda;g),\end{aligned}$$

and analogously for β and ξ.

Furthermore,

$$\sum_{\vec{l}_{n-1}=1}^{N}\equiv\sum_{l_1=1}^{N}\sum_{l_2=1}^{N}\cdots\sum_{l_{n-1}=1}^{N},$$

and

$$F(j,\vec{l}_{n-1},k)\equiv F_{jl_1}F_{l_1l_2}\cdots F_{l_{n-1}}k,$$

where $F\in\mathfrak{B}(\mathbb{C}^N)$, and $F_{jk}=(e_j,Fe_k)_{\mathbb{C}^N}$. Notice that

$$\gamma(f,-t)=-\gamma(f,t),$$

$$\begin{aligned}&\beta(m_\kappa,m_{\kappa+1},...,m_n;t_{\kappa+1},t_{\kappa+2},...,t_n;f,\lambda;g)\\&\quad=\mathrm{Im}(f,\xi(m_\kappa,m_{\kappa+1},...,m_n;\\&\qquad t_{\kappa+1},t_{\kappa+2},...,t_n;\lambda;g)),\end{aligned}$$

and that if we scale λ and f as $\lambda\to x\lambda$, $f\to x^yf$, $x,y\in\mathbb{R}$, then

$$\alpha\to x^2\alpha,\quad\beta\to x^{y+1}\beta,\quad\text{and}\quad\xi\to x\xi. \tag{4.2}$$

Let us remark for future reference that if $f\in\mathrm{Dom}(\mathcal{A}^{-1})$, then

$$\zeta(f,t)=(\omega_t-1)\mathcal{A}^{-1}f \tag{4.3}$$

holds true as is easily seen by differentiation. From (4.3) one deduces that

$$\gamma(f,t)=\mathrm{Im}(\mathcal{A}^{-1}f,\omega_t\mathcal{A}^{-1}f)+(f,\mathcal{A}^{-1}f)\,t, \tag{4.4}$$

for $f\in\mathrm{Dom}(\mathcal{A}^{-1})$.

We now state and prove the main technical result.

Proposition: Let $\rho\in\mathfrak{H}$ satisfy (3.9), and let H^λ and $\Phi_{A,g}(f,\lambda,t)$, where $f,g\in\mathfrak{H}$, $A\in\mathfrak{B}(\mathbb{C}^N)$, $t\in\mathbb{R}$, be given by (3.8) and (3.14), respectively. The jkth matrix element of $\Phi_{A,g}(f,\lambda,t)$ is given by

$$\Phi_{A,g}(f,\lambda,t)_{jk}=\sum_{l=1}^{N}\sum_{m=1}^{N}A_{lm}\,T^{jk}_{lm}(g;f,\lambda,t),$$

where

$$\begin{aligned}&T^{jk}_{lm}(g;f,\lambda,t)\\&=\exp\{i\,\mathrm{Im}(\omega_t f,g)\}\left[\exp\left\{-\frac{i(\lambda_j+\lambda_k)\,\mathrm{Im}(g,\zeta(\rho,t))}{2}+i\alpha(j,k;t;\lambda;\rho)+i\beta(j,k;t;f;\lambda;\rho)\right.\right.\\&\left.\left.-\frac{\|\xi(j,k;t;\lambda;\rho)+\omega_{-t}g\|^2}{4}\right\}\delta_{jl}\delta_{mk}+\delta_{jl}\left\{\exp\left\{-\frac{i(\lambda_j+\lambda_m)\,\mathrm{Im}(g,\zeta(\rho,t))}{2}\right\}\sum_{n=1}^{\infty}(-i)^n\sum_{\vec{l}_{n-1}=1}^{N}F(m,\vec{l}_{n-1},k)\right.\right.\end{aligned}$$

$$\times T\int_0^t dt^{(n)}\exp\left\{i\alpha(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)-\frac{i\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t}g,\lambda;\rho)}{2}+i\beta(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;f,\lambda;\rho)\right.$$
$$\left.-\frac{\|\xi(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)+\omega_{-t}g\|^2}{4}\right\}+\delta_{mk}\left\{\exp\left\{-\frac{i(\lambda_l+\lambda_k)\mathrm{Im}(g,\xi(\rho,t))}{2}\right\}\sum_{n=1}^{\infty}i^n\sum_{\vec{l}_{n-1}=1}^{N}\overline{F(l,\vec{l}_{n-1},j)}\right.$$
$$\times T\int_0^t dt^{(n)}\exp\left\{-i\alpha(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda;\rho)-\frac{i\beta(l,\vec{l}_{n-1},j;\vec{t}_n;\omega_{-t}g,\lambda;\rho)}{2}\right.$$
$$\left.\left.-i\beta(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;f,\lambda;\rho)-\frac{\|\xi(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda,\rho)-\omega_{-t}g\|^2}{4}\right\}\right\}+\exp\left\{-\frac{i(\lambda_l+\lambda_m)\mathrm{Im}(g,\xi(\rho,t))}{2}\right\}$$
$$\times\sum_{p=1}^{\infty}\sum_{n=1}^{\infty}(-i)^n i^p\sum_{\vec{m}_{p-1}=1}^{N}\sum_{\vec{l}_{n-1}=1}^{N}\overline{F(l,\vec{m}_{p-1},j)}\,F(m,\vec{l}_{n-1},k)\,T\int_0^t ds^{(p)}\,T\int_0^t dt^{(n)}$$
$$\times\exp\left\{-\frac{i\beta(l,\vec{m}_{p-1},j;\vec{s}_p;\omega_{-t}g,\lambda;\rho)}{2}-\frac{i\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t}g,\lambda;\rho)}{2}\right.$$
$$+\frac{i\,\mathrm{Im}(\xi(l,\vec{m}_{p-1},j;\vec{s}_p;\lambda;\rho),\xi(m,\vec{l}_{n-1},k;\vec{t}_n;\lambda;\rho))}{2}-i\alpha(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;\lambda;\rho)+i\alpha(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)$$
$$-i\alpha(l,m;t;\lambda;\rho)-i\beta(l,m;t;f,\lambda;\rho)-i\beta(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;f,\lambda;\rho)+i\beta(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;f,\lambda;\rho)$$
$$\left.\left.-\frac{\|\xi(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)-\xi(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;\lambda;\rho)-\xi(l,m;t;\lambda;\rho)+\omega_{-t}g\|^2}{4}\right\}\right].$$

The above series converge absolutely and uniformly in f, g, ρ, λ, and t for t in any bounded subset of $\mathbb{R}$.

Proof: We first prove the assertions concerning convergence of the series. Let $\Delta := \max\{|F_{jk}|: j,k\in\{1,2,\ldots,N\}\}$. Then, $|F(j,\vec{l}_{n-},k)|\leqslant\Delta^n$. The absolute value of both simple infinite series in the expression for T^{jk}_{lm} is bounded, for $t\geqslant 0$, by

$$\sum_{n=1}^{\infty}T\int_0^t dt^{(n)}\sum_{\vec{l}_{n-1}=1}^{N}\Delta^n=\sum_{n=1}^{\infty}\Delta^n\frac{t^n}{n!}\sum_{\vec{l}_{n-1}=1}^{N}=\sum_{n=1}^{\infty}\frac{\Delta^n N^{n-1}t^n}{n!}=N^{-1}(\exp\{N\Delta t\}-1).$$

Analogously, the double infinite series in the expression for T^{jk}_{lm} is bounded, for $t\geqslant 0$, by $N^{-2}(\exp\{N\Delta t\}-1)^2$.

Let $U_0^{\lambda}(t) := \exp\{-itH_0^{\lambda}\}$, $t\in\mathbb{R}$, where H_0^{λ} is given by (3.8). We have, with $Q(f) := (1/\sqrt{2})(a(f)+a^*(f))$, $f\in\mathfrak{H}$, that

$$H_0^{\lambda}=\bigoplus_{n=1}^{N}\{\Omega(\mathcal{h})+\lambda v_n Q(\rho)\}=\bigoplus_{n=1}^{N}\{\Omega(\mathcal{h})+\lambda_n Q(\rho)\},\tag{4.5}$$

so H_0^{λ} is diagonal, and so is $U_0^{\lambda}(t)$. Let $F(t) := U_0^{\lambda}(-t)(\mathbb{1}\otimes F)U_0^{\lambda}(t)$, $t\in\mathbb{R}$. The Dyson series for the propagator of $F(t)\in\mathfrak{B}(\mathfrak{F}\otimes\mathbb{C}^N)$ leads to the following expression for $U^{\lambda}(t)$:

$$U^{\lambda}(t)=U_0^{\lambda}(t)\left\{\mathbb{1}+\sum_{n=1}^{\infty}(-i)^n T\int_0^t dt^{(n)}F(t_1)F(t_2)\cdots F(t_n)\right\}.\tag{4.6}$$

The series is convergent in operator norm by the boundedness of $\mathbb{1}\otimes F$. from (3.14) with (4.1) we obtain

$$\Phi_{A,g}(f,\lambda,t)_{jk}=\langle\Omega(f),[U^{\lambda}(-t)(W(g)\otimes A)\,U^{\lambda}(t)]_{jk}\Omega(f)\rangle$$
$$=\sum_{l=1}^{N}\sum_{m=1}^{N}A_{lm}\langle\Omega(f),[U^{\lambda}(t)^*]_{jl}W(g)[U^{\lambda}(t)]_{mk}\Omega(f)\rangle.$$

Thus,

$$T^{jk}_{lm}(g,f\lambda,t)=\langle\Omega(f),[U^{\lambda}(t)]^*_{lj}W(g)[U^{\lambda}(t)]_{mk}\Omega(f)\rangle.\tag{4.7}$$

To compute T^{jk}_{lm} from (4.7), we proceed in two steps, (A) and (B). In (A), we assume that $\rho\in\mathrm{Dom}(\mathcal{h}^{-1})$ and compute T^{jk}_{lm}. This assumption is removed in (B). Notice that $\rho\in\mathrm{Dom}(\mathcal{h}^{-1})$ means

$$\int_{\mathbf{R}^3}d^3k\,|\mathbf{k}|^{-3}\rho(\mathbf{k})|^2<\infty,$$

which is incompatible with the interpretation of the Fourier transform of ρ as a charge distribution.

(A) We assume $\rho\in\mathrm{Dom}(\mathcal{h}^{-1})$, and from a result of Cook,[11] we obtain from (4.5)

$$H_0=\bigoplus_{n=1}^{N}\{W(-\lambda_n\mathcal{h}^{-1}\rho)\,\Omega(\mathcal{h})W(\lambda_n\mathcal{h}^{-1}\rho)-\lambda_n^2(\rho,\mathcal{h}^{-1}\rho)\,\mathbb{1}/2\}.$$

Thus, with (3.5),

$$[U_0^{\lambda}(t)]_{jk}=\delta_{jk}\exp\{i\lambda_j^2(\rho,\mathcal{h}^{-1}\rho)t/2\}\,W(-\lambda_j\mathcal{h}^{-1}\rho)\,U_0(t)\,W(\lambda_j\mathcal{h}^{-1}\rho).$$

From this, using the commutation relations (3.3) and (3.6) we obtain the matrix elements of $F(t)$

$$[F(t)]_{jk} = F_{jk} \exp\{i(\lambda_k^2 - \lambda_j^2)(t(\rho,\hbar^{-1}\rho) + \mathrm{Im}(\hbar^{-1}\rho,\omega_t\hbar^{-1}\rho))/2\} W((\lambda_j - \lambda_k)\{\omega_t - 1\}\hbar^{-1}\rho).$$

With the notation introduced in (d) above, and (4.4) and (4.3),

$$[F(t)]_{jk} = F_{jk} \exp\{i\alpha(j,k;t;\lambda;\rho)\} W(\xi(j,k;t;\lambda;\rho)). \tag{4.8}$$

Repeated use of the commutation relations (3.3) and (4.8) leads to

$$[F(t_1)F(t_2)\cdots F(t_n)]_{jk} = \sum_{\vec{l}_{n-1}=1}^{N} F(j,\vec{l}_{n-1},k) \exp\{i\alpha(j,\vec{l}_{n-1},k;\vec{t}_n;\lambda;\rho)\} W(\xi(j,\vec{l}_{n-1},k;\vec{t}_n;\lambda;\rho)). \tag{4.9}$$

Furthermore,

$$[U_0^\lambda(-t)]_{jj} W(g)[U_0^\lambda(t)]_{kk} = \exp\{i\alpha(j,k;t;\lambda;\rho) - i(\lambda_j + \lambda_k)\,\mathrm{Im}(g,\zeta(\rho,t))/2\} W(\xi(j,k;t;\lambda;\rho) + \omega_{-t} g). \tag{4.10}$$

Using (4.9), (4.10), and the formula

$$\alpha(j,k;t;\lambda;\rho) + \alpha(k,\vec{l}_{n-1},m;\vec{t}_n;\lambda;\rho) - \mathrm{Im}(\xi(j,k;t;\lambda;\rho),\xi(k,\vec{l}_{n-1},m;\vec{t}_n;\lambda;\rho))/2 = \alpha(j,k,\vec{l}_{n-1},m;t,\vec{t}_n;\lambda;\rho),$$

we get from (4.6)

$$\begin{aligned}
&[U^\lambda(t)]^*_{lj} W(g)[U^\lambda(t)]_{mk} \\
&= \delta_{jl}\delta_{mk} \exp\left\{-\frac{i(\lambda_j+\lambda_k)\,\mathrm{Im}(g,\zeta(\rho,t))}{2} + i\alpha(j,k;t;\lambda;\rho)\right\} W(\xi(j,k;t;\lambda;\rho) + \omega_{-t} g) \\
&\quad + \delta_{jl}\left\{\exp\left\{-\frac{i(\lambda_j+\lambda_m)\,\mathrm{Im}(g,\zeta(\rho,t))}{2}\right\} \sum_{n=1}^{\infty} (-i)^n \sum_{\vec{l}_{n-1}=1}^{N} F(m,\vec{l}_{n-1},k)\, T\int_0^t dt^{(n)}\right. \\
&\quad \times \exp\left\{i\alpha(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho) - \frac{i\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t} g,\lambda;\rho)}{2}\right\} \\
&\quad \left.\times W(\xi(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho) + \omega_{-t} g)\right\} + \delta_{mk}\left\{\exp\left\{-\frac{i(\lambda_l+\lambda_k)\,\mathrm{Im}(g,\zeta(\rho,t))}{2}\right\}\right. \\
&\quad \times \sum_{n=1}^{\infty} i^n \sum_{\vec{l}_{n-1}}^{N} \overline{F(l,\vec{l}_{n-1},j)}\, T\int_0^t dt^{(n)} \exp\left\{-i\alpha(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda;\rho)\right. \\
&\quad \left.\left.-\frac{i\beta(l,\vec{l}_{n-1};j;\vec{t}_n;\omega_{-t} g,\lambda;\rho)}{2}\right\} W(-\xi(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda;\rho) + \omega_{-t} g)\right\} \\
&\quad + \exp\left\{-\frac{i(\lambda_l+\lambda_m)\mathrm{Im}(g,\zeta(\rho,t))}{2}\right\} \sum_{n=1}^{\infty}\sum_{p=1}^{\infty} (-i)^n i^p \sum_{\vec{l}_{n-1}=1}^{N}\sum_{\vec{m}_{p-1}=1}^{N} \overline{F(l,\vec{m}_{p-1},j)} \\
&\quad \times F(m,\vec{l}_{n-1},k)\, T\int_0^t ds^{(p)}\, T\int_0^t dt^{(n)} \exp\{i\alpha(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho) - i\alpha(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;\lambda;\rho) \\
&\quad + i\alpha(m,l;t;\lambda;\rho) - i\beta(l,\vec{m}_{p-1},j;\vec{s}_p;\omega_{-t} g,\lambda;\rho)/2 - i\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t} g,\lambda;\rho)/2 \\
&\quad + i\,\mathrm{Im}(\xi(l,\vec{m}_{p-1},j;\vec{s}_p;\lambda;\rho),\xi(m,\vec{l}_{n-1},k;\vec{t}_n;\lambda;\rho))/2\} W(\xi(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho) \\
&\quad - \xi(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;\lambda;\rho) + \xi(m,l;t;\lambda;\rho) + \omega_{-t} g).
\end{aligned} \tag{4.11}$$

Using (3.7), (4.7), and (4.11), we obtain our expression for $T_{lm}^{jk}(g;f,\lambda,t)$.

(B) For $\rho \in \mathfrak{S}$, and $\delta \geqslant 0$, we let

$$\rho_\delta(\mathbf{k}) = \begin{cases} \rho(\mathbf{k}), & \text{if } |\mathbf{k}| \geqslant \delta, \\ 0 & \text{if } |\mathbf{k}| < \delta. \end{cases}$$

We let $H^{\lambda,\delta}$ [resp. $U^{\lambda,\delta}(t)$] be the Hamiltonian (resp. unitary group) given by (3.8) [resp. (3.10)] with ρ replaced by ρ_δ. Omitting superscripts and indices 0 for $\delta = 0$, we assume (i) $\rho \in \mathrm{Dom}(\hbar^{-1/2})$, i.e., (3.9); (ii) $\rho_\delta \in \mathrm{Dom}(\hbar^{-1})$, for $\delta > 0$; and (iii) $\lim_{\delta\to 0} U^{\lambda,\delta}(t) = U^\lambda(t)$ in strong operator topology. All these conditions are satisfied as soon as ρ belongs to the Schwartz space over $\mathbb{R}^3$. We infer from (iii) and the definition $\Phi_{A,g}(f,\lambda,t)$ that $T_{lm}^{jk}(g;f,\lambda,t;\rho_\delta)$ depends continuously on δ. We have obtained an infinite series expression for T_{lm}^{jk} for $\delta > 0$. Since the infinite series involved are uniformly convergent with respect to $\|\rho_\delta\|$, we may let $\delta \to 0$ term by term. Since $\gamma(\rho_\delta,t)$ and $\zeta(\rho_\delta,t)$ are continuous functions of δ, so are α, β, and ξ. This completes the proof.

Corollary 1: Let $\epsilon \geqslant 0$. Then

$$\lim_{\epsilon\to 0} \Phi_{A,\epsilon g}(\epsilon^{-1} f,\epsilon\lambda,t)_{jk} = \exp\{i \operatorname{Im}(\omega_t f, g)\} \sum_{l=1}^{N}\sum_{m=1}^{N} A_{lm}\Theta^{jk}_{lm}(f,\lambda,t),$$

where

$$\Theta^{jk}_{lm}(f,\lambda,t) = \delta_{jl}\delta_{mk}\exp\{i\beta(j,k;t;f,\lambda;\rho)\} + \delta_{jl}\left\{\sum_{n=1}^{\infty}(-i)^n \sum_{\vec{l}_{n-1}=1}^{N} F(m,\vec{l}_{n-1},k)\right.$$

$$\times T\int_0^t dt^{(n)} \exp\{i\beta(j,m,\vec{l}_{n-1},k,t,\vec{t}_n;f,\lambda;\rho)\}\Big\} + \delta_{mk}\left\{\sum_{n=1}^{\infty} i^n \sum_{\vec{l}_{n-1}=1}^{N} \overline{F(l,\vec{l}_{n-1},j)}\right.$$

$$\times T\int_0^t dt^{(n)} \exp\{-i\beta(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda,\rho)\}\Big\} + \sum_{n=1}^{\infty}\sum_{p=1}^{\infty}(-i)^n i^p$$

$$\times \sum_{\vec{m}_{p-1}=1}^{N}\sum_{\vec{l}_{n-1}=1}^{N} \overline{F(l,\vec{m}_{p-1},j)}F(m,\vec{l}_{n-1},k)\, T\int_0^t ds^{(p)}\, T\int_0^t dt \exp\{-i\beta(l,m;t;f,\lambda;\rho)$$

$$-i\beta(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;f,\lambda;\rho) + i\beta(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;f,\lambda;\rho)\}.$$

The series are absolutely convergent uniformly in f, ρ, λ, and t, for t in any bounded subset of $\mathbb{R}$.

Proof: Using the fact that the infinite series in the expression of the proposition for T^{jk}_{lm} are uniformly convergent in f, g, and λ, we perform the limit $\lim_{\epsilon\to 0} T^{jk}_{lm}(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t)$ term by term. Using the scaling properties (4.2), we get

$$\lim_{\epsilon\to 0} T^{jk}_{lm}(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t) = \exp\{i \operatorname{Im}(\omega_t f, g)\}\, \Theta^{jk}_{lm}(f,\lambda,t).$$

The assertion concerning convergence of the series follows exactly as in the Proposition. This completes the proof.

Corollary 2: Let $\epsilon > 0$. Then $T^{jk}_{lm}(\sigma, g; f,\lambda,t)$, $\sigma = 1,2,\ldots$, defined recursively by

$$T^{jk}_{lm}(\sigma; g; f,\lambda,t) := \lim_{\epsilon\to 0} T^{jk}_{lm}(\epsilon;\sigma; g; f,\lambda,t),$$

where

$$T^{jk}_{lm}(\epsilon;\sigma; g; f,\lambda,t) := \epsilon^{-2\sigma}\left\{T^{jk}_{lm}(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t) - \sum_{\nu=0}^{\sigma-1}\epsilon^{2\nu}\, T^{jk}_{lm}(\nu; g; f,\lambda,t)\right\}$$

and

$$T^{jk}_{lm}(0; g; f,\lambda,t) := \lim_{\epsilon\to 0} T^{jk}_{lm}(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t) = \exp\{i \operatorname{Im}(\omega_t f, g)\}\, \Theta^{jk}_{lm}(f,\lambda,t),$$

exists and is given by

$$T^{jk}_{lm}(\sigma; g; f,\lambda,t)$$

$$= \frac{1}{\sigma!}\exp\{i \operatorname{Im}(\omega_t f, g)\}\left[\delta_{jl}\delta_{mk}\exp\{i\beta(j,k;t;\lambda;\rho)\}\left[-\frac{i(\lambda_j+\lambda_k)\operatorname{Im}(g,\zeta(\rho,t))}{2}\right.\right.$$

$$\left.+ i\alpha(j,k;t;\lambda;\rho) - \frac{\|\xi(j,k;t;\lambda;\rho) + \omega_{-t}g\|^2}{4}\right]^{\sigma} + \delta_{jl}\left\{\sum_{n=1}^{\infty}(-i)^n \sum_{\vec{l}_{n-1}=1}^{N} F(m,\vec{l}_{n-1},k)\, T\int_0^t dt^{(n)}\right.$$

$$\times\exp\{i\beta(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;f,\lambda;\rho)\}\left[-\frac{i(\lambda_j+\lambda_m \operatorname{Im}(g,\zeta(\rho,t))}{2} + i\alpha(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)\right.$$

$$\left.\left.-\frac{i\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t}g,\lambda;\rho)}{2} - \frac{\|\xi(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho) + \omega_{-t}g\|^2}{4}\right]^{\sigma}\right\} + \delta_{mk}\left\{\sum_{n=1}^{\infty} i^n \sum_{\vec{l}_{n-1}=1}^{N} \overline{F(l,\vec{l}_{n-1},j)}\right.$$

$$\times T\int_0^t dt^{(n)}\exp\{-i\beta(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda;\rho)\}\left[-\frac{i(\lambda_l+\lambda_k \operatorname{Im}(g,\zeta,(\rho,t))}{2}\right.$$

$$\left.\left.- i\alpha(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda;\rho) - \frac{i\beta(l,\vec{l}_{n-1},j;\vec{t}_n;\omega_{-t}g,\lambda;\rho)}{2} - \frac{\|\xi(k,l,\vec{l}_{n-1},j;t,\vec{t}_n;\lambda,\rho) - \omega_{-t}g\|^2}{4}\right]^{\sigma}\right\} + \sum_{n=1}^{\infty}\sum_{p=1}^{\infty}(-i)^n i^p$$

$$\times \sum_{\vec{m}_{p-1}=1}^{N}\sum_{\vec{l}_{n-1}=1}^{N} \overline{F(l,\vec{m}_{p-1},j)}\, F(m,\vec{l}_{n-1},k)\, T\int_0^t ds^{(p)}\, T\int_0^t dt^{(n)}$$

$$\times\exp\{-i\beta(l,m;t;f,\lambda;\rho) + i\beta(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;f,\lambda;\rho) - i\beta(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;f,\lambda;\rho)\}$$

$$\times\Bigg[-\frac{i(\lambda_l+\lambda_m\ \mathrm{Im}(g,\zeta(\rho,t))}{2}-\frac{i\beta(l,\vec{m}_{p-1},j;\vec{s}_p;\omega_{-t}g,\lambda;\rho)}{2}-\frac{i\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t}g,\lambda;\rho)}{2}$$
$$+\frac{i\,\mathrm{Im}(\xi(l,\vec{m}_{p-1},j;\vec{s}_p;\lambda;\rho),\xi(m,\vec{l}_{n-1},k;\vec{t}_n;\lambda;\rho))}{2}-i\alpha(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;\lambda;\rho)+i\alpha(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)$$
$$-i\alpha(l,m;t;\lambda;\rho)-\frac{\|\xi(l,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)-\xi(m,l,\vec{m}_{p-1},j;t,\vec{s}_p;\lambda;\rho)-\xi(l,m;t;\lambda;\rho)+\omega_{-t}g\|^2}{4}\Bigg]^{\sigma}\Bigg].$$

The series are absolutely convergent, uniformly in f,g,λ,ρ, and t, for $0\leqslant\lambda\leqslant\lambda_0$, $\|g\|\leqslant X_1$, $\|\rho\|\leqslant X_2$, $|t|\leqslant X_3$, where λ_0, X_1, X_2, X_3 are arbitrary positive reals.

Proof: In the formula for T^{jk}_{lm} obtained in the Proposition, we scale $g\to\epsilon g$, $\lambda\to\epsilon\lambda$, $f\to(1/\epsilon)f$. In the summands of the infinite series, ϵ appears as [recall (4.2)]

$$\exp\{i\epsilon^2\alpha(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)$$
$$-i\epsilon^2\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t}g,\lambda;\rho)/2$$
$$-\epsilon^2\|\xi(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)+\omega_{-t}g^2\|/4\}.$$

Using the exponential series, and interchanging the order of summations, we get

$$T^{jk}_{lm}(\epsilon g;\epsilon^{-1}f,\epsilon\lambda,t)\sim\sum_{\sigma=0}^{\infty}\epsilon^{2\sigma}T^{jk}_{lm}(\sigma;g;f,\lambda,t),$$

where $T^{jk}_{lm}(\sigma;g;f,\lambda,t)$ is given by the expression in the statement. If $T^{jk}_{lm}(\sigma;g;f,\lambda,t)$ is well defined by this expression, then, by construction,

$$T^{jk}_{lm}(\sigma,g;f,\lambda,t)=\lim_{\epsilon\to0}T^{jk}_{lm}(\epsilon;\sigma;g;f,\lambda,t).$$

We now pass to the proof of the convergence properties. The estimates [see (a)–(c)]

$$|\lambda_n|\leqslant\|V\|,\quad n=1,2,\ldots,N,$$

where $\|V\|$ is the norm of $V\in\mathfrak{B}(\mathbb{C}^N)$,

$$|\gamma(f,t)|\leqslant\|f\|^2t^2/2,\quad\text{for } t\geqslant0,$$
$$\|\xi(f,t)\|\leqslant\|f\|t,\quad\text{for } t\geqslant0$$

are elementary. They lead to

$$|\alpha(m_\kappa,m_{\kappa+1},\ldots,m_n;t_{\kappa+1},t_{\kappa+2},\ldots,t_n;\lambda;g)|$$
$$\leqslant\lambda^2\|V\|^2\|g\|^2t^2[n-\kappa+3n(n-2\kappa-1)]/2,$$
$$|\beta(m_\kappa,m_{\kappa+1},\ldots,m_n;t_{\kappa+1},t_{\kappa+2},\ldots,t_n;f,\lambda;g)|$$
$$\leqslant2\lambda\|V\|\,\|f\|\,\|g\|\,t(n-\kappa),$$
$$\|\xi(m_\kappa,m_{\kappa+1},\ldots,m_n;t_{\kappa+1},t_{\kappa+2},\ldots,t_n;\lambda;g)\|$$
$$\leqslant2\lambda\|V\|\,\|g\|t(n-\kappa),$$

when $t\geqslant t_{\kappa+1}\geqslant t_{\kappa+2}\geqslant\cdots t_n\geqslant0$. These estimates can be used to obtain bounds on the absolute values of the infinite series defining $T^{jk}_{lm}(\sigma;g;f,\lambda,t)$. We treat the first of these series in detail; the other two can be treated analogously. Using Δ, as introduced in the proof of the Proposition, the absolute value of the first infinite series is bounded by ($t\geqslant0$)

$$\frac{1}{\sigma!}\sum_{n=1}^{\infty}\sum_{\vec{l}_{n-1}=1}^{N}\Delta^nT\int_0^t dt^{(n)}$$
$$\times|-i(\lambda_j+\lambda_m)\,\mathrm{Im}(g,\zeta(\rho,t))/2$$
$$+i\alpha(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)$$
$$-i\beta(m,\vec{l}_{n-1},k;\vec{t}_n;\omega_{-t}g;\lambda;\rho)/2$$
$$-\|\xi(j,m,\vec{l}_{n-1},k;t,\vec{t}_n;\lambda;\rho)+\omega_{-t}g\|^2/4|^{\sigma}$$
$$\leqslant\frac{1}{\sigma!}\sum_{n=1}^{\infty}\sum_{\vec{l}_{n-1}=1}^{N}\Delta^nT\int_0^t dt^{(n)}\,a(t,n)^{\sigma}$$
$$=\frac{1}{\sigma!}\sum_{n=1}^{\infty}\sum_{\vec{l}_{n-1}=1}^{N}\frac{a(t,n)^{\sigma}\Delta^nt^n}{n!}$$
$$=\frac{1}{\sigma!}N^{-1}\sum_{n=1}^{\infty}\frac{a(t,n)^{\sigma}\Delta^nt^nN^n}{n!},\tag{4.12}$$

where

$$a(t,n):=2\lambda\|V\|\,\|g\|\,\|\rho\|t+2\lambda\|V\|\,\|g\|\,\|\rho\|tn$$
$$+\lambda^2\|V\|^2\|\rho\|^2t^2[n+1+3n(n+1)]$$
$$\times\tfrac{1}{2}+[\|g\|+2\lambda\|V\|\,\|\rho\|t(n+1)]^2/4$$
$$=\lambda^2\|V\|^2\|\rho\|^2t^2(n+1)((5n/2)+\tfrac{3}{2})$$
$$+4\lambda\|V\|\,\|g\|\,\|\rho\|\,t(n+1)+\|g\|^2/4.$$

Now, since $\frac{3}{2}\leqslant\frac{5}{2}\leqslant16$ and $1\leqslant n+1\leqslant2n$, we have

$$a(t,n)\leqslant\tfrac{5}{2}\lambda^2\|V\|^2\|\rho\|^2t^2(n+1)^2$$
$$+4\lambda\|V\|\,\|g\|\,\|\rho\|t(n+1)+\|g\|^2/4$$
$$\leqslant16\lambda^2\|V\|^2\|\rho\|^2t^2(n+1)^2$$
$$+4\lambda\|V\|\,\|g\|\,\|\rho\|t(n+1)+\|g\|^2/4$$
$$=[4\lambda\|V\|\,\|\rho\|t(n+1)+\|g\|^2/2]^2$$
$$\leqslant[4\lambda\|V\|\,\|\rho\|t(n+1)+\|g\|^2(n+1)/2]^2$$
$$=(n+1)^2[4\lambda\|V\|\,\|\rho\|t+\|g\|^2/2]^2$$
$$\leqslant4n^2[4\lambda\|V\|\,\|\rho\|t+\|g\|^2/2]^2.$$

Inserting this into the bound (4.12) we get the absolute value of the first infinite series is bounded by

$$\left[\frac{[2(4\lambda\|V\|\,\|\rho\|t+\|g\|^2/2)]^{2\sigma}N^{-1}}{\sigma!}\right]\sum_{n=1}^{\infty}\frac{n^{2\sigma}(\Delta Nt)^n}{n!}$$
$$=[[2(4\lambda\|V\|\,\|\rho\|t+\|g\|^2/2)]^{2\sigma}N^{-1}/\sigma!]$$
$$\times\exp\{\Delta Nt\}\sum_{m=1}^{2\sigma}(\Delta Nt)^m\mathscr{S}^{(m)}_{2\sigma}.$$

Here, the $\mathscr{S}^{(m)}_n$ are the Stirling numbers of the second kind,[12] defined as the number of ways an n-element set can be partitioned into m disjoint subsets. The $\mathscr{S}^{(m)}_n$ are given by

$$\mathscr{S}_n^{(m)} = \sum_{k=0}^{m} \frac{(-)^{m-k} k^n}{k!(m-k)!}, \quad m = 1,2,\ldots,n, \quad n = 1,2,\ldots,$$

and one has

$$\sum_{k=0}^{\infty} \frac{x^k k^n}{k!} = \exp\{x\} \sum_{m=1}^{n} x^m \mathscr{S}_n^{(m)}.$$

This completes the proof.

Proof of Theorem 1(a)(1): With $g = 0$ in Corollary 1, we have

$$A_{(0)}(f,\lambda,t)_{jk} := \lim_{\epsilon \to 0} A(\epsilon^{-1} f, \epsilon\lambda, t)_{jk}$$
$$= \sum_{l=1}^{N} \sum_{m=1}^{N} A_{lm} \Theta_{lm}^{jk}(f,\lambda,t),$$

proving existence, and providing an explicit expression for $A_{(0)}(f,\lambda,t)$. We have to prove that the equation of motion

$$i\dot{A}_{(0)}(f,\lambda,t) = [A_{(0)}(f,\lambda,t), H(f,\lambda,t)]_-,$$
$$A_{(0)}(f,\lambda,0) = A \tag{4.13}$$

is satisfied. The solution of (4.13) is given by

$$\tilde{U}^{\lambda}(t)^* A\, \tilde{U}^{\lambda}(t), \quad t \in \mathbb{R}, \tag{4.14}$$

where

$$\dot{\tilde{U}}^{\lambda}(t) = -iH(f,\lambda,t)\, \tilde{U}^{\lambda}(t), \quad \tilde{U}^{\lambda}(0) = \mathbf{1}.$$

Let

$$H\partial_0(f,\lambda,t) = \lambda\, \mathrm{Re}(\omega_t f, \rho) V.$$

The solution of

$$\dot{\tilde{U}}_0^{\lambda}(t) = -iH_0(f,\lambda,t)\, \tilde{U}_0^{\lambda}(t), \quad \tilde{U}_0^{\lambda}(0) = \mathbf{1}$$

is given by

$$\tilde{U}_0^{\lambda}(t) = \exp\left\{ -i \int_0^t ds\, H_0(f,\lambda,s) \right\},$$

with matrix elements (recall that V is diagonal)

$$\tilde{U}_0^{\lambda}(t)_{jk} = \delta_{jk} \exp\left\{ -i\lambda v_j \int_0^t ds\, \mathrm{Re}(\omega_s f, \rho) \right\}.$$

With (c), we have

$$\tilde{U}_0^{\lambda}(t)_{jk} = \delta_{jk} \exp\{ -i\lambda_j\, \mathrm{Im}(f, \xi(\rho, -t)) \}.$$

Letting

$$\tilde{F}(t) := \tilde{U}_0^{\lambda}(t)^* F \tilde{U}_0^{\lambda}(t), \quad t \in \mathbb{R},$$

we have

$$\tilde{U}^{\lambda}(t) = \tilde{U}_0^{\lambda}(t) \left\{ \mathbf{1} + \sum_{n=1}^{\infty} (-i)^n T \int_0^t dt^{(n)} \right.$$
$$\left. \times \tilde{F}(t_1) \tilde{F}(t_2) \cdots \tilde{F}(t_n) \right\}.$$

For the matrix elements of $\tilde{F}(t)$ we get

$$\tilde{F}(t)_{jk} = F_{jk} \exp\{ i\beta(j,k;t;f,\lambda;\rho) \}.$$

Thus

$$(\tilde{F}(t_1)\tilde{F}(t_2)\cdots\tilde{F}(t_n))_{jk}$$
$$= \sum_{\vec{l}_{n-1}=1}^{N} F(j,\vec{l}_{n-1},k) \exp\{ i\beta(j,\vec{l}_{n-1},k;\vec{t}_n;f,\lambda;\rho) \}.$$

We then have, for the matrix elements of $\tilde{U}^{\lambda}(t)$,

$$\tilde{U}^{\lambda}(t)_{jk} = \exp\{ -i\lambda_j\, \mathrm{Im}(f, \xi(\rho,t)) \}$$
$$\times \left\{ \delta_{jk} + \sum_{n=1}^{\infty} (-\mathrm{i})^n \sum_{\vec{l}_{n-1}=1}^{N} F(j,\vec{l}_{n-1},k)\, T \int_0^t dt^{(n)} \right.$$
$$\left. \times \exp\{ i\beta(j,\vec{l}_{n-1},k;\vec{t}_n;f,\lambda;\rho) \} \right\}. \tag{4.15}$$

From (4.14)

$$(\tilde{U}^{\lambda}(t)^* A \tilde{U}^{\lambda}(t))_{jk} = \sum_{l=1}^{N} \sum_{m=1}^{N} A_{lm} (\tilde{U}^{\lambda}(t)^*)_{jl}\, \tilde{U}^{\lambda}(t)_{mk},$$

but from (4.15),

$$\Theta_{lm}^{jk}(f,\lambda,t) = (\tilde{U}^{\lambda}(t)^*)_{jl}\, \tilde{U}^{\lambda}(t)_{mk}.$$

Thus,

$$A_{(0)}(f,\lambda,t) = \tilde{U}^{\lambda}(t)^* A \tilde{U}_{\lambda}(t), \tag{4.16}$$

which completes the proof.

Remark: Since $\tilde{U}^{\lambda}(t)$ is unitary, (4.16) with $A = \mathbf{1} \in \mathfrak{B}(\mathbb{C}^N)$ implies

$$\sum_{m=1}^{N} \Theta_{mm}^{jk}(f,\lambda,t) = \delta_{jk}. \tag{4.17}$$

Proof of Theorem 1(a)(2): It follows directly from Corollary 1, upon setting $A = \mathbf{1}$, and using (4.17), that

$$\lim_{\epsilon \to 0} \Phi_{\mathbf{1},\epsilon g}(\epsilon^{-1} f, \epsilon\lambda, t) = \exp\{ \mathrm{i}\, \mathrm{Im}(\omega_t f, g) \} \mathbf{1}.$$

This implies our assertion and exhibits again the classical nature of the limit in the field.

Proof of Theorem 1(b): This follows directly from Corollary 2. We get

$$A_{(\sigma)}(f,\lambda,t)_{jk} = \sum_{l=1}^{N} \sum_{m=1}^{N} A_{lm}\, T_{lm}^{jk}(\sigma;0;f,\lambda,t),$$
$$\omega_{(\sigma)}(g;f,\lambda,t) = \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{m=1}^{N} T_{mm}^{jk}(\sigma;g;f,\lambda,t)\, D_{kj}.$$

This completes the proof.

V. CONCLUSION

The objective of this paper was to give a mathematically rigorous justification of the use of time-dependent Hamiltonians in quantum mechanics. Such Hamiltonians have a wide range of applications, especially in the fields of atomic and molecular spectroscopy, laser physics, and nonlinear optical phenomena.

The fully quantized Hamiltonian used as a starting point for the derivation of a time-dependent Hamiltonian has itself approximate character. It is the Hamiltonian commonly used in practical applications to describe quantum effects of the electromagnetic field.[13]

Apart from the proof of the existence of the semiclassical dynamics we have shown that the observables describing the fluctuations around the semiclassical motion exist in any order, for the field as well as for the N-level system. The fluctuation observables account for those phenomena like spontaneous emission and back action of the N-level system on the field that are beyond the semiclassical dynamics. For all the observables pertaining to the semiclassical and the

fluctuation dynamics we obtained explicit expressions for their time dependence in the form of infinite series.

The dynamics of the fully quantized system was represented by a formal asymptotic series with the semiclassical dynamics as asymptotic representation. These series cannot be used to numerically improve the semiclassical solutions. However, in practical applications one is never interested in the entire series. Instead, only those terms need to be considered that account for the effects outside the scope of the semiclassical approximation one wishes to describe.

ACKNOWLEDGMENT

We gratefully acknowledge discussions with Peter Pfeifer and Hans Primas.

[1]H. Primas, *Chemistry, Quantum Mechanics and Reductionism* (Springer, Berlin, 1983).
[2]K. Hepp, Commun. Math. Phys. **35**, 265 (1974).
[3]E. A. Power, *Introductory Quantum Electrodynamics* (Longmans, Green, London, 1964).
[4]For a quantitative discussion of the relevance of the A^2 terms, see P. Pfeifer, J. Phys. A: Math. Gen. **14**, L129 (1981). If the A^2 terms are kept, then they can be transformed away by a so-called Bogolyubov transformation, if the field is assumed to consist of finitely many modes; see Theorem 8.1 of F. A.Berezin, *The Method of Second Quantization* (Academic, New York, 1968), or alternatively J. L. van Hemmen, Z. Phys. B **38**, 271 (1980).
[5]P. Pfeifer, "Chiral molecules—a superselection rule induced by the radiation field," dissertation ETH No. 6551, ok Gotthard S + D AG, Zürich, 1980, pp. 31 ff.
[6]This idea is perhaps as old as the quantum theory of radiation itself. It is widely used, useful, successful, and thus, *a posteriori*, well motivated. Our results provide some mathematical basis for this well-established folklore.
[7]G. G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Wiley–Interscience, New York, 1972), Sec. 1.1.c. This is our standard reference for the Fock space methods which are used freely in this work.
[8]E. B. Davies, Ann. Inst. H. Poincaré A **35**, 149 (1981), particularly Lemma 1.
[9]O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics II* (Springer, Berlin, 1981), Proposition 5.2.4.
[10]See Ref. 7, Theorem 7 of Chap. 3 and the discussion following.
[11]J. M. Cook, J. Math. Phys. **2**, 33 (1961), Theorem 1. See also Ref. 7, p. 19.
[12]*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965), 5th printing, Sec. 24.1.4.
[13]H. Haken, "Laser Physics," in *Handbuch der Physik* (Springer, Berlin, 1969), Vol. 25/2C.

On the algebras of local observables in the generalized sense in quantum mechanics

K. Kong Wan
Physics Department, St. Andrews University, St. Andrews, Scotland

R. G. D. McLean
Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada

(Received 29 October 1984; accepted for publication 29 March 1985)

The notion of observables localized with respect to a general spectral measure is introduced as a generalization of the localization with respect to the spectral measure of the position observable. The physical features of these observables are discussed and certain of their algebraic and lattice theoretic properties are presented.

I. INTRODUCTION

Local observables in quantum field theory have been extensively studied since Haag and Kastler[1] introduced their algebraic formulation of quantum field theory. However, less attention was paid to local observables in the context of quantum mechanics. Based on a study of asymptotic localization and separation of states[2,3] the authors recently formulated an algebraic theory of asymptotically separable quantum mechanics[4] in which local observables play an important part. In quantum measurement theory local observables also play a part and they can be related to measuring devices of finite size.[5,6] It appears that these observables do deserve a systematic study in the context of quantum mechanics. In this paper we shall introduce a generalized notion of local observables and present some mathematical results relevant to quantum mechanics.

II. LOCAL OBSERVABLES IN THE GENERALIZED SENSE

Let us consider quantum systems whose Hilbert space is $\mathscr{H} = L^2(\mathbf{R}^n)$ where $\mathbf{R}^n$ is the configuration space. Let $\mathbf{B}(\mathscr{H})$ denote the von Neumann algebra of all bounded operators on $\mathscr{H}$. Denote by $\mathscr{B}(\mathbf{R}^k)$, $k = 1,2,...$, the Borel sets of $\mathbf{R}^k$ and by $\mathscr{B}_c(\mathbf{R}^k)$ the bounded Borel sets of $\mathbf{R}^k$. A spectral measure on the measurable space $(\mathbf{R}^k, \mathscr{B}(\mathbf{R}^k))$, whose values are projectors on $\mathscr{H}$, will be denoted by E (see Ref. 7). We shall now introduce observables which are localized with respect to a general spectral measure E. In particular, E may be the spectral measure of the position operator, then we will obtain the observables localized in the configuration space.[5,6] Alternatively if E is a momentum spectral measure then we are establishing a "localization" in the momentum space of the system. If $k = 1$ then E may be identified with the spectral measure of a self-adjoint operator in $\mathscr{H}$ (see Ref. 7). Then we have a "localization" in the spectrum of that operator. Throughout this paper $\mathscr{H}(E(R))$ will denote the range of the projector $E(R)$, $R \in \mathscr{B}(\mathbf{R}^k)$, i.e., $\mathscr{H}(E(R)) = \{\Phi \in \mathscr{H} : E(R)\Phi = \Phi\}$. The orthogonal complement of $\mathscr{H}(E(R))$ in $\mathscr{H}$ is then $\mathscr{H}(E(R^c))$, $R^c = \mathbf{R}^k - R$. The spectral measure of a self-adjoint operator A in $\mathscr{H}$ will be denoted by $E(A;\cdot)$.

Definition D2.1: A self-adjoint operator A in $\mathbf{B}(\mathscr{H})$ is said to be localized with respect to $E(R)$, $R \in \mathscr{B}_c(\mathbf{R}^k)$ [or simply $E(R)$ local] if

(1) for all unit vectors $\psi \in \mathscr{H}(E(R^c))$,

$$\langle \psi | A\psi \rangle = a, \tag{2.1}$$

where $a \in \mathbf{R}$ is independent of ψ, and

(2) for all vectors $\psi \in \mathscr{H}(E(R^c))$ and $\psi \in \mathscr{H}(E(R))$,

$$\langle \psi | A\varphi \rangle = 0. \tag{2.2}$$

This definition can be compared with a similar definition in field theory given by de Muynck.[8] Before discussing the physical meaning of this definition we shall present some immediate consequences of this definition.

Lemma 2.1: Let M be a projector and M^c its orthogonal complement in $\mathscr{H}$ and let $A \in \mathbf{B}(\mathscr{H})$ and $a \in \mathbf{C}$. Then the following are equivalent:

(1) $A = MXM + aI$, for some $X \in \mathbf{B}(\mathscr{H})$; (2.3)

(2) $A = MXM + aM^c$, for some $X \in \mathbf{B}(\mathscr{H})$; (2.4)

(3) $A = MAM + aM^c$. (2.5)

If A is self-adjoint and $a \in \mathbf{R}$ then each of these is also equivalent to

(4) $Af = af$, for all f in the range of M^c (2.6)

[range of M^c will be denoted by ran (M^c)].

Proof: [(1)⇒(2).] This follows from the identity

$$MXM + aI = M(X + aM)M + aM^c.$$

[(2)⇒(3).] If $A = MXM + aM^c$ then

$$\begin{aligned} MAM + AM^c &= M(MXM + aM^c)M + aM^c \\ &= MXM + aM^c \\ &= A. \end{aligned}$$

Since (3)⇒(1) is obvious the first three statements are equivalent and we now assume A is self-adjoint and $a \in \mathbf{R}$.

[(3)⇒(4).] If $A = MAM + AM^c$ and $f \in \text{ran}(M^c)$ then $Mf = 0$ and $Af = aM^c f = af$.

[(4)⇒(3).] It follows from (4) that $A(M^c f) = a(M^c f)$ for all f in $\mathscr{H}$, so $AM^c = aM^c$ and (taking adjoints) $M^c A = aM^c$. Now

$$\begin{aligned} A &= (M + M^c)A(M + M^c) \\ &= MAM + M^c AM + MAM^c + M^c AM^c \\ &= MAM + aM^c M + aMM^c + aM^c \\ &= MAM + aM^c. \end{aligned}$$

 0022-2488/85/102540-06$02.50

Theorem 2.1: Let A be a self-adjoint operator in $\mathbf{B}(\mathscr{H})$. Then the following statements are equivalent:

(1) A is $E(R)$ local;

(2) $A = E(R)AE(R) + aE(R^c)$, for some $a \in \mathbf{R}$; (2.7)

(3) $A = E(R)XE(R) + aI$, for some $a \in \mathbf{R}$ and some $X \in \mathbf{B}(\mathscr{H})$. (2.8)

Proof: The implications (3)$\Leftrightarrow$(2)$\Rightarrow$(1) follow from Lemma 2.1 and Definition D2.1. To prove (1)$\Rightarrow$(2) we have

$$A = (E(R) + E(R^c))A(E(R) + E(R^c))$$
$$= E(R)AE(R) + E(R^c)AE(R^c) \quad \text{by (2.2).}$$

Now (2.1) implies that for all $\psi \in \mathscr{H}(E(R^c))$,

$$\langle \psi | (A - a)\psi \rangle = 0$$

or

$$\langle \psi | [E(R^c)AE(R^c) - aE(R^c)]\psi \rangle = 0.$$

Hence

$$E(R^c)AE(R^c) - aE(R^c) = 0.$$

So

$$A = E(R)AE(R) + aE(R^c).$$

Theorem 2.2: A self-adjoint operator A in $\mathbf{B}(\mathscr{H})$ is $E(R)$ local if and only if the following conditions hold for all Borel sets S in $\mathbf{R}$:

(1) if $\psi \in \mathscr{H}(E(R^c))$, $\|\psi\| = 1$,

then $E(A;S)\psi = \|E(A;S)\psi\|\psi$, (2.9)

(2) if $\varphi \in \mathscr{H}(E(R))$,

then $E(R)E(A;S)\varphi = E(A;S)\varphi$. (2.10)

Proof: (a) Suppose A is $E(R)$ local. Then there exists X in $\mathbf{B}(\mathscr{H})$ such that $A = E(R)XE(R) + aI$. Since $E(R)$ clearly commutes with A it also commutes with $E(A;S)$ for all S. Hence $E(R)E(A;S)\varphi = E(A;S)E(R)\varphi = E(A;S)\varphi$ if $\varphi \in \mathscr{H}(E(R))$. Having demonstrated (2.10), let us prove (2.9). Since (1) and (2) clearly hold if $E(R) = I$ we may assume that $E(R) \neq I$ and it then follows from Theorem 2.1 that a is an eigenvalue of A and so $E(A;\{a\})$ projects onto the corresponding eigenspace,[9] and that $E(R^c) \leqslant E(A;\{a\})$ since every $\psi \in \mathscr{H}(E(R^c))$ satisfied $A\psi = a\psi$. For a given Borel set S in $\mathbf{R}$ consider the cases $a \notin S$ and $a \in S$ separately as follows.

(i) Suppose $a \notin S$ and $\psi \in \mathscr{H}(E(R^c))$. The projectors $E(A;\{a\})$ and $E(A;S)$ are then orthogonal, so

$$E(R^c) \leqslant E(A;\{a\}) \leqslant E(A;S^c), \quad S^c = \mathbf{R} - S$$

giving $E(R) \geqslant E(A;S)$. Hence $E(A;S)\psi = 0$ since $\|E(A;S)\psi\| \leqslant \|E(R)\psi\| = 0$. So (2.9) is satisfied.

(ii) Suppose $a \in S$ and ψ is a unit vector in $\mathscr{H}(E(R^c))$. Let $S' = S - \{a\}$, then $E(A;S) = E(A;S') + E(A;\{a\})$, and since $a \notin S'$ we have $E(A;S')\psi = 0$, as in (i) above. Hence $E(A;S)\psi = E(A;\{a\})\psi = \psi$. Since $\|\psi\| = 1$ this implies (2.9) in this case.

(b) To prove the converse suppose A satisfies (2.9) and (2.10). By (2.10) $E(A;S)E(R) = E(R)E(A;S)E(R)$ and taking adjoints gives

$$E(R)E(A;S) = E(A;S)E(R),$$

implying

$$[E(R), A] = 0 = [E(R^c), A].$$

Hence

$$A = (E(R) + E(R^c))A(E(R) + E(R^c))$$
$$= E(R)AE(R) + AE(R^c). \quad (2.11)$$

We want to show that $AE(R^c) = aE(R^c)$, $a \in \mathbf{R}$. First, observe that given $\psi \in \mathscr{H}(E(R^c))$, $\|\psi\| = 1$, Eq. (2.8) implies

$$E(A;S)\psi = \|E(A;S)\psi\|E(A;S)\psi.$$

So, $\|E(A;S)\psi\|$ equals either 0 or 1 for all S. Now for such a vector ψ we can define a function $f: \mathbf{R} \to \mathbf{R}$ by

$$f(\lambda) = \|E(A;(-\infty, \lambda])\psi\|^2, \quad \lambda \in \mathbf{R}.$$

Then f is nondecreasing and right continous.[7] It can assume values 0 and 1 only. It follows that for suitably large negative and positive λ the function f takes the values 0 and 1, respectively. Hence f has a unique discontinuity at some point a in $\mathbf{R}$ and

$$\|E(A;\{a\})\psi\|^2 = \|E(A;(-\infty, a])\psi\|^2 - \|E(A;(-\infty, a))\psi\|^2.$$

Now,[7]

$$\|E(A;(-\infty, a))\psi\|^2 = \|E\left(A; \bigcup_{n=1}^{\infty} (-\infty, a - 1/n)\right)\psi\|^2$$
$$= \|\lim_{n\to\infty} E(A;(-\infty, a - 1/n))\psi\|^2$$
$$= \lim_{n\to\infty} \|E(A;(-\infty, a - 1/n))\psi\|^2$$
$$= \lim_{n\to\infty} f(a - 1/n).$$

Hence

$$\|E(A;\{a\})\psi\|^2 = f(a) - \lim_{n\to\infty} f(a - 1/n) = f(a) = 1.$$

So, $E(A;\{a\})\psi = \psi$ and $A\psi = a\psi$. Finally we will show that this value a is independent of ψ. Let ψ' be another unit vector in $\mathscr{H}(E(R^c))$ linearly independent of ψ. Then there are real constants b, c, such that

$$A\psi' = b\psi' \quad \text{and} \quad A(\psi + \psi') = c(\psi + \psi'),$$

which clearly imply $b = a$ and $c = a$. We can now conclude that $AE(R^c) = aE(R^c)$. It follows from this and (2.11) that A is $E(R)$ local.

Theorem 2.3: Let M be a projector on $\mathscr{H}$ and let $\mathbf{B}$ denote $\mathbf{B}(\mathscr{H})$. Then $M\mathbf{B}M$ is a C^* subalgebra of $\mathbf{B}$ and the von Neumann algebra in $\mathscr{H}$ generated by $M\mathbf{B}M$ is

$$(M\mathbf{B}M)'' = M\mathbf{B}M + \mathbf{C}M^c = M\mathbf{B}M + \mathbf{C}I. \quad (2.12)$$

A self-adjoint A in $\mathbf{B}$ belongs to $(M\mathbf{B}M)''$ if and only if $A = MAM + aM^c$ for some $a \in \mathbf{R}$.

Proof: It is readily proved that $M\mathbf{B}M$ is a C^* subalgebra. To prove $(M\mathbf{B}M)'' = M\mathbf{B}M + \mathbf{C}M^c$ we begin by showing that the commutant $(M\mathbf{B}M)'$ of $M\mathbf{B}M$ (see Ref. 10) is

$$(M\mathbf{B}M)' = M^c\mathbf{B}M^c + \mathbf{C}M. \quad (2.13)$$

If $X = M^cXM + aM$ and $Y = MYM$ then $XY = aM = YX$ and this implies $M^c\mathbf{B}M^c + \mathbf{C}M \subseteq (M\mathbf{B}M)'$. Conversely, suppose $X \in (M\mathbf{B}M)'$ then X commutes with M and so

$$X = (M + M^c)X(M + M^c) = MXM + M^cXM^c.$$

Also MXM commutes with MAM for every $A \in \mathbf{B}$ so the re-

striction of MXM to ran(M) commutes with every bounded operator on ran(M) and it follows that $MXM = aM$ for some $a \in \mathbb{C}$. Hence $X = M^c XM + aM$ so $(\mathbf{M}\mathbf{B}M)' \subseteq M^c \mathbf{B} M^c + \mathbb{C}M$ and (2.13) has been verified. Now two applications of (2.13) give

$$\begin{aligned} A \in (M\mathbf{B}M)'' &\Leftrightarrow A \in (M^c \mathbf{B} M^c + \mathbb{C}M)' \\ &\Leftrightarrow A\,(M^c X M^c + aM) = M^c X M^c + aM)A \\ &\qquad (\forall X \in \mathbf{B})\ (\forall a \in \mathbb{C}) \\ &\Leftrightarrow AM^c X M^c = M^c X M^c A \quad (\forall X \in \mathbf{B}) \\ &\Leftrightarrow A \in (M^c \mathbf{B} M^c)' \\ &\Leftrightarrow A \in M\mathbf{B}M + \mathbb{C}M^c. \end{aligned}$$

Finally, the last assertion in the statement of the theorem follows from Lemma 2.1.

To conclude this section let us comment on the physical background for our Definition D2.1. The physics behind this definition is basically the same as that for the usual observables localized in the configuration space.[5,6] To be definite and explicit let us examine the case where $k = 1$ in the measurable space $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$. Then E is the spectral measure $E(B;\cdot)$ of a self-adjoint operator B on $\mathscr{H}$. Assuming B to be an observable, $E(R)$ becomes a proposition corresponding to B having a value in the Borel set R (see Ref. 11). Now suppose an experimenter has at his disposal a measuring device which cannot detect any state of the system lying outside the range $\mathscr{H}(E(R))$ of $E(R)$. In other words the device is sensitive only to values of B lying in R. An example would be a measuring device sensitive only to, say, a certain range of energy values. The question posed now is what kind of observables A are measurable with a measuring device possessing the above properties. Our Definition D2.1 answers this question with (2.1) and (2.2) reflecting the characteristic feature of the measuring device. This becomes clearer if we had set $a = 0$. Condition (2.2) implies that A does not correlate states lying outside the range of $E(R)$ with states inside the range of E(R). Condition (2.1) means the vanishing of expectation value with respect to any state not in the range of $E(R)$, a fact together with the absence of correlations implied by (2.2) indicating that a null measurement of A will result for such states. In other words A being $E(R)$ local means that A is nontrivial only on the sensitive range $\mathscr{H}(E(R))$ of $E(R)$. The result $A = E(R)XE(R)$ from Theorem 2.1 (when $a = 0$) confirms this intuition. Since the effect of $a \neq 0$ adds to $E(R)XE(R)$ only a constant term aI, which does not change the basic physical character of $E(R)XE(R)$, we have allowed $a \neq 0$ in Definition D2.1 to secure certain desirable mathematical properties of such observables. These will become apparent later in Sec. IV. A further motivation for $E(R)$-local observables manifests itself in Theorem 2.2. Despite its somewhat abstract appearance, Theorem 2.2 has a direct and straightforward physical interpretation. Intuitively one would expect an observable A that is nontrivial only on the sensitive range of $E(R)$ to be related to states and measurement as follows: (1) if the state vector does not lie in the range of $E(R)$ before a measurement of A then the state is not affected by a measurement of A; and (2) if the state vector lies in the range of $E(R)$ before a measurement of A then it lies in the range of $E(R)$ right after the measurement. Now if we confine ourselves to measurements to which Lüdders' rule[12] applies, i.e., if Φ_i is the state vector before a measurement of A, then the normalized state vector right after the measurement of A obtaining a value in the Borel set S is given by

$$\Phi_f = E(A;S)\Phi_i / \|E(A;S)\Phi_i\|.$$

We can see that (2.9) and (2.10) in Theorem 2.2 correspond, respectively, to (1) and (2) above exactly. In fact we could base our notion of $E(R)$-local observables on (1) and (2) and formally define $E(R)$-local observables through (2.9) and (2.10).

III. THE ALGEBRA OF ALL LOCAL OBSERVABLES IN THE GENERALIZED SENSE

In this section we investigate the * algebra generated by all E-local observables and show that in general this * algebra is not closed and is not dense in $\mathbf{B}(\mathscr{H})$ in the operator norm.

Definitions D3.1: (1) The E-local algebra is defined to be $\mathscr{A}_L(E) = \{E(R)AE(R):\ A \in \mathbf{B}(\mathscr{H}),\ R \in \mathscr{B}_c(\mathbb{R}^k)\}$.

(2) The E-local algebra with unity is defined to be $\mathscr{A}_I(E) = \{E(R)AE(R) + aI:\ A \in \mathbf{B}(\mathscr{H}),\ R \in \mathscr{B}_c(\mathbb{R}^k),\ a \in \mathbb{C}\}$. The norm closures of $\mathscr{A}_L(E)$ and $\mathscr{A}_I(E)$ will be denoted by $\overline{\mathscr{A}}_L(E)$ and $\overline{\mathscr{A}}_I(E)$, respectively. When there is no danger of confusion over the spectral measure involved we shall just write $\mathscr{A}_L$, $\mathscr{A}_I$, $\overline{\mathscr{A}}_L$, and $\overline{\mathscr{A}}_I$ to denote these sets.

Theorem 3.1: (1) $\mathscr{A}_I = \mathscr{A}_L + \mathbb{C}I$, $\overline{\mathscr{A}}_I = \overline{\mathscr{A}}_L + \mathbb{C}I$.

(2) $\mathscr{A}_L$ and $\mathscr{A}_I$ are irreducible * subalgebras of $\mathbf{B}(\mathscr{H})$ and the von Neumann algebra generated by each of these * subalgebras is equal to $\mathbf{B}(\mathscr{H})$.

Proof: (1) The statement $\mathscr{A}_I = \mathscr{A}_L + \mathbb{C}I$ is clearly true. Now we prove the other assertion. $\mathbf{B}(\mathscr{H})$ is a Banach space under the operator norm and is hence a topological vector space.[13] Moreover, $\overline{\mathscr{A}}_L$ is a closed subspace of $\mathbf{B}(\mathscr{H})$. Hence $\overline{\mathscr{A}}_L + \mathbb{C}I$ is also closed.[13] It follows that $\overline{\mathscr{A}}_I = \overline{\mathscr{A}}_L + \mathbb{C}I$.

(2) By the multiplicative property of strong convergence[14] every A in $\mathbf{B}(\mathscr{H})$ is the strong limit (as $j \to \infty$) of $E(R_j) \times AE(R_j)$, where $\{R_j\}$ is an increasing sequence in $\mathscr{B}_c(\mathbb{R}^k)$ converging to $\mathbb{R}^k$. Since a von Neumann algebra is closed in the strong topology, it follows that $\mathbf{B}(\mathscr{H})$ is the von Neumann algebra generated by $\mathscr{A}_L$ and hence by $\mathscr{A}_I$ also.

Let $A, B \in \mathscr{A}_L$, then there are bounded Borel sets R and Q with $A = E(R)AE(R)$ and $B = E(Q)BE(Q)$. It is easily verified that

$$A + B = E(R \cup Q)(A + B)E(R \cup Q),$$

$$AB = E(R \cup Q)(AE(R)E(Q)B)E(R \cup Q).$$

It is clear then that $\mathscr{A}_L$ is a * subalgebra and so is $\mathscr{A}_I$.

Finally noting that the von Neumann algebra generated by $\mathscr{A}_L$ equals its bicommutant[14,15] we have[14]

$$\mathscr{A}_L' = \mathscr{A}_L''' = \mathbf{B}(\mathscr{H})' = \mathbb{C}I.$$

Therefore $\mathscr{A}_L$, and hence $\mathscr{A}_I$, are irreducible.[15]

The properties of the * algebras $\mathscr{A}_L$ and $\mathscr{A}_I$ are closely related to the spectrum[16] of the spectral measure E.

Lemma 3.1: Suppose E is not a compact spectrum measure[16] on the measureable space $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$. Then the spectrum of E is the union of a disjoint sequence Q_r, $r = 1,2,3...$, in $\mathscr{B}(\mathbb{R}^k)$ such that for each r, (1) $E(Q_r) \neq 0$, (2) Q_r lies outside the closed ball of radius $r - 1$ centered at the origin in $\mathbb{R}^k$.

Proof: Let Λ be the spectrum of E; then $E(\Lambda)=I$ (see Ref. 16) and by hypothesis Λ is not compact. Define, for $j=1,2,3,\ldots$,

$$V_j=\{x\in\mathbb{R}^k:\ j-1<|x|<j\}.$$

Then for any subset Q of $\mathbb{R}^k$, $Q\cap V_j$ is a disjoint sequence of bounded Borel sets with union Q.

Let $R_j=\Lambda\cap V_j$, then $E(\cup_j R_j)=E(\Lambda)=I$. Let R be the union of all R_j for which $E(R_j)\neq 0$, then $E(R)=I$. Now suppose that $E(R_j)\neq 0$ for only finitely many j, then R is bounded. It follows that Λ is a subset of $\bar{R}$, the closure of R and Λ is hence bounded, contradicting our premise that E, hence its spectrum Λ which is a closed set, is not compact. So there are infinitely many such j with $E(R_j)\neq 0$. Now let R_{j_r} be the subsequence of $\{R_j\}$ consisting of those R_j for which $E(R_j)\neq 0$. Let D be the union of all those R_j for which $E(R_j)=0$ and define $Q_r=R_{j_r}\cup(D\cup V_r)$. Clearly we have (i) $E(Q_r)\neq 0$, (ii) Q_r lies outside the closed ball of radius $r-1$ centered at the origin since $j_r\geqslant r$, and (iii) the union of Q_r is Λ.

Theorem 3.2: The following are equivalent:

(1) E is a compact spectral measure;
(2) $\mathscr{A}_L=\mathscr{A}_I$;
(3) $\mathscr{A}_L=\mathbf{B}(\mathscr{H})$;
(4) $\mathscr{A}_I=\mathbf{B}(\mathscr{H})$;
(5) $\mathscr{A}_L$ is closed in $\mathbf{B}(\mathscr{H})$ in the uniform topology;
(6) $\mathscr{A}_I$ is closed in $\mathbf{B}(\mathscr{H})$ in the uniform topology.

Proof: It is sufficient to prove the following implications:

$$(1)\Rightarrow(2)\Rightarrow(3)\Rightarrow(5)\Rightarrow(6)\Rightarrow(1),$$

$$(3)\Rightarrow(4)\Rightarrow(6).$$

Those not proved below are obvious or follow from $\mathscr{A}_I=\mathscr{A}_L+\mathbb{C}I$ and $\bar{\mathscr{A}}_I=\bar{\mathscr{A}}_L+\mathbb{C}I$.

[(1)$\Rightarrow$(2).] Suppose E is compact, then the spectrum Λ of E is a bounded Borel set with $E(\Lambda)=I$. Hence $I\in\mathscr{A}_L$ so $\mathscr{A}_L+\mathbb{C}I=\mathscr{A}_L$ and (2) follows.

[(2)$\Rightarrow$(3).] If $\mathscr{A}_L=\mathscr{A}_I$ then $I\in\mathscr{A}_L$, which implies $I=E(R)$ for some bounded R and now

$$\mathbf{B}(\mathscr{H})=E(R)\mathbf{B}(\mathscr{H})E(R)\subseteq\mathscr{A}_L.$$

[(6)$\Rightarrow$(1).] We assume E is not compact and show that $\mathscr{A}_I$ is not closed. Let Q_r be a disjoint sequence of bounded Borel sets satisfying properties (1) and (2) of Lemma 3.1. Now the partial sums of the series $\Sigma\, 2^{-r}E(Q_{2r})$ belong to $\mathscr{A}_L$ and since this series is clearly absolutely convergent in $B(\mathscr{H})$ it converges to an element A of $\bar{\mathscr{A}}_L$. Clearly $A\in\bar{\mathscr{A}}_L+\mathbb{C}I=\bar{\mathscr{A}}_I$ and we show by contradiction that $A\notin\mathscr{A}_I$.

Suppose A does belong to $\mathscr{A}_I$, then $A=E(R)AE(R)+aE(R^c)$ for some bounded Borel set R and some $a\in\mathbb{C}$ so

$$\sum_{r=1}^{\infty}2^{-r}E(Q_{2r})=\sum_{r=1}^{\infty}2^{-r}E(R\cap Q_{2r})+aE(R^c).$$

It follows from property (2) of Lemma 3.1 that $R\cap Q_{2j+1}=\varnothing$ for some j, so multiplying each side of the above equality by $E(Q_{2j+1})$ and using the fact that Q_r is a disjoint sequence we have

$$0=aE(Q_{2j+1})$$

and it follows from property (1) of the lemma that $a=0$. Hence for all f in $\mathscr{H}$,

$$\sum_{r=1}^{\infty}2^{-r}E(Q_{2r})f=\sum_{r=1}^{\infty}2^{-r}E(R\cap Q_{2r})f.$$

Now by property (2) of the lemma $R\cap Q_{2i}=\varnothing$ for some integer i and by (1) of the lemma there is a nonzero vector f in the range of $E(Q_{2i})$. For such an f the last identity gives

$$2^{-i}f=0,$$

which is a contradition since $f\neq 0$. We conclude that $A\in\bar{\mathscr{A}}_I-\mathscr{A}_I$ so $\mathscr{A}_I$ is not closed.

IV. GENERALIZED LOCAL OBSERVABLES AND QUANTUM LOGIC

In this section we shall study some properties of projectors in the * algebras $\mathscr{A}_I$ and $\bar{\mathscr{A}}_I$ defined with respect to a given spectral measure E.

In the quantum logic approach to quantum mechanics it is usually assumed that the propositions associated with a physical system form an orthocomplemented partially ordered set usually possessing some additional structure. Physically a proposition may be interpreted as an observable that only gives the value 0 or 1 whenever it is measured.[11,17]

In the formulation of quantum logic of Jauch and Piron it is possible to get quite close to deducing the existence of a Hilbert space describing the system from the axioms for the propositional system. The propositions may then be identified with the orthomodular lattice of projectors on this Hilbert space.[11,16]

We now show that the projectors in the * algebra $\mathscr{A}_I$ form an orthomodular lattice[11,16] satisfying all but one of Piron's axioms for a propositional system,[11,16] namely the completeness property.

We shall denote the lattice of all projectors in the Hilbert space $\mathscr{H}$ by $\mathscr{P}(\mathscr{H})$. Since the projectors in $\mathscr{A}_I$ and $\bar{\mathscr{A}}_I$ all belong to $\mathscr{P}(\mathscr{H})$ they inherit a partial ordering from $\mathscr{P}(\mathscr{H})$ and we shall assume that they are ordered in this way throughout this section. Our theorem exhibits some properties of these partially ordered sets of projectors.

Theorem 4.1: The projectors in $\mathscr{A}_I$ are an irreducible atomic orthomodular sublattice[11] of $\mathbf{B}(\mathscr{H})$ in which the covering law[11] holds. This lattice is complete if and only if E is a compact spectral measure.

Proof: It is easily verified that a projector M belongs to $\mathscr{A}_I$ if and only if $M\in\mathscr{A}_L$ or $M^c\in\mathscr{A}_L$ (since $\mathscr{A}_I=\mathscr{A}_L+\mathbb{C}I$). A similar property holds for projections in $\bar{\mathscr{A}}_I=\bar{\mathscr{A}}_L+\mathbb{C}I$.

(1) Let M and N be projectors in $\mathscr{A}_I$. If $M\in\mathscr{A}_L$ then $M=E(R)ME(R)$ for some bounded Borel set R and now

$$M\wedge N\leqslant M\leqslant E(R),$$

which implies $M\wedge N=E(R)(M\wedge N)E(R)$ so $M\wedge N\in\mathscr{A}_L$. Hence if one of M,N belongs to $\mathscr{A}_L$ we have $M\wedge N\in\mathscr{A}_L$. Now suppose neither M nor N belong to $\mathscr{A}_L$; then M^c and N^c belong to $\mathscr{A}_L$ so there are bounded Borel sets R and Q in $\mathbb{R}^k$ with

$$M^c=E(R)M^cE(R)\quad\text{and}\quad N^c=E(Q)N^cE(Q).$$

This implies $M^c\leqslant E(R)$ and $N^c\leqslant E(Q)$ so

$$M^c\vee N^c=E(R)\vee E(Q)=E(R\cup Q),$$

giving

$$M^c \vee N^c = E(R\cup Q)(M^c \vee N^c)E(R\cup Q).$$

Since $R\cup Q$ is bounded we have $M^c \vee N^c \in \mathscr{A}_L$ and now by de Morgan's law[18]

$$M \wedge N = I - M^c \vee N^c \in \mathscr{A}_I.$$

We have now shown that $M \wedge N$ belongs to $\mathscr{A}_I$ whenever M and N belong to $\mathscr{A}_I$. If M and N belong to $\mathscr{A}_I$ then this property together with de Morgan's law implies that $M \vee N = I - M^c \wedge N^c$ also belongs to $\mathscr{A}_I$. Thus the projectors in $\mathscr{A}_I$ are a sublattice of $\mathscr{P}(\mathscr{H})$.

Next we show that the projector lattice of $\mathscr{A}_I$ is irreducible, i.e., 0 and I are the only projectors in $\mathscr{A}_I$ which are compatible with every projector in $\mathscr{A}_I$. Since compatibility of projectors is equivalent to commutativity[11,16] this follows from the fact that $\mathscr{A}_I' = \mathbb{C}I$ since $\mathscr{A}_I$ is irreducible by Theorem 3.1.

We now show that the projector lattice of $\mathscr{A}_I$ is atomic. Let M be a nonzero projector in $\mathscr{A}_I$. If $M \in \mathscr{A}_L$ let f be a unit vector in the range of M and let P_f be the projector onto the subspace spanned by f. Then

$$P_f < M = E(R)ME(R) < E(R)$$

for some bounded Borel set R and now P_f is clearly an atom. If $M \in \mathscr{A}_L$, then $M^c \in \mathscr{A}_L$, so $M^c = E(R)M^cE(R)$ for some bounded Borel set R. Let Q be a bounded Borel set with $R \cap Q = \varnothing$ and $E(Q) \neq 0$ [if no such Q exists then the spectrum of E must be a subset of R so $\mathscr{A}_I = \mathbf{B}(\mathscr{H})$ by Theorem 3.2 and in this case the projector lattice of $\mathscr{A}_I$ is clearly atomic]. Now [by considering a unit vector in the range of $E(Q)$] there is a one-dimensional projector N with

$$N < E(Q) < E(R)^c < M.$$

It follows that N is an atom in the projector lattice of $\mathscr{A}_I$ with $N < M$. Hence the projector lattice of $\mathscr{A}_I$ is atomic.

It is easily checked that every atom in $\mathscr{A}_I$ is a one-dimensional projector and the covering law in the projector lattice of $\mathscr{A}_I$ now follows from the covering law in $\mathscr{P}(\mathscr{H})$. The orthomodular law in the projector lattice of $\mathscr{A}_I$ follows from the orthomodular law in $\mathscr{P}(\mathscr{H})$.

Finally let us consider the completeness of the projector lattice of $\mathscr{A}_I$. If E is compact then $\mathscr{A}_I = \mathbf{B}(\mathscr{H})$ by Theorem 3.2. The projector lattice of $\mathscr{A}_I$ is simply $\mathscr{P}(\mathscr{H})$, which is well known to be complete. Next we show that the projectors in $\mathscr{A}_I$ are not a complete lattice when the spectral measure E is not compact. Suppose that the spectrum Λ of E is not compact, then, by Lemma 3.1, Λ is the union of a disjoint sequence Q_r of bounded Borel sets such that for each r, $E(Q_r) \neq 0$ and Q_r lies outside a closed ball of radius $r-1$ centered at the origin. We shall show that the set $\{E(Q_{2r-1}): r = 1,2,\ldots\}$ has no supremum in the projector lattice of $\mathscr{A}_I$. By way of contradiction suppose that $M \in \mathscr{A}_I$ is a supremum for this set. Note that $E(Q_{2r})^c$ is an upper bound for the above set for every r. Now

$$\bigwedge_{r=1}^{\infty} E(Q_{2r})^c = \left(\bigvee_{r=1}^{\infty} E(Q_{2r})\right)^c = I - \bigvee_{r=1}^{\infty} E(Q_{2r})$$
$$\Rightarrow M < E(Q_1 \cup Q_2 \cup \cdots)$$

since $M < E(Q_{2r})^c$. But M is an upper bound for $\{E(Q_{2r-1}): r = 1,2,\ldots\}$ in $\mathscr{P}(\mathscr{H})$ so $M \geq E(Q_1 \cup Q_3 \cup \cdots)$. Hence $M = E(Q_1 \cup Q_3 \cup \cdots)$, so $E(Q_1 \cup Q_3 \cup \cdots)$ belongs to $\mathscr{A}_I$. Now we must have $E(Q_1 \cup Q_3 \cup \cdots) \in \mathscr{A}_L$ or $E(Q_2 \cup Q_4 \cup \cdots) = E(Q_1 \cup Q_3 \cup \cdots)^c \in \mathscr{A}_L$. Suppose $E(Q_1 \cup Q_3 \cup \cdots) \in \mathscr{A}_L$, then

$$\begin{aligned} E(Q_1 \cup Q_3 \cup \cdots) &= E(R)E(Q_1 \cup Q_3 \cup \cdots)E(R) \\ &= E(R \cap (Q_1 \cup Q_3 \cup \cdots)) \\ &= E\left(\bigcup_{r=1}^{\infty} R \cap Q_{2r-1}\right), \end{aligned}$$

for some bounded Borel set R. Now since R is bounded and each Q_r lies outside a ball of radius $r-1$ we have $R \cap Q_{2j-1} = \varnothing$ for some j. Since $E(Q_{2j-1}) \neq 0$ we can choose a nonzero vector f in the range of $E(Q_{2j-1})$, i.e., we have $f \neq 0$ satisfying

$$\begin{aligned} f &= E(Q_{2j-1})f \\ &= E(Q_1 \cup Q_3 \cup \cdots)f \\ &= E\left(\bigcup_{r=1}^{\infty} (R \cap Q_{2r-1})\right)f. \end{aligned}$$

Since the sequence Q_r is disjoint this implies

$$f = E(R \cap Q_{2j-1})f = 0,$$

which is a contradiction since f was chosen to be nonzero. If $E(Q_2 \cup Q_4 \cup \cdots) \in \mathscr{A}_L$, then we arrive at a contradiction in a similar way. Hence the assumption that the set $\{E(Q_{2r-1}): r = 1,2,\ldots\}$ has a supremum in the projection lattice of $\mathscr{A}_L$ leads to a contradiction so this lattice is not complete.

Before moving on, we would point out that if the spectral measure E is not compact then the projectors in $\mathscr{A}_L$ do not constitute an orthomodular lattice. There is no maximum element in this set of projectors.

Lemma 4.1: Let $A \in \mathbf{B}(\mathscr{H})$, then these are equivalent: (1) $A \in \bar{\mathscr{A}}_L$; (2) there is a sequence R_r of bounded Borel sets in $\mathbb{R}^k$ such that $E(R_r)AE(R_r)$ converges to A in the operator norm; (3) for every increasing sequence Q_r of open balls converging to $\mathbb{R}^k$ the sequence $E(Q_r)AE(Q_r)$ converges to A in the operator norm.

Proof: [(1)$\Rightarrow$(2).] Suppose $A \in \bar{\mathscr{A}}_L$. Then there is a sequence A_r in $\mathscr{A}_L$ converging in the operator norm to A and for each r, $A_r = E(R_r)A_rE(R_r)$ for some bounded Borel set R_r. Now

$$\begin{aligned} &\|E(R_r)AE(R_r) - A\| \\ &\quad \leq \|E(R_r)AE(R_r) - E(R_r)A_rE(R_r)\| + \|A_r - A\| \\ &\quad \leq 2\|A_r - A\| \\ &\quad \to 0 \quad (r \to \infty), \end{aligned}$$

so (2) holds.

[(2)$\Rightarrow$(3).] Assume (2) holds and define $V_k = \cup_{r=1}^{k} R_r$. Then V_k is an increasing sequence of bounded Borel sets and $R_k < V_k$ for each integer k. Let Q_r be an increasing sequence of open balls converging to $\mathbb{R}^n$. Then we can extract a strictly increasing subsequence Q_{r_k} such that

$$V_k < Q_{r_k}, \quad \text{for every } k = 1,2,3,\ldots.$$

Now $R_k < Q_{r_k}$ so $E(Q_{r_k})E(R_k) = E(R_k)$ giving

$$\begin{aligned} &\|E(Q_{r_k})AE(Q_{r_k}) - A\| \\ &\quad \leq \|E(Q_{r_k})AE(Q_{r_k}) \\ &\qquad - E(Q_{r_k})E(R_k)AE(R_k)E(Q_{r_k})\| \\ &\qquad + \|E(R_k)AE(R_k) - A\| \\ &\quad \leq 2\|E(R_k)AE(R_k) - A\|. \end{aligned}$$

Hence a subsequence of $E(Q_r)AE(Q_r)$ converges to A. To show that the sequence itself converges to A observe that for $r_k \leqslant r$ similar inequalities to the above give

$$\|E(Q_r)AE(Q_r) - A\| \leqslant 2\|E(Q_{r_k})AE(Q_{r_k}) - A\| \leqslant 4\|E(R_k)AE(R_k) - A\|$$

and this implies $E(Q_r)AE(Q_r)$ converges to A.

Finally (3)$\Rightarrow$(1) is obvious from the definition of $\bar{\mathscr{A}}_L$.

Theorem 4.2: If E is not compact spectral measure then projectors in $\bar{\mathscr{A}}_I$ are not a complete lattice.

Proof: We now show that the projectors in $\bar{\mathscr{A}}_I$ do not form a complete lattice. As in the proof of Theorem 4.1 we can find a sequence Q_r of disjoint Borel sets in $\mathbb{R}^k$ such that for each r, $E(Q_r) \neq 0$, Q_r lies outside a ball of radius $r-1$ centered at the origin and the set $\{E(Q_{2r-1}): r = 1,2,\ldots\}$ has no supremum in $\mathscr{A}_I$. By way of contradiction suppose that $M \in \bar{\mathscr{A}}_I$ is a supremum for this set, then as in the proof of Theorem 4.1 we must have $M = E(Q_1 \cup Q_3 \cup \cdots)$. It follows that $E(Q_1 \cup Q_3 \cup \cdots) \in \bar{\mathscr{A}}_L$ or $E(Q_2 \cup Q_4 \cup \cdots) \in \bar{\mathscr{A}}_L$. If $E(Q_1 \cup Q_3 \cup \cdots) \in \bar{\mathscr{A}}_L$, then by Lemma 4.1 there is a sequence R_r of bounded Borel sets converging to $\mathbb{R}^k$ such that

$$E(R_r)E(Q_1 \cup Q_3 \cup \cdots)E(R_r) \to E(Q_1 \cup Q_3 \cup \cdots)$$

in the operator norm. Now

$$E(R_r)E(Q_1 \cup Q_3 \cup \cdots)E(R_r) = E(R_r \cap (Q_1 \cup Q_3 \cup \cdots)).$$

It follows from the properties of Q_r that this sequence is not eventually constant and so does not converge in the operator norm giving a contradiction. If $E(Q_2 \cup Q_4 \cup \cdots)$ belongs to $\bar{\mathscr{A}}_L$ then we arrive at a contradiction in a similar way. Hence the projectors in $\bar{\mathscr{A}}_I$ are not a complete lattice.

V. CONCLUDING REMARKS

The completeness property usually built into a propositional system is seen to be absent in the projector lattice in $\mathscr{A}_I$. Furthermore the projectors in $\mathscr{A} = \mathscr{A}_0^s + L^\infty(p)$ do not form a complete lattice either (see the Appendix). Here $\mathscr{A}$ is the C^* algebra associated with a free quantum particle in the asymptotically separable theory of quantum mechanics recently put forward by the authors.[4] It would therefore be of interest to investigate further the properties of incomplete lattices and to search for possible physical implications.

APPENDIX: PROJECTOR LATTICE OF $\mathscr{A}_0^s$ AND $L^\infty(p)$

We shall adhere to the notation in our two previous papers[4] here.

Theorem A: The projectors in $\mathscr{A} = \mathscr{A}_0^s + L^\infty(p)$ do not form a complete lattice.

Proof: It follows from the proof of Theorem 1 in the first paper of Ref. 4 that there is an increasing b_r of balls in $\mathbb{R}^n$ with the property that if we define

$$\Delta_1 = b_1,$$

$$\Delta_r = b_r - b_{r-1} \quad (r = 2,3,\ldots),$$

and

$$\Delta = \bigcup_{r=1}^{\infty} \Delta_{2r},$$

then $s^* - \lim U_t^{o\dagger} E_x(\Delta) U_t^o$ does not exist. So, $E_x(\Delta)$ does not belong to $\mathscr{A}^s$ and hence does not belong to $\mathscr{A}$ [this follows from Theorem 8, part (3), of the first paper in Ref. 4]. Note that $E_x(\Delta_r)$ belongs to $\mathscr{A}_L$ and hence belongs to $\mathscr{A}^s$.

We shall show that the set $\{E_x(\Delta_{2r}): r = 1,2,\ldots\}$ has no supremum in the ordered set of projectors in $\mathscr{A}$. By way of contradiction suppose P is a supremum of this set, then for each r we have $E_x(\Delta_{2r}) \leqslant P$, and taking a supremum in $\mathscr{P}(\mathscr{H})$ gives

$$E_x(\Delta) = \bigvee_{r=1}^{\infty} E_x(\Delta_{2r}) \leqslant P.$$

Also for all i and j we have $E_x(\Delta_{2j}) \leqslant E_x(\Delta_{2j-1})^c$, the orthogonal complement of $E_x(\Delta_{2i-1})$. So for all r, $P \leqslant E_x(\Delta_{2j-1})^c$, and taking an infimum in $\mathscr{P}(\mathscr{H})$ gives

$$P \leqslant \bigwedge_{r=1}^{\infty} E_x(\Delta_{2r-1})^c = \left(\bigvee_{r=1}^{\infty} E_x(\Delta_{2r-1}) \right)^c = E_x(\Delta).$$

Hence $P = E_x(\Delta)$ and if P were in the ordered set of projectors in $\mathscr{A}$ we would have that $E_x(\Delta)$ should belong to $\mathscr{A}$ giving the required contradiction.

ACKNOWLEDGMENTS

The authors express their thanks to Professor V. Enss for a very useful communication and for bringing to their notice the work of Dollard.[3]

R. G. D. McLean wishes to acknowledge a postdoctoral fellowship from the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, and to thank R. Giles and H. Kummer for their hospitality.

[1] R. Haag and D. Kastler, J. Math. Phys. **5**, 848 (1964).
[2] K. K. Wan and D. McLean, Phys. Lett. A **94**, 198 (1983), **95**, 76 (1983); R. S. Strichartz, J. Functional Anal. **40**, 341 (1981); V. Enss, Commun. Math. Phys. **89**, 245 (1983).
[3] J. D. Dollard, Commun. Math. Phys. **12**, 193 (1969).
[4] K. K. Wan and D. McLean, J. Phys. A **17**, 825 (1984), **17**, 2363 (1984); K. K. Wan and D. McLean, J. Phys. **17**, 837 (1984), **17**, 2363 (1984); K. K. Wan and D. J. McLean, Phys. A Lett. A **102**, 163 (1984).
[5] K. K. Wan, T. D. Jackson, and I. H. McKenna, Nuovo Cimento B **81**, 165 (1984).
[6] K. K. Wan and T. D. Jackson, Phys. Lett. A **106**, 219 (1984); P. Pfeifer, Helv. Phys. Acta **53**, 410 (1980).
[7] E. Prugovečki, *Quantum Mechanics in Hilbert Space* (Academic, New York, 1981), pp. 231–236, 240.
[8] W. M. de Muynck, Found. Phys. **14**, 199 (1984).
[9] P. Roman, *Some Modern Mathematics for Physicists and Other Outsiders* (Pergamon, New York, 1975), Vol. 2, p. 650.
[10] O. Bratteli and D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics* (Springer, New York, 1979), Vol. I, pp. 47 and 71.
[11] J. M. Jauch *Foundations of Quantum Mechanics* (Addison–Wesley, Reading, MA 1968), Sec. 5; C. Piron, *Foundations of Quantum Mechanics* (Benjamin, London, 1976).
[12] G. Lüdders, Ann. Phys. **8**, 322 (1951); J. Bub, *Proceedings of the International School of Physics "Enrico Fermi" Course LXXII on Problems in the Foundations of Physics* (North-Holland, Amsterdam, 1979).
[13] S. K. Berberian, *Lectures in Functional Analysis and Operator Theory* (Springer, New York, 1974), pp. 66 and 90.
[14] J. Weidmann, *Linear Operators in Hilbert Spaces* (Springer, New York, 1980), p. 80.
[15] T. F. Jordon, *Linear Operators for Quantum Mechanics* (Wiley, New York 1969), Sec. 4.
[16] P. R. Halmos, *Introduction to Hilbert Space* (Chelsea, New York, 1957), p. 62.
[17] R. J. Greechie and S. P. Gudder, *Contemporary Research in the Foundations of Quantum Theory*, edited by C. A. Hooker (Reidel, Dordrecht, 1975).
[18] T. S. Blyth, M. F. Janowitz, *Residuation Theory* (Pergamon, Oxford, 1972), p. 162.

On the connection between Schrödinger and Dirichlet forms

S. Albeverio
ZiF, Universität Bielefeld, D-4800 Bielefeld 1, Federal Republic of Germany and Mathematisches Institut, Ruhr-Universität, D-4630 Bochum 1, Federal Republic of Germany

F. Gesztesy[a] and W. Karwowski[b]
ZiF, Universität Bielefeld, D-4800 Bielefeld 1, Federal Republic of Germany

L. Streit
ZiF and Fakultät für Physik, Universität Bielefeld, D-4800 Bielefeld 1, Federal Republic of Germany

(Received 3 October 1984; accepted for publication 8 March 1985)

Relations between Schrödinger forms associated with Schrödinger operators in $L^2(\Omega;d^n x)$, $\Omega \subset \mathbb{R}^n$ open, $n \geqslant 1$ and the corresponding Dirichlet forms are investigated. Various concrete examples are presented.

I. INTRODUCTION

In this paper we consider the relation between Schrödinger forms Q associated with Schrödinger operators $H = -\Delta + V$ in $L^2(\Omega;d^n x)$, $\Omega \subset \mathbb{R}^n$ open, $n \in \mathbb{N}$ and the corresponding Dirichlet (energy) form E_ϕ in $L^2(\Omega;\phi^2 d^n x)$, where $0 < \phi \in L^2_{\text{loc}}(\Omega;d^n x)$ represents an appropriate distributional solution of $H\phi = E\phi$ for some $E \in \mathbb{R}$. In contrast to earlier treatments on this subject where some emphasis has been devoted to the construction of strongly singular interactions H in $L^2(\Omega;d^n x)$ by means of introducing first E_ϕ and then obtaining Q and H by unitary equivalence between $L^2(\Omega;\phi^2 d^n x)$ and $L^2(\Omega;d^n x)$ (cf. Refs. 1–6) we now intend to study the opposite direction: Given a Schrödinger operator H in $L^2(\Omega;d^n x)$ with Q its associated form we show how to use appropriate distributional solutions ϕ of $H\phi = E\phi$ for some $E \in \mathbb{R}$ to construct the corresponding Dirichlet form E_ϕ with associated Hamiltonian H_ϕ in $L^2(\Omega;\phi^2 d^n x)$. In particular we are concerned with a certain kind of uniqueness question regarding the associated Schrödinger and Dirichlet forms by establishing that $C_0^\infty(\Omega)$ as well as $\phi C_0^\infty(\Omega)$ are cores for Q and $C_0^\infty(\Omega)$ is a core for E_ϕ implying (cf. Theorem 2.5)

$$E_\phi(f,g) = Q(\phi f,\phi g) - E(\phi f,\phi g)$$
$$\text{for all } f,g \in \mathscr{D}(E_\phi) \subset L^2(\Omega;\phi^2 d^n x) \tag{1.1}$$

[where $(\cdot,\cdot)$ denotes the scalar product in $L^2(\Omega;d^n x)$ and $\mathscr{D}(\cdot)$ the domain]. We note that in the case $\Omega = \mathbb{R}^n$ and under certain conditions on ϕ (see Lemma 2.1) strong uniqueness results concerning operator cores for H and H_ϕ (instead of form cores) have been obtained in Refs. 1 and 7 (cf. also Refs. 8). Here we follow the approach in Ref. 9 by studying directly Q and E_ϕ. In particular, taking $\Omega = \mathbb{R}^n$ and ϕ to be the ground state of H, it was shown[9] that $C_0^\infty(\mathbb{R}^n)$ as well as $\phi C_0^\infty(\mathbb{R}^n)$ are form cores for H and that Eq. (1.1) holds in this case. Our results essentially represent a "localized" version of this approach in the sense that first $\mathbb{R}^n$ is replaced by an open set $\Omega \subset \mathbb{R}^n$ and second ϕ is not assumed to lie in the form domain of H. Finally, in Sec. III, we use the detailed results of Refs. 10–13 concerning distributional solutions of $H\phi = E\phi$ to discuss explicit realizations of our general framework described in Sec. II. For a recent treatment of ultracontractivity (ultraboundedness) of Schrödinger semigroups, where also core properties of Dirichlet forms have been investigated see Ref. 14.

II. RELATIONS BETWEEN SCHRÖDINGER AND DIRICHLET FORMS

In order to describe the results of Refs. 1, 7, and 9 concerning the relation between Schrödinger forms and the corresponding Dirichlet forms, mentioned in the Introduction, we first introduce some notations.

Let $0 \leqslant \phi \in L^2_{\text{loc}}(\mathbb{R}^n)$ and define the minimal energy form $\dot{E}_\phi$ in $L^2(\mathbb{R}^n;\phi^2 d^n x)$ by

$$\dot{E}_\phi(f,g) = \int_{\mathbb{R}^n} \phi^2(x) d^n x\, \overline{(\nabla f)(x)}(\nabla g)(x)$$
$$\text{on } \mathscr{D}(\dot{E}_\phi) = C_0^\infty(\mathbb{R}^n). \tag{2.1}$$

[Abbreviations of the type $L^p_{(\text{loc})}(\Omega)$ for some $\Omega \subset \mathbb{R}^n$, $p \geqslant 1$ always refer to $L^p_{(\text{loc})}(\Omega;d^n x)$, where $d^n x$ denotes the Lebesgue measure.] If $\dot{E}_\phi$ is closable we denote its closure by E_ϕ and the unique self-adjoint, non-negative operator in $L^2(\mathbb{R}^n;\phi^2 d^n x)$ associated with E_ϕ is denoted by H_ϕ. As a sufficient condition for closability of $\dot{E}_\phi$ we mention, for example,[1] $(\nabla\phi)_j \in L^2_{\text{loc}}(\mathbb{R}^n)$, $1 \leqslant j \leqslant n$ (see Ref. 15) (see also Refs. 5, 7, and 16–18 for recent results in this direction).

Next assume $V = V_1 + V_2$ with $c \leqslant V_1 \in L^1_{\text{loc}}(\mathbb{R}^n)$ for some $c \in \mathbb{R}$ and V_2 to be form-bounded with respect to $H_0 = -\Delta$, $\mathscr{D}(H_0) = H^{2,2}(\mathbb{R}^n)$ (see Ref. 19) with bound strictly less than 1. The minimal Schrödinger form $\dot{Q}$ in $L^2(\mathbb{R}^n)$ is then defined by

$$\dot{Q}(f,g) = \int_{\mathbb{R}^n} d^n x\, \overline{(\nabla f)(x)}(\nabla g)(x)$$
$$+ \int_{\mathbb{R}^n} d^n x\, V(x)\overline{f(x)}g(x), \quad \mathscr{D}(\dot{Q}) = C_0^\infty(\mathbb{R}^n). \tag{2.2}$$

By the hypotheses on V, $\dot{Q}$ is closable and its closure Q has the domain $\mathscr{D}(Q) = H^{2,1}(\mathbb{R}^n) \cap \mathscr{D}(|V_1|^{1/2})$. By $H = -\Delta + V$ we

[a] Alexander von Humboldt Research Fellow. On leave of absence from Institut für Theoretische Physik, Universität Graz, Graz, Austria.
[b] On leave of absence from Institute for Theoretical Physics, University of Wrocław, Wrocław, Poland.

 0022-2488/85/102546-08$02.50

denote the unique self-adjoint and semibounded operator in $L^2(\mathbb{R}^n)$ corresponding to Q.

Now we are in position to describe the results contained in Refs. 1, 7, and 9.

Lemma 2.1 [1,7]: (a) Let $0 \leqslant \phi \in L^2_{\rm loc}(\mathbb{R}^n)$, $\phi^{-1} \in L^\infty_{\rm loc}(\mathbb{R}^n)$, $\Delta\phi \in L^2_{\rm loc}(\mathbb{R}^n)$ (see Ref. 15), and define $V = \phi^{-1}(\Delta\phi)$. If $\dot{H} = -\Delta + V$, $\mathscr{D}(\dot{H}) = C_0^\infty(\mathbb{R}^n)$, is essentially self-adjoint in $L^2(\mathbb{R}^n)$, then $\dot{E}_\phi$ and $\dot{Q}$ are closable and H_ϕ is essentially self-adjoint on $\phi^{-1}C_0^\infty(\mathbb{R}^n)$.

Moreover

$$\mathscr{D}(H_\phi) = \phi^{-1}\mathscr{D}(H), \quad H_\phi = \phi^{-1}H\phi. \tag{2.3}$$

In particular, $\phi^{-1}C_0^\infty(\mathbb{R}^n)$ is a form core for E_ϕ and

$$\mathscr{D}(E_\phi) = \phi^{-1}\mathscr{D}(Q),$$
$$E_\phi(\phi^{-1}f, \phi^{-1}g) = Q(f,g),$$
$$\text{for all } f,g \in \mathscr{D}(Q). \tag{2.4}$$

(b) Assume $\phi \in L^2_{\rm loc}(\mathbb{R}^n)$, $\phi > 0$ on compact sets, and ϕ is locally Lipshitz continuous. Then H_ϕ is essentially self-adjoint on $C_0^\infty(\mathbb{R}^n)$.

Remark 2.2: For the sake of brevity we have chosen rather strong conditions on ϕ in part (a) of Lemma 2.1. For considerably weaker assumptions cf. Ref. 1. Sufficient conditions on $V = \phi^{-1}(\Delta\phi)$ in order to guarantee essential self-adjointness of $\dot{H}$ can be found in Ref. 20.

It remains to state the following lemma.

Lemma 2.3 [9]: Let $V = V_1 + V_2$ with $c \leqslant V_1 \in L^1_{\rm loc}(\mathbb{R}^n)$ for some $c \in \mathbb{R}$ and $V_2 \in L^p(\mathbb{R}^n)$ real with $p \geqslant n/2$ if $n \geqslant 3$, $p > 1$ if $n = 2$ and $p = 1$ if $n = 1$, then $\dot{Q}$ is closable. Furthermore assume that $E = \inf \sigma(H)$ is an eigenvalue (necessarily simple) with ϕ the corresponding eigenvector (ground state). Then $\phi \in L^\infty(\mathbb{R}^n)$, $\phi^{-1} \in L^\infty_{\rm loc}(\mathbb{R}^n)$, and $\phi C_0^\infty(\mathbb{R}^n)$ is a core for Q. Moreover ϕ can be chosen to be >0 a.e. and

$$Q(\phi f, \phi g) - E(\phi f, \phi g) = E_\phi(f,g)$$
$$\text{for all } f,g \in \mathscr{D}(E_\phi). \tag{2.5}$$

□

For the rest of this section we aim at a twofold generalization of the above result. First, in analogy to Lemma 2.1, we want to weaken the global hypothesis $\phi \in \mathscr{D}(H)$ (implicit in ϕ being an eigenvector) and second we are interested in replacing $\mathbb{R}^n$ by an open set $\Omega \subset \mathbb{R}^n$. We mainly follow the ideas of Ref. 9 "localizing" the corresponding proofs step by step.

The following assumptions are made.

(I) Let $\Omega \subset \mathbb{R}^n$ be open (nonempty), $\Sigma \subset \Omega$ be a closed set of Lebesgue-measure zero, and $V = V_1 + V_2$, where $c \leqslant V_1 \in L^1_{\rm loc}(\Omega - \Sigma)$ for some $c \in \mathbb{R}$ and $V_2 \in L^1_{\rm loc}(\Omega - \Sigma)$ is real valued and form bounded with respect to the Dirichlet Laplacian $-\Delta_{\rm D}$ defined below with bound strictly less than 1.

Here the minimal Dirichlet form $\dot{Q}_{\rm D}$ in $L^2(\Omega)$ is defined by

$$\dot{Q}_{\rm D}(f,g) = (\nabla f, \nabla g), \quad \mathscr{D}(\dot{Q}_{\rm D}) = C_0^\infty(\Omega - \Sigma), \tag{2.6}$$

with $Q_{\rm D}$ its closure and $-\Delta_{\rm D}$ the associated Hamiltonian.

We also define the closed Schrödinger form

$$Q(f,g) = Q_{\rm D}(f,g) + (|V|^{1/2}f, \operatorname{sgn}(V)|V|^{1/2}g),$$
$$\mathscr{D}(Q) = \mathscr{D}(Q_{\rm D}) \cap \mathscr{D}(|V_1|^{1/2}), \tag{2.7}$$

and denote by $H = -\Delta_{\rm D} + V$ the corresponding self-adjoint operator in $L^2(\Omega)$.

(II) There exists an $E \in \mathbb{R}$ and $\phi \in L^\infty_{\rm loc}(\Omega - \Sigma)$ real valued such that

$$\phi^{-1} \in L^\infty_{\rm loc}(\Omega - \Sigma), \quad (\nabla\phi)_j \in L^2_{\rm loc}(\Omega - \Sigma), \quad 1 \leqslant j \leqslant n,$$
$$\text{and } -\Delta\phi + V\phi = E\phi,$$

in the sense of distributions in $C_0^\infty(\Omega - \Sigma)'$.

Given the above assumptions we state the following lemma.

Lemma 2.4: Assume hypotheses (I) and let $\phi \in L^\infty_{\rm loc}(\Omega - \Sigma)$ be such that $\phi^{-1} \in L^\infty_{\rm loc}(\Omega - \Sigma)$ and $(\nabla\phi)_j \in L^2_{\rm loc}(\Omega - \Sigma)$, $1 \leqslant j \leqslant n$. Then $C_0^\infty(\Omega - \Sigma)$ as well as $\phi C_0^\infty(\Omega - \Sigma)$ are form cores for $H = -\Delta_{\rm D} + V$.

Proof: We first note that by Theorem 2 in Ref. 21, $C_0^\infty(\Omega - \Sigma)$ is a form core for H. Thus it suffices to prove that for each $f \in C_0^\infty(\Omega - \Sigma)$ there is a sequence $\{f_k\}_{k \in \mathbb{N}} \subset C_0^\infty(\Omega - \Sigma)$ such that [$\|\cdot\|_p$ denotes the norm in $L^p(\Omega)$ for $1 \leqslant p \leqslant \infty$]

$$\| |H|^{1/2}(f - \phi f_k)\|_2^2 + \|f - \phi f_k\|_2^2 \to 0, \quad \text{as } k \to \infty.$$

Let $f \in C_0^\infty(\Omega - \Sigma)$, then

$$\int_\Omega d^n x |V_1(x)| \, |\phi^{-1}(x) f(x)|^2$$
$$\leqslant \|\phi^{-2}f^2\|_\infty \int_{\operatorname{supp}(f)} d^n x |V_1(x)| < \infty,$$

and similarly ($|\mathscr{M}|$ denotes the Lebesgue measure of some measurable set $\mathscr{M} \subset \mathbb{R}^n$)

$$\int_\Omega d^n x |\nabla(\phi^{-1}f)(x)|^2$$
$$\leqslant 2 \int_{\operatorname{supp}(f)} d^n x |\phi(x)|^{-2} |(\nabla f)(x)|^2$$
$$+ 2 \int_{\operatorname{supp}(f)} d^n x |\phi(x)|^{-4} |f(x)|^2 |(\nabla\phi)(x)|^2$$
$$\leqslant 2\|\phi^{-2}|\nabla f|^2\|_\infty |\operatorname{supp}(f)|$$
$$+ 2\|\phi^{-2}f\|_\infty^2 \int_{\operatorname{supp}(f)} d^n x |(\nabla\phi)(x)|^2,$$

and this is finite by the assumptions on ϕ. Next we introduce an approximate identity j_k,

$$j_k(x) = k^n j(kx), \quad x \in \mathbb{R}^n, \quad k \in \mathbb{N},$$

where

$$0 \leqslant j \in C_0^\infty(\mathbb{R}^n), \quad \operatorname{supp}(j) \subset \{x \in \mathbb{R}^n \mid |x| \leqslant 1\},$$
$$\int_{\mathbb{R}^n} d^n x \, j(x) = 1,$$

and introduce

$$f_k = j_k * (\phi^{-1}f), \quad f \in C_0^\infty(\Omega - \Sigma).$$

Since

$$|f_k(x)| \leqslant \|\phi^{-1}f\|_\infty < \infty, \tag{2.8}$$

$\{f_k\}_{k \in \mathbb{N}}$ is uniformly bounded, and from the fact that supp $(\phi^{-1}f)$ is compact we infer that

$$\|f_k - \phi^{-1}f\|_2 \underset{k \to \infty}{\to} 0, \quad \|\nabla f_k - \nabla(\phi^{-1}f)\|_2 \to 0, \tag{2.9}$$
$$\text{as } k \to \infty.$$

By passing to a subsequence (again denoted by f_k) we also get

$$f_k \underset{k\to\infty}{\to} \phi^{-1}f \quad \text{pointwise a.e.} \tag{2.10}$$

Finally, if $k^{-1} < k_0^{-1} = \text{dist}\ (\text{supp}\ (\phi^{-1}f),\ \partial(\Omega-\Sigma))$ [$\partial\mathscr{M} = \overline{\mathscr{M}} \cap (\overline{\mathbb{R}^n - \mathscr{M}})$ denotes the boundary of a subset $\mathscr{M} \subset \mathbb{R}^n$], we obtain[22]

$$f_k \in C_0^\infty(\Omega-\Sigma), \quad k > k_0. \tag{2.11}$$

Thus

$$\phi^{-1}C_0^\infty(\Omega-\Sigma) \subset \mathscr{D}(Q) = \mathscr{D}(Q_D) \cap \mathscr{D}(|V_1|^{1/2}).$$

Consequently, $V \in L^1_{\text{loc}}(\Omega-\Sigma)$, Eqs. (2.8)–(2.11) and dominated convergence imply ($k \geqslant k_0$)

$$\begin{aligned}
&Q(f-\phi f_k f-\phi f_k) + \|f-\phi f_k\|^2 \\
&\quad = \int_\Omega d^n x |\nabla(f-\phi f_k)(x)|^2 \\
&\qquad + \int_\Omega d^n x [V(x)+1] |(f-\phi f_k)(x)|^2 \\
&\quad \leqslant \int_\Omega d^n x |\phi(x)|^2 |\nabla(\phi^{-1}f - f_k)(x)|^2 \\
&\qquad + \int_\Omega d^n x |\phi(x)|^2 [|V(x)|+1] |(\phi^{-1}f - f_k)(x)|^2 \\
&\quad \leqslant \Big[\operatorname*{ess\,sup}_{y\in \text{supp}(f_{k_0})} |\phi(y)|^2\Big] \Big\{ \int_{\text{supp}(f_{k_0})} d^n x |\nabla(\phi^{-1}f - f_k)(x)|^2 \\
&\qquad + \int_{\text{supp}(f_{k_0})} d^n x [|V(x)|+1] |(\phi^{-1}f - f_k)(x)|^2 \Big\} \underset{k\to\infty}{\to} 0.
\end{aligned}$$

□

In order to state our main result we assume that $\phi \in L^2_{\text{loc}}(\Omega-\Sigma)$ and define the minimal Dirichlet (energy) form $\dot{E}_\phi$ in $L^2(\Omega;\phi^2\, d^n x)$ by

$$\begin{aligned}
&\dot{E}_\phi(f,g) = \int_\Omega \phi^2(x) d^n x\, \overline{(\nabla f)(x)}(\nabla g)(x) \\
&\qquad \text{on } \mathscr{D}(\dot{E}_\phi) = C_0^\infty(\Omega-\Sigma).
\end{aligned} \tag{2.12}$$

In case $\dot{E}_\phi$ is closable its closure is again denoted by E_ϕ with H_ϕ, the uniquely associated self-adjoint, non-negative operator in $L^2(\Omega;\phi^2\, d^n x)$.

Theorem 2.5: Assume conditions (I) and (II). Then $\dot{E}_\phi$ is closable, $\phi C_0^\infty(\Omega-\Sigma)$ is a form core for Q, and

$$E_\phi(f,g) = Q(\phi f,\phi g) - E(\phi f,\phi g), \quad \text{for all } f,g \in \mathscr{D}(E_\phi). \tag{2.13}$$

Proof: By Lemma 2.4 it suffices to prove Eq. (2.13) for $f,g \in C_0^\infty(\Omega-\Sigma)$. But the fact that $(\Delta\phi) = V\phi$ a.e., $\phi \in L^\infty_{\text{loc}}(\Omega-\Sigma)$ and $V\phi \in L^1_{\text{loc}}(\Omega-\Sigma)$ immediately shows

$$\begin{aligned}
&Q(\phi f,\phi g) - E(\phi f,\phi g) \\
&\quad = \int_\Omega d^n x\, \overline{(\nabla\phi f)(x)}(\nabla\phi g)(x) \\
&\qquad + \int_\Omega d^n x [V(x) - E] |\phi(x)|^2 \overline{f(x)} g(x) \\
&\quad = \int_\Omega d^n x\, \overline{\phi(x)} \{(-\Delta\phi)(x) \\
&\qquad + [V(x) - E]\phi(x)\}\, \overline{f(x)} g(x) \\
&\qquad + \int_\Omega |\phi(x)|^2 d^n x\, \overline{(\nabla f)(x)}(\nabla g)(x) \\
&\quad = E_\phi(f,g), \quad f,g \in C_0^\infty(\Omega-\Sigma).
\end{aligned}$$

□

The form E_ϕ by definition is the closure of the minimal Dirichlet form $\dot{E}_\phi$ defined on $C_0^\infty(\Omega-\Sigma)$ [cf. Eq. (2.12)]. On the other hand the main object studied in Refs. 1–6 is the minimal Dirichlet form

$$\begin{aligned}
&\dot{\tilde{E}}_\phi(f,g) = \int_\Omega \phi^2(x) d^n x (\nabla f)(x)(\nabla g)(x), \\
&\qquad \text{on } \mathscr{D}(\dot{\tilde{E}}_\phi) = C_0^\infty(\Omega).
\end{aligned} \tag{2.14}$$

If, e.g., $\phi \in L^2_{\text{loc}}(\Omega)$, $\nabla\phi \in L^2_{\text{loc}}(\Omega-\Sigma)$ (which will be assumed in the following), the corresponding closure of $\dot{\tilde{E}}_\phi$ will be denoted by $\tilde{E}_\phi$. Obviously we have

$$E_\phi \subset \tilde{E}_\phi, \tag{2.15}$$

but it is natural to ask under which additional conditions on Σ one actually has equality in (2.15). The rest of this section is entirely devoted to give a complete answer to this question using the concept of capacity.

For convenience of the reader we briefly summarize some basic facts concerning the notion of capacity (for more details see Refs. 1,4,16, and 23). Let $\mathscr{O}$ denote the family of open measurable subsets of Ω and define for all $A \in \mathscr{O}$,

$$\mathscr{L}_A = \{f \in \mathscr{D}(\tilde{E}_\phi) | f \geqslant 1 \quad m\text{-a.e. on } A\}, \tag{2.16}$$

where m is the measure $dm = \phi^2(x) d^n x$, i.e.,

$$m(A) = \int_A \phi^2(x) d^n x. \tag{2.17}$$

We assume assumption (II) so that $\phi^2 > 0\ dx$-a.e. on $\Omega-\Sigma$ and, since Σ has Lebesgue measure zero, also $\phi^2 > 0\ dx$-a.e., which implies $m(A) > 0$ for $A \in \mathscr{O}$. Assuming $\phi \in L^2_{\text{loc}}(\Omega)$, from now on we have $m(A) < \infty$ for all bounded, measurable sets A so that we can define the capacity of sets according to[16]

$$\text{Cap}(A) = \begin{cases} \inf\limits_{f\in\mathscr{L}_A} \{\tilde{E}_\phi(f,f) + \|\phi f\|_2^2\}, & \text{if } \mathscr{L}_A \neq \emptyset, \\ \infty, & \text{if } \mathscr{L}_A = \emptyset. \end{cases} \tag{2.18}$$

For an arbitrary set $B \subset \Omega$ we define

$$\text{Cap}(B) = \inf_{\substack{A\in\mathscr{O} \\ B\subset A}} \text{Cap}(A). \tag{2.19}$$

By Lemma 3.1.1 of Ref. 16 we obtain for each $A \in \mathscr{O}$ the existence of a unique element $e_A \in \mathscr{L}_A$, the so-called equilibrium potential of A, such that

$$\begin{aligned}
&\tilde{E}_\phi(e_A,e_A) + \|\phi e_A\|_2^2 = \text{Cap}(A), \\
&\qquad 0 \leqslant e_A \leqslant 1 \quad \text{and} \quad e_A = 1 \quad m\text{-a.e. on } A.
\end{aligned} \tag{2.20}$$

One has always $m(A) \leqslant \text{Cap}(A)$ but in general $m(A) = 0$ does not imply $\text{Cap}(A) = 0$. Whether $m(A) = 0$ indeed implies $\text{Cap}(A) = 0$ strongly depends on the behavior of ϕ near the set A and on the dimension n. The following facts which can be derived from Sec. 4 of Ref. 4 give a partial answer to this question.

Lemma 2.6[4]: Let $n = 1$, $(a,b) \subset \mathbb{R}$, $b > a$, and $x_0 \in (a,b)$.

(a) If there exists a $C > 0$ such that

$$|\phi(x)| \leqslant C|x - x_0|^{1/2}, \quad \text{for all } x \in (a,b),$$

then $\text{Cap}(\{x_0\}) = 0$.

(b) If there exist positive constants C,γ with $0 < \gamma < \frac{1}{2}$ such that

$$|\phi(x)| \geqslant C|x-x_0|^{\gamma}, \quad \text{for all } x \in (a,b),$$

then $\mathrm{Cap}(\{x_0\}) > 0$.

The above result has the following simple extension to $n \geqslant 2$ if we assume that Σ is a finite union of C^1-smooth manifolds[24] [i.e., $\Sigma = \cup_{r=1}^{N}(\cap_{s=1}^{M_r}\Sigma_s)$, N, $M_r \in \mathbb{N}$, $\Sigma_s = \{x \in \Omega \mid F_s(x) = 0\}$, $F_s \in C^1(\overline{\Omega}_s)$, $\Omega_s \subset \mathbb{R}^n$ open, $(\nabla F_s)^2(x) > 0$ for $x \in \overline{\Omega}_s$]: If $\phi \in L^{\infty}_{\mathrm{loc}}(U)$ for some open set $U \subset \Omega$ with $\Sigma \subset U$ and the codimension of Σ is greater than 1, we have $\mathrm{Cap}(\Sigma) = 0$. Moreover a one-point set $\{x_0\} \subset \Omega \subset \mathbb{R}^n$ has zero capacity for any $n \geqslant 1$ if in a neighborhood of x_0 the function ϕ obeys

$$|\phi(x)| \leqslant C|x-x_0|^{(2-n)/2}, \quad \text{for some } C > 0. \tag{2.21}$$

[This follows from Lemma 2.6 since for f spherically symmetric around x_0 we have $\tilde{E}_{\phi}(f,f) \leqslant \mathrm{const}\int_0^{\infty} r\,dr|f'(r)|^2$.] Similarly, if for some positive constants $C, \gamma, 0 < \gamma < 1$, ϕ fulfills

$$|\phi(x)| \geqslant C|x-x_0|^{(\gamma+1-n)/2}, \tag{2.22}$$

for all x in a neighborhood of x_0, then $\mathrm{Cap}(\{x_0\}) > 0$.

If one assumes that $|\phi|$ is bounded away from zero in a neighborhood of Σ and Σ is a smooth manifold of codimension 1, then $\mathrm{Cap}(\Sigma) > 0$. For example, the surface of a sphere in $\mathbb{R}^3$ has positive finite capacity. Under the same conditions on ϕ a plane in $\mathbb{R}^3$ has nonzero (possibly infinite) capacity. If for some $C > 0$, $\delta > 0$, $|\phi(x)| > C$ holds for all x such that the distance from x to the plane is smaller than δ the capacity of the plane is infinite. The following example shows how to construct a ϕ such that a plane in $\mathbb{R}^3$ has finite capacity.

Example 2.7: Let $\Omega = \mathbb{R}^3$, $\Sigma = \{x \in \mathbb{R}^3 | x_3 = 0\}$ and assume

$$\phi(x) = \varphi(x_1,x_2)\psi(x_3), \quad \varphi, \Delta\varphi \in L^2_{\mathrm{loc}}(\mathbb{R}^2),$$
$$\psi, \psi' \in L^2_{\mathrm{loc}}(\mathbb{R}), \quad \varphi, \psi \text{ real valued.}$$

If $C(r) \subset \Sigma$ denotes the closed circle of radius r centered at the origin we have[16]

$$\mathrm{Cap}(\Sigma) = \sup_{r>0} \mathrm{Cap}(C(r)) = \lim_{r \to \infty} \mathrm{Cap}(C(r)). \tag{2.23}$$

Next we assume that σ represents the capacity of $\{0\} \subset \mathbb{R}$ computed with the form

$$\dot{\tilde{E}}_{\psi}(f,f) = \int_{\mathbb{R}} \psi^2(x)dx|f'(x)|^2 + \int_{\mathbb{R}} \psi^2(x)dx|f(x)|^2, \quad \mathscr{D}(\dot{\tilde{E}}_{\psi}) = C_0^{\infty}(\mathbb{R}), \tag{2.24}$$

and denote by e the corresponding equilibrium potential. Let $P(r)$ represent the capacity of $C(r)$ computed with the form

$$\dot{\tilde{E}}_{\varphi}(g,g) = \int_{\mathbb{R}^2} \varphi^2(x)d^2x|(\nabla g)(x)|^2 + \int_{\mathbb{R}^2} \varphi^2(x)d^2x|g(x)|^2, \quad \mathscr{D}(\dot{\tilde{E}}_{\varphi}) = C_0^{\infty}(\mathbb{R}^2), \tag{2.25}$$

with e_r the associated equilibrium potential. Then

$$\begin{aligned}\mathrm{Cap}(C(r)) &\leqslant \int_{\mathbb{R}^3} \varphi^2(x_1,x_2)\psi^2(x_3)d^3x \\ &\quad \times \{|(\partial_1 e_r)(x_1,x_2)e(x_3)|^2 \\ &\quad + |(\partial_2 e_r)(x_1,x_2)e(x_3)|^2 \\ &\quad + |e_r(x_1,x_2)(\partial_3 e)(x_3)|^2 \\ &\quad + |e_r(x_1,x_2)e(x_3)|^2\} \\ &\leqslant P(r)\int_{\mathbb{R}} \psi^2(x_3)dx_3|e(x_3)|^2 \leqslant P(r)\sigma\end{aligned} \tag{2.26}$$

(where $\partial_j = \partial/\partial x_j$, $j = 1,2,3$). Hence $\mathrm{Cap}(\Sigma)$ is finite if $P(r)$ stays bounded as $r \to \infty$. Next we assume

$$\varphi(x_1,x_2) = \omega(\rho), \quad \rho = (x_1^2 + x_2^2)^{1/2},$$
$$\omega(\rho) = \rho^{\alpha}, \quad \alpha < -1, \quad \text{for } \rho \geqslant 1,$$

and $g \in \mathscr{D}(\tilde{E}_{\varphi})$,

$$g(x_1,x_2) = h(\rho), \quad h(\rho) = \begin{cases} 1, & 0 \leqslant \rho \leqslant r, \\ r^{-\beta}\rho^{\beta}, & \rho \geqslant r, \quad \beta < 0. \end{cases}$$

Then for $r > 1$ we have

$$P(r) \leqslant 2\pi\beta^2 r^{-2\beta}\int_r^{\infty} d\rho\,\rho^{2\alpha+2\beta-1} + 2\pi\int_0^1 \rho\,d\rho|\omega(\rho)|^2 + 2\pi\int_1^r d\rho\,\rho^{2\alpha+1} + 2\pi r^{-2\beta}\int_r^{\infty} d\rho\,\rho^{2\alpha+2\beta+1}, \tag{2.27}$$

which is obviously bounded as $r \to \infty$.

After these preliminaries we are ready to state the following theorem.

Theorem 2.8: Assume in addition to conditions (I) and (II) that $\phi \in L^2_{\mathrm{loc}}(\Omega)$. Then both $\dot{E}_{\phi}$ and $\dot{\tilde{E}}_{\phi}$ are closable and $E_{\phi} = \tilde{E}_{\phi}$ if and only if $\mathrm{Cap}(\Sigma) = 0$.

Proof: Since ϕ is admissable in the sense of Ref. 1, $\dot{\tilde{E}}_{\phi}$ is closable. Closability of $\dot{E}_{\phi}$ follows from Theorem 2.5 (or simply from the fact that $\dot{E}_{\phi} \subset \tilde{E}_{\phi}$).

(a) Assume $\mathrm{Cap}(\Sigma) = 0$: Let $U_m \subset \Omega$, $m \in \mathbb{N}$ be a sequence of open sets such that $\Sigma \subset U_m$, $m \in \mathbb{N}$, and $\lim_{m \to \infty} \mathrm{Cap}(U_m) = \mathrm{Cap}(\Sigma) = 0$ (since Σ is closed the existence of U_m follows from Refs. 1 or 16). Let e_m be the equilibrium potential of U_m. Since $e_m \in \mathscr{D}(\tilde{E}_{\phi})$ there is a sequence $\{g_{m,k}\}_{k \in \mathbb{N}} \subset C_0^{\infty}(\Omega)$ converging to e_m in $\tilde{E}_{\phi}$-graph norm as $k \to \infty$. Without loss of generality we may assume that for any $k \in \mathbb{N}$ there is an open set $U_{m,k} \subset \Omega$ such that $\overline{U_m} \subset U_{m,k}$ and $g_{m,k}(x) = 1$ for $x \in U_{m,k}$, $m \in \mathbb{N}$. Let $f \in C_0^{\infty}(\Omega)$, then $(f - g_{m,k}f) \in C_0^{\infty}(\Omega - \Sigma)$. From

$$\begin{aligned}&\tilde{E}_{\phi}(f-(f-e_mf), f-(f-e_mf)) + \|\phi[f-(f-e_mf)]\|_2^2 \\ &\quad = \int_{\Omega} \phi^2(x)d^nx|\nabla(e_mf)(x)|^2 + \int_{\Omega} \phi^2(x)d^nx|(e_nf)(x)|^2 \\ &\quad \leqslant \sup_{x \in \Omega}[|\nabla f(x)|^2 + |f(x)|^2]\{\tilde{E}_{\phi}(e_m,e_m) + \|\phi e_m\|_2^2\} \\ &\qquad + 2\left[\tilde{E}_{\phi}(e_m,e_m)\int_{\Omega} \phi^2(x)d^nx|\nabla f(x)|^2|(e_mf)(x)|^2\right]^{1/2} \\ &\quad \leqslant c\,\mathrm{Cap}(U_m) \underset{m \to \infty}{\to} 0,\end{aligned}$$

where c is independent of m (and only depends on f), we infer that $\{f_j\}_{j\in\mathbb{N}} = \{f - g_{m_j,k_j} f\}_{j\in\mathbb{N}} \subset C_0^\infty(\Omega - \Sigma)$ converges to f in $\tilde{E}_\phi$-graph norm as $j \to \infty$. Hence $C_0^\infty(\Omega - \Sigma)$ is a core for $\tilde{E}_\phi$.

(b) Suppose that Σ is compact and $\mathrm{Cap}(\Sigma) > 0$: We first note that in this case $\mathrm{Cap}(\Sigma)$ is necessarily finite. Indeed there is an open bounded set A such that $\Sigma \subset A$ and $\overline{A} \subset \Omega$. Hence we can find a function $h \in C_0^\infty(\Omega)$ such that $h(x) > 1$ for $x \in A$. Since by definition $C_0^\infty(\Omega) \subset \tilde{E}_\phi$ we obtain $h \in \mathscr{L}_A$ and thus $\mathscr{L}_A \neq 0$ implying the finiteness of $\mathrm{Cap}(A)$. By the monotonicity property of capacity, $\mathrm{Cap}(\Sigma)$ is also finite. Next let $f \in C_0^\infty(\Omega)$ and without loss of generality we assume $f(x) > 1$ for $x \in \Sigma$. Assume that $\{f_m\}_{m\in\mathbb{N}} \subset C_0^\infty(\Omega - \Sigma)$ converges to f in $\tilde{E}_\phi$-graph norm. Let $U'_m = \Omega - \mathrm{supp} f_m$, $m \in \mathbb{N}$, then U'_m are open and $\Sigma \subset U'_m$ for all $m \in \mathbb{N}$. Define open sets U_m such that $\Sigma \subset U_m \subset U'_m$, $m \in \mathbb{N}$ and $\lim_{m\to\infty} \mathrm{Cap}(U_m) = \mathrm{Cap}(\Sigma)$. Since $(f - f_m)(x) \geqslant 1$ for $x \in U_m$ and m large enough we get the contradiction

$$0 < \mathrm{Cap}(\Sigma) \leqslant \mathrm{Cap}(U_m) \leqslant \tilde{E}_\phi(f - f_m, f - f_m) + \|\phi(f - f_m)\|_2^2 \underset{m\to\infty}{\to} 0.$$

(c) Let Σ be unbounded and $\mathrm{Cap}(\Sigma) > 0$: We define the sequence $\{\Sigma_m\}_{m\in\mathbb{N}} = \{S_m \cap \Sigma\}_{m\in\mathbb{N}}$ of compact sets, where S_m denotes the closed sphere in $\mathbb{R}^n$ of radius m centered at the origin. By the theorem of capacitability (cf. Refs. 16 and 23) and Theorem 3.1.1 of Ref. 16 we have

$$\mathrm{Cap}(\Sigma) = \sup_{m\in\mathbb{N}} \mathrm{Cap}(\Sigma_m).$$

Since $\mathrm{Cap}(\Sigma) > 0$ it follows that $\mathrm{Cap}(\Sigma_{m_0}) > 0$ for some $m_0 \in \mathbb{N}$. If we introduce

$$\dot{E}_{\phi,m_0}(f,g) = \int_\Omega \phi^2(x) d^n x\, \overline{(\nabla f)(x)}(\nabla g)(x),$$
$$\mathscr{D}(\dot{E}_{\phi,m_0}) = C_0^\infty(\Omega - \Sigma_{m_0}),$$

with E_{ϕ,m_0} the corresponding closure we get that $E_{\phi,m_0} \subsetneq \tilde{E}_\phi$ by part (b). But $E_\phi \subset E_{\phi,m_0} \subsetneq \tilde{E}_\phi$ then concludes the proof. □

Remark 2.9: To the regular Dirichlet form $\tilde{E}_\phi$ (in the sense of Ref. 16) there corresponds an m-symmetric Hunt process X_t of diffusion type with generator given by $\tilde{E}_\phi$. The condition $\mathrm{Cap}(\Sigma) = 0$ is equivalent with Σ exceptional for X_t in the sense of Ref. 16, i.e., such that there exists a Borel set $\tilde{\Sigma} \supset \Sigma$ such that $P_m(\sigma_{\tilde{\Sigma}} < \infty) = 0$, where $\sigma_{\tilde{\Sigma}} = \inf\{t > 0 | X_t \in \tilde{\Sigma}\}$ is the hitting time for X_t and P_m is the measure defined on Ω by $P_m(\Lambda) = \int_\Lambda P_x(\Lambda) m(dx)$, $dm = \phi^2\, dx$, P_x the law of the process X_t started at x, and Λ a measurable subset of Ω. Intuitively, thus, $\mathrm{Cap}(\Sigma) = 0$ corresponds to the fact that the process X_t a.s. does not reach $\tilde{\Sigma}$ in finite time, hence we are not free to prescribe boundary conditions on Σ.

We note that the fact that $C_0^\infty(\mathbb{R}^n - \{x_0\})$, $x_0 \in \mathbb{R}^n$ is a form core for $H_0 = -\Delta$, $\mathscr{D}(H_0) = H^{2,2}(\mathbb{R}^n)$ for all $n \geqslant 2$ as proved, e.g., in Ref. 25, trivially follows from the above theorem by taking $\Omega = \mathbb{R}^n$, $\Sigma = \{x_0\}$, $\phi(x) = 1$, observing that $\mathrm{Cap}(\{x_0\}) = 0$ for $n \geqslant 2$ [cf. Eq. (2.21)].

For numerous explicitly solved examples where the above phenomenon (viz. the disappearance of a boundary condition on Σ in $\dot{E}_\phi$ when the closure of $\dot{E}_\phi$ is considered) can be studied directly; see Ref. 1.

III. SOME PARTICULAR CASES

A. $\Omega = \mathbb{R}^n$, $\Sigma = \phi$

Following Refs. 12 and 13 we first introduce the class of potentials K_n: A measurable function $V: \mathbb{R}^n \to \mathbf{R}$ lies in K_n iff

$$\sup_{x\in\mathbf{R}} \int_{|x-y|<1} dy |V(y)| < \infty, \quad \text{if } n = 1,$$

$$\lim_{\epsilon\to 0_+} \left\{ \sup_{x\in\mathbb{R}^2} \int_{|x-y|<\epsilon} d^2y [-\ln|x-y|] |V(y)| \right\} = 0, \quad \text{if } n = 2,$$

$$\lim_{\epsilon\to 0_+} \left\{ \sup_{x\in\mathbb{R}^n} \int_{|x-y|<\epsilon} d^n y |x-y|^{-(n-2)} |V(y)| \right\} = 0, \quad \text{if } n \geqslant 3, \tag{3.1}$$

Moreover $V \in K_n^{\mathrm{loc}}$ iff $\chi_R V \in K_n$ for all $R > 0$ (χ_R being the characteristic function of the ball $\{x \in \mathbb{R}^n | |x| \leqslant R\}$), in particular $K_1^{\mathrm{loc}} = L^1_{\mathrm{loc}}(\mathbb{R})$.

Assuming

(A1) $V = V_1 + V_2$, $c \leqslant V_1 \in K_n^{\mathrm{loc}}$, for some $c \in \mathbb{R}$ and $V_2 \in K_n$,

then hypothesis (I) with $\Omega = \mathbb{R}^n$, $\Sigma = \varnothing$ is fulfilled. Moreover $C_0^\infty(\mathbb{R}^n)$ is a form core for $H = -\Delta + V$ in $L^2(\mathbb{R}^n)$ (see Refs. 12 and 13). [If in addition $V \in L^2_{\mathrm{loc}}(\mathbb{R}^n)$ then $C_0^\infty(\mathbb{R}^n)$ is an operator core for H.] In order to show that also assumption (II) is obeyed we recall the following.

Lemma 3.1 [12]: Assume (A1). Then, for some $E \in \mathbb{R}$, $-\Delta\phi + V\phi = E\phi$ has a nonzero distributional solution ϕ [in the sense that ϕ, $V\phi \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $-(\Delta f, \phi) + (f, V\phi) = E(f,\phi)$ for all $f \in C_0^\infty(\mathbb{R}^n)$] which is nonnegative iff $\inf \sigma(H) \geqslant E$. Moreover if ϕ is such a solution it equals a.e. a continuous function and $(\nabla\phi)_j \in L^2_{\mathrm{loc}}(\mathbb{R}^n)$, $1 \leqslant j \leqslant n$. In addition, Harnack's inequality

$$\phi(x) \leqslant C\phi(y), \quad x, y \in \mathbb{R}^n,$$

where C is independent of x, y if $x, y \in K$ for compact subsets $K \subset \mathbb{R}^n$, implies the existence of $C_R > 0$ such that

$$C_R^{-1} \leqslant \phi(x) \leqslant C_R, \quad \text{if } |x| < R.$$

Consequently assumptions (I) and (II) with $\Omega = \mathbb{R}^n$, $\Sigma = \varnothing$ are fulfilled if V obeys (A1). In particular the assertions of Theorem 2.5 are valid.

In general, if $E < \inf \sigma(H)$ there is no ϕ_0 that is distinguished among all possible ϕ's as the example $V = 0$ and $\phi(x) = e^{(-E)^{1/2}\hat{e}\cdot x}$, $E < 0$, $\hat{e}$ some unit vector in $\mathbb{R}^n$ (satisfying $-\Delta\phi = E\phi$) shows. This situation changes if $E \to \inf \sigma(H)$ as described below.

Lemma 3.2: In addition to hypothesis (A1) assume that $E_0 = \inf \sigma(H)$ belongs to the discrete spectrum of H. Then E_0 is simple and the corresponding eigenfunction ϕ_0 (the

ground state) can be chosen to be continuous and strictly positive. In particular ϕ_0 is the only distributional solution of $(-\Delta + V)\phi_0 = E_0\phi_0$ (up to multiplicative constants) with $\phi_0, (\nabla\phi_0)_j \in L^2(\mathbb{R}^n)$, $1 \leqslant j \leqslant n$.

Sketch of proof: Strict positivity of ϕ follows from positivity improving of the semigroup e^{-tH}, $t>0$ (see Refs. 26 and 27). The remaining assertions are contained in Lemma 3.1. □

It remains to discuss the case where $E_0 = \inf \sigma_{ess}(H)$ and the discrete spectrum of H is empty. We first treat the case $n \geqslant 3$. Instead of Assumption (A1) we now introduce the stronger condition

(A2) for some $0 < \epsilon < n/2 - 1$,
$V \in L^{n/2-\epsilon}(\mathbb{R}^n) \cap L^{n/2+\epsilon}(\mathbb{R}^n)$ real valued, $n \geqslant 3$.

Hypothesis (A2) implies $\sigma_{ess}(H) = [0, \infty)$ (cf. Theorem XIII.15 in Ref. 27).

Lemma 3.3 [10]: Assume condition (A2) and $\inf \sigma(H) = 0$. Then there is a unique $\phi_0 \in L^\infty(\mathbb{R}^n)$ satisfying

$$\phi(x) = \phi_\infty - c_n \int d^n y |x-y|^{-(n-2)} V(y)\phi(y),$$

$$c_n = \Gamma(n/2)/2(n-2)\pi^{n/2}, \quad n \geqslant 3,$$

for some $\phi_\infty \in [0, \infty)$ [$\Gamma(z)$ being the gamma function[28]]. Moreover ϕ_0 is continuous, strictly positive, $(\nabla\phi_0)_j \in L^2_{loc}(\mathbb{R}^n)$, $1 \leqslant j \leqslant n$, and $(-\Delta + V)\phi_0 = 0$ in the sense of distributions.

Remark 3.4: (a) If $\phi_\infty > 0$, $\phi_0(x)$ corresponds to the zero-energy scattering solution of H in the "subcritical case,"[29] whereas if $\phi_\infty = 0$, $\phi_0(x)$ describes a zero-energy resonance (if $n = 3,4$) [resp. a zero-energy bound state (if $n \geqslant 5$)] of H in the so-called "critical case"[32] associated with the absorption of the ground state.[33] For the notion of zero-energy resonances (resp. bound states and scattering wave functions) we refer to Refs. 31 and 35–39.

(b) In contrast to assumption (A1), condition (A2) implies a strong fall of V_+ at infinity. As a consequence $\phi_0 \in L^\infty(\mathbb{R}^n) n \geqslant 3$ in Lemma 3.3. In fact the example $V(x) = \gamma/|x|$, $\gamma > 0$, $n = 3$ shows that in general $\phi_0 \in L^\infty_{loc}(\mathbb{R}^n)$ in Lemma 3.1 cannot be replaced by $L^\infty(\mathbb{R}^n)$ since[40]

$$\phi_0(x) = (\gamma|x|)^{-1/2} I_1((4\gamma|x|)^{1/2}), \quad \gamma > 0, \quad x \in \mathbb{R}^3$$

[$I_1(z)$ the modified Bessel function of order 1 (see Ref. 28)] obeys

$$(-\Delta + \gamma/|x|)\phi_0 = 0, \qquad 0 < \phi_0 \in C^\infty(\mathbb{R}^3)$$

and

$$\phi_0(x) \underset{|x| \to \infty}{=} O(|x|^{-3/4} \exp(4\gamma|x|)^{1/2})$$

in this particular case. (The example trivially extends to $n = 2$ and $n \geqslant 4$ and an appropriate modification of V at $x = 0$ also applies for $n = 1$.) For a general discussion concerning bounds on zero-energy (distributional) solutions ϕ_0 of H we refer to Refs. 12 and 41.

In condition (A2) we treated only the case $n \geqslant 3$. The main reason to single out $n = 1,2$ of course lies in the fact that under hypothesis (A2) for $\lambda \in \mathbb{R}$, $|\lambda|$ small enough, $H(\lambda) = -\Delta + \lambda V$ and $H_0 = -\Delta$ are unitarily equivalent [the wave operators $\Omega_\pm(\lambda) = \text{s-lim}_{t \to \pm\infty} e^{itH(\lambda)} e^{-itH_0}$ are unitary and hence $H(\lambda) = \Omega_\pm(\lambda) H_0 \Omega_\pm(\lambda)^*$ (see Refs. 27 and 42)]. If $V \leqslant 0$, $V \in L^p(\mathbb{R}^n) + L^q(\mathbb{R}^n)$, $1 < q < \infty$, $p = 1$ if $n = 1$, $p > 1$ if $n = 2$ then $H(\lambda) = -\Delta + \lambda V$ has a negative eigenvalue for all $\lambda > 0$ [and $\sigma_{ess}(H) = [0, \infty)$].[43] If $E_0 = \inf \sigma(H)$ belongs to the discrete spectrum of H and $n = 1,2$ then Lemma 3.2 applies. If, e.g., $n = 1$, $(1 + |x|^2)V \in L^1(\mathbb{R})$ and $E_0 = \inf \sigma(H) = 0$ it suffices to study integral equations of the type

$$\phi_0(x) = c_1 + c_2 x + 2^{-1} \int_{\mathbb{R}} dy |x-y| V(y)\phi_0(y) \qquad (3.2)$$

for appropriate $c_1, c_2 \in \mathbb{R}$ (cf. Refs. 44–46). By Lemma 3.1 or by the one-to-one correspondence between solutions of Eq. (3.2) and distributional solutions of $(-d^2/dx^2 + V(x))\phi = 0$ (which automatically have a locally absolutely continuous derivative) one obtains the existence of a strictly positive solution ϕ_0 of Eq. (3.2). In particular ϕ_0 obeys the assertions of Lemma 3.3 except that $\phi_0 \in L^\infty(\mathbb{R})$ only if V is critical. (For related results see also Refs. 47 and 48.) A corresponding treatment of the case $n = 2$ appeared in Ref. 49.

For possible applications of the above results (especially for $n = 3$) to prove the absence of negative bound states we refer to Ref. 50.

B. $\Omega \subset \mathbb{R}^n$ bounded, $\Sigma = \emptyset$

Let $\Omega \subset \mathbb{R}^n$ be an open, connected, and bounded set and denote by $-\Delta_D$ the corresponding Dirichlet Laplacian in $L^2(\Omega)$ defined in Eq. (2.7). Assuming

(B1) $V \in K_n^{loc}$,

V is relatively form-bounded with respect to $-\Delta_D$ with relative bound zero.[13] Thus the form sum $H = -\Delta_D + V$ is well defined, $C_0^\infty(\Omega)$ is a form core for H, and H has a compact resolvent. Moreover we have the following lemma.

Lemma 3.5: Assume condition (B1) and let $E_0 = \inf \sigma(H)$. Then E_0 is a simple eigenvalue of H (the ground state energy) and the corresponding eigenfunction ϕ_0 can be chosen to be continuous and strictly positive on Ω. Moreover $\phi_0 \in L^\infty(\Omega)$ and (up to multiplicative constants) ϕ_0 is the only distributional solution of $(-\Delta + V)\phi_0 = E_0\phi_0$ with $\phi_0, (\nabla\phi_0)_j \in L^2(\Omega)$, $1 \leqslant j \leqslant n$.

Sketch of proof: Since $(H - E)^{-1}$, $E < E_0$ is compact and positivity improving, E_0 is a simple eigenvalue of H, and the corresponding eigenfunction $\phi_0 \in L^2(\Omega)$ can be chosen to be non-negative[51] (cf. also Ref. 52). By the results of Refs. 12 and 13, ϕ_0 equals a.e. a continuous function and Harnack's inequality implies strict positivity of ϕ_0. Clearly $\phi_0 \in L^\infty(\Omega)$ since ϕ_0 vanishes on $\partial\Omega$. Thus conditions (I) and (II) are fulfilled and Theorem 2.5 applies. □

The above results can be generalized to unbounded domains $\Omega \subset \mathbb{R}^n$. In this context we also refer to Refs. 53–56

for results in connection with positive distributional solutions of $H\phi = E\phi$. In the special case $n = 1$, cf. also Ref. 48.

C. $\Sigma = \{x_0\}$

We finally show how different distributional solutions $\phi_j, j = 1,2$ of $-\Delta\phi_j + V\phi_j = E_j\phi_j, j = 1,2$ can lead to the same conclusions as in Lemma 2.4.

Example 3.6: Let $n = 1$, $\Omega = (0,2\pi)$, $\Sigma = \{\pi\}$, $V = 0$ and define the form

$$\dot{Q}_D(f,g) = (f',g'), \mathscr{D}(\dot{Q}_D) = C_0^\infty((0,2\pi) - \{\pi\}).$$

Then its closure is given by

$$\begin{aligned} &Q_D(f,g) = (f',g'), \\ &\mathscr{D}(Q_D) = \{f \in L^2((0,2\pi) | f \in \mathrm{AC}((0,2\pi)); \\ &\qquad f(0_+) = f(\pi_\pm) = f(2\pi_-) = 0; \\ &\qquad f' \in L^2((0,2\pi))\}, \end{aligned}$$

where $\mathrm{AC}((0,2\pi))$ denotes the set of absolutely continuous functions on $(0,2\pi)$. Let

$$\phi_0(x) = \sin(x/2), \quad \phi_1 = \sin x, \quad \phi_2 = |\sin x|,$$

then obviously

$$\phi_j C_0^\infty((0,2\pi) - \{\pi\}) = C_0^\infty((0,2\pi) - \{\pi\}), \quad 1 \leqslant j \leqslant 3,$$

and hence $\phi_j C_0^\infty((0,2\pi) - \{\pi\})$ is a core for Q_D. In fact it follows from the results in Ref. 1 that also $\phi_j C_0^\infty((0,2\pi))$ if $j = 2,3$ is a core for Q_D. That is, if ϕ is nodeless (as it is the case for ϕ_1 above) the boundary condition at Σ has to be put in by hand [i.e., by taking $C_0^\infty(\Omega - \Sigma)$ as the minimal domain]. On the contrary if $\tilde{\phi}$ corresponds to an "excited state" with a node on Σ, then also $\tilde{\phi} C_0^\infty(\Omega)$ represents a core for Q_D (resp. Q).

ACKNOWLEDGMENTS

The first three authors are indebted to Professor L. Streit for kind invitations to join the Research project Nr. 2 in mathematics and physics at ZiF, University of Bielefeld, Federal Republic of Germany, the stimulating atmosphere of which made possible this collaboration. We also thank Professor R. Høegh-Krohn for useful discussions and invitations to Oslo University.

We thank Mrs. Mischke and Mrs. Richter for their skillful typing. Financial support by the Alexander von Humboldt Foundation and GmbH-Kernforschung Karlsruhe are gratefully acknowledged by F. Gesztesy and W. Karwowski, respectively.

[1] S. Albeverio, R. Høegh-Krohn, and L. Streit, J. Math. Phys. **18**, 907 (1977).
[2] S. Albeverio, R. Høegh-Krohn, and L. Streit, J. Math. Phys. **21**, 1636 (1980).
[3] L. Streit, Phys. Rep. 77, 363 (1981).
[4] S. Albeverio, M. Fukushima, W. Karwowski, and L. Streit, Commun. Math. Phys. **81**, 501 (1981).
[5] S. Albeverio and R. Høegh-Krohn, "Some Remarks on Dirichlet Forms and their Applications to Quantum Mechanics and Statistical Mechanics," and "Diffusions, quantum fields, and groups of mappings," in *Functional Analysis and Markov Processes, Lecture Notes in Mathematics, Vol. 923*, edited by M. Fukushima (Springer, Berlin, 1982), pp. 120 and 133, respectively.
[6] S. Albeverio, W. Karwowski, M. Röckner and L. Streit, "Some remarks on Dirichlet forms, Green's functions and Schrödinger operators," BiBoS Preprint, 1985; S. Albeverio and R. Høegh-Krohn, "Diffusion fields, quantum fields, fields with values in Lie groups," in *Advances in Probability*, edited by M. Pinsky (Dekker, New York, 1984), pp. 1–98.
[7] N. Wielens, Diploma Thesis, Universität Bielefeld, 1982; N. Wielens, J. Funct. Anal. **61**, 98 (1985).
[8] M. Fukushima, Phys. Rep. **77**, 255 (1981); "Markov Processes and Functional Analysis," in *Proceedings of the International Mathematical Conference*, Singapore, 1981 (North-Holland, Amsterdam, 1981); I. G. Hooton, Trans. Am. Math. Soc. **253**, 237 (1979).
[9] R. Carmona, J. Funct. Anal. **33**, 259 (1979).
[10] B. Simon, J. Funct. Anal. **35**, 215 (1980).
[11] B. Simon, J. Funct. Anal. **40**, 66 (1981).
[12] B. Simon, Bull. Am. Math. Soc. **7**, 447 (1982).
[13] M. Aizenman and B. Simon, Commun. Pure Appl. Math. **35**, 209 (1982).
[14] (a) E. B. Davies and B. Simon, "Ultracontractive semigroups and some problems in analysis," to appear in *Aspects of Mathematics and its Applications* in honor of L. Nachbin, and J. Funct. Anal. **59**, 335 (1984); (b) S. Kusuoka and D. Stroock, J. Funct. Anal. **60**, 243 (1985).
[15] Derivatives are understood in the sense of distributions in $C_0^\infty(\mathbf{R}^n)'$.
[16] M. Fukushima, *Dirichlet Forms and Markov Processes* (North-Holland/Kodansha, Amsterdam, 1980).
[17] J. F. Brasche, Diploma Thesis, Universität Bielefeld, 1983; J. Math. Phys. **26**, 621, 1984; K. Rullkötter and U. Spönemann, Diploma Thesis, Universität Bielefeld, 1983, and paper in preparation.
[18] M. Röckner and N. Wielens, "Dirichlet forms—closability and change of speed measure," in *Infinite Dimensional Analysis and Stochastic Processes*, edited by S. Albeverio (Pitman, London, 1985); J. Marion, "Closability of some singular Dirichlet forms," preprint, Univ. Marseille-Luminy, 1984.
[19] $H^{m,p}(\Omega)$ denotes the corresponding Sobolev space.
[20] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. II: Fourier Transform, Essential Self-Adjointness* (Academic, New York, 1975).
[21] H. L. Cycon, J. Operator Theory **6**, 75 (1981).
[22] R. A. Adams, *Sobolev Spaces* (Academic, New York, 1975), p. 30.
[23] P. A. Meyer, *Probability and Potentials* (Ginn, Boston, MA, 1966).
[24] Y. Choquet-Bruhat, C. De Witt-Morette, and M. Dillard-Bleick, "*Analysis, Manifolds and Physics*" (North-Holland, Amsterdam, 1982), revised edition.
[25] W. G. Faris, "Self-adjoint operators," in *Lecture Notes in Mathematics, Vol. 437* (Springer, Berlin, 1975).
[26] W. G. Faris and B. Simon, Duke Math. J. **42**, 559 (1975).
[27] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. IV: Analysis of Operators* (Academic, New York, 1978).
[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
[29] That is, $\sigma(-\Delta + (1+\epsilon)V) \cap (-\infty, 0) = \emptyset$ for $0 \leqslant \epsilon \leqslant \epsilon_0$ for some $\epsilon_0 > 0$ (cf. Refs. 10–12, 30, 31).
[30] B. Simon, J. Funct. Anal. **25**, 338 (1977).
[31] M. Klaus and B. Simon, Ann. Phys. **130**, 251 (1980).
[32] That is, $\sigma(-\Delta + V) \cap (-\infty, 0) = \emptyset$, but $\sigma(-\Delta + (1+\epsilon)V) \cap (-\infty, 0) \neq \emptyset$ for $0 < \epsilon < \epsilon_0$, $0 < \epsilon_0$ suitable (cf. Refs. 10–12, 30, 31).
[33] Note that the wave operators $\Omega_\pm = \text{s-lim}_{t \to \pm\infty} e^{itH} e^{-itH_0}$ exist due to condition (A2) (cf. Theorem XI.26 in Ref. 34 and its analog for $n > 3$).
[34] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. III: Scattering Theory* (Academic, New York, 1979).
[35] R. G. Newton, J. Math. Phys. **18**, 1348 (1977).
[36] A. Jensen and T. Kato, Duke Math. J. **46**, 583 (1979).
[37] A. Jensen, Duke Math. J. **47**, 57 (1980); J. Math. Anal. Appl. **101**, 397 (1984).
[38] S. Albeverio, F. Gesztesy, and R. Høegh-Krohn, Ann. Inst. H. Poincaré A **37**, 1 (1982).
[39] M. Murata, J. Funct. Anal. **49**, 10 (1982).
[40] S. Albeverio, D. Bollé, F. Gesztesy, R. Høegh-Krohn, and L. Streit, Ann. Phys. **148**, 308 (1983).
[41] B. Simon, J. Funct. Anal. **42**, 347 (1981).
[42] T. Kato, Math. Ann. **162**, 258 (1966).

[43]B. Simon, "On the number of bound states of two-body Schrödinger operators—a review," in *Studies in Mathematical Physics, Essays in Honor of V. Bargman*, edited by E. H. Lieb, B. Simon, and A. S. Wightman (Princeton U. P., Princeton, NJ, 1976).
[44]P. A. Deift and E. Trubowitz, Commun. Pure Appl. Math. **32**, 121 (1979).
[45]M. Klaus, Helv. Phys. Acta **55**, 49 (1982).
[46]D. Bollé, F. Gesztesy, and S. F. J. Wilk, Phys. Lett. A **97**, 30 (1983); J. Operator Theory **13**, 3 (1985).
[47]U. W Schmincke, Proc. R. Soc. Edinburgh Sect. A **80**, 67 (1978).
[48]P. A. Deift, Duke Math. J. **45**, 267 (1978).
[49]M. Murata, J. Funct. Anal. **56**, 300 (1984).
[50]A. Martin and P. C. Sabatier, J. Math. Phys. **18**, 1623 (1977).
[51]H. W. Goelden, Math. Z. **155**, 239 (1977).
[52]F. Gesztesy, Rep. Math. Phys. **20**, 93 (1984).
[53]W. F. Moss and J. Piepenbrink, Pac. J. Math. **75**, 219 (1978).
[54]W. Allegretto, Pac. J. Math. **92**, 15 (1981).
[55]U. W. Schmincke, Arch. Rat. Mech. Anal. **75**, 147 (1981).
[56]E. Müller-Pfeiffer, Proc. R. Soc. Edingburgh Sect. A **94**, 287 (1983).

Not necessary but sufficient condition for the positivity of generalized Wigner functions

G. Mourgues,[a] M. R. Feix, and J. C. Andrieux
Equipe Physique Mathematique, Modelisation, Simulation/Centre National de la Recherche Scientifique, 3 A Avenue de la Recherche Scientifique, 45045 Orleans Cedex, France

P. Bertrand
Laboratoire de Physique Théorique, Université de Nancy I, 54506 Vandoeuvre Les Nancy, France

(Received 18 September 1984; accepted for publication 7 June 1985)

Many authors, both in quantum mechanics and signal theory, are concerned with a condition of positivity for the generalized Wigner functions. They give (and use) as necessary and sufficient the condition that the kernel be the characteristic function of another Wigner function. It is shown that the condition is not necessary, but only sufficient.

I. INTRODUCTION

The generalized Wigner functions introduced by Cohen[1] have been mathematically justified and studied by Ruggeri,[2] Srinivas and Wolf,[3] Amiet and Huguenin,[4] Kruger and Poffyn,[5] and finally Springborg.[6] These generalized distributions are sesquilinear with respect to the wave function.

The initial aim was to interpret quantum mechanics in terms of classical concepts, which were expected to be especially useful in the "quasiclassical" limit (See Tatarskii).[7] Each of these generalized distributions are linked to a correspondence principle associating operators (quantum observables) with "corresponding" classical phase-space functions.

In this brief report, we are mainly interested in the positivity of these generalized distributions.

II. PROCEDURE

Without loss of generality, we consider that the quantum system is in a pure state ψ.

The generalized Cohen[1] distributions are given by

$$f_\psi^G(q,p) = \frac{1}{(2\pi\hbar)^2}\int_{R^3} \Omega(u,v) e^{-(i/\hbar)(up+vq-bv)} \times \overline{\psi\left(b-\frac{u}{2}\right)}\psi\left(b+\frac{u}{2}\right) du\, dv\, db, \tag{1}$$

and the associated correspondence principle (see Refs. 3 and 4) imposes that $\Omega(u,v)$ must satisfy the following properties: (a) the integral (1) must be convergent; (b) $\Omega(0,0)=1$; (c) $\overline{\Omega(u,v)} = \Omega(-u,-v)$; and (d) $\Omega(u,v)$ must have neither zero nor pole, and must be infinitely differentiable.

Note that the original Wigner Weyl distribution $f_\psi^w(q,p)$ is obtained by taking the particular choice $\Omega(u,v)=1$. The essential property of these generalized distributions is that for any observable B [whose corresponding classical function is $b(q,p)$] we have

$$\langle B\rangle_\psi = \mathrm{tr}(|\psi\rangle\langle\psi|B) = \langle\psi|B|\psi\rangle = \int_{R^2} b(q,p) f_\psi^G(q,p)\, dq\, dp.$$

Unfortunately, the marginal distributions are only obtained if $\Omega(0,v)=\Omega(u,0)=1$ for every u and v, and the positivity property [$f_\psi^G(q,p)\geqslant 0$ everywhere in the phase space] for every state ψ is incompatible with the precedent condition, as proved by Wigner.[8]

Srinivas and Wolf[3] have claimed that a *necessary* and *sufficient* condition giving this positivity is that $\Omega(u,v)$ be the characteristic function of the Wigner Weyl distribution $f_\theta^w(q,p)$ of some arbitrary state θ, i.e.,

$$\Omega(u,v) = \mathscr{F}_{v,u}(f_\theta^w(q,p)) = \int_{R^2} e^{(i/\hbar)(up+vq)} f_\theta^w(q,p)\, dq\, dp. \tag{2}$$

This condition is certainly sufficient. Noticing that (1) gives

$$\mathscr{F}_{v,u}(f_\psi^G(q,p)) = \Omega(u,v)\langle e^{(i/\hbar)(uP+vQ)}\rangle_\psi, \tag{3}$$

where Q and P are the usual position and momentum operators, we obtain, taking $\Omega(u,v)=1$ and substituting θ to ψ in (3)

$$\mathscr{F}_{v,u}(f_\theta^w(q,p)) = \langle e^{(i/\hbar)(uP+vQ)}\rangle_\theta. \tag{4}$$

This is exactly the "quasicharacteristic function" of the state θ as defined by Cushen and Hudson.[9]

Rewriting (3) with $\Omega(u,v)$ as chosen in (2), with (4), we obtain

$$\mathscr{F}_{v,u}(f_\psi^G(q,p)) = \langle e^{(i/\hbar)(uP+vQ)}\rangle_\theta \langle e^{(i/\hbar)(uP+vQ)}\rangle_\psi. \tag{5}$$

Cushen and Hudson[9] have proved that the product of two quasicharacteristic functions is a characteristic function and therefore is a function of positive type in Bochner's sense (see Ref. 10).

Taking the double inverse Fourier transform of (5), we obtain

$$f_\psi^G(q,p) = f_\theta^w(q',p') \underset{q,p}{*} f_\psi^w(q',p') \geqslant 0$$

(where $*$ denotes double convolution), and

$$\int_{R^2} f_\theta^w(q-q', p-p') f_\psi^w(q',p')\, dq'\, dp'.$$

However, condition (2) is not necessary.

The origin of the error can be traced back in Srinivas and

[a] Institut Universitaire de Technologie de Montluçon, Université de Clermont Ferrand II, France.

 0022-2488/85/102554-02$02.50

Wolf,[3] p. 1481, where, after formula 4.4b, the authors claim "...$\Delta(\hat{q}, \hat{p}, q, p)$ is a self-adjoint operator with constant trace C and hence $\Delta(\hat{q}, \hat{p}, q, p)/C$ satisfies for every q and p all the properties of a density operator." The word "hence" has no justification whatsoever and consequently Theorem 4 is not correct. We show this assertion through the following counterexample issued from the Cartwright[11] smoothing method.

We choose as state θ the centered Gaussian which will be written with the notation of the coherent fundamental state (namely 0) of the harmonic oscillator (mass m and frequency ω).

$$\langle q|0\rangle = (m\omega/\hbar)^{1/4} e^{-(1/2)(m\omega/\hbar)q^2},$$

so

$$\Omega(u, v) = \langle e^{(i/\hbar)(uP + vQ)}\rangle_0 = e^{-(1/4\hbar)(m\omega u^2 + v^2/m\omega)}.$$

Taking for $\xi(u, v)$ a double Gaussian

$$\xi(u, v) = e^{-(1/2\hbar^2)(\Sigma_p{}^2 u^2 + \Sigma_q{}^2 v^2)},$$

with Σ_q and Σ_p arbitrary chosen, but positive, we obtain

$$\tilde{\Omega}(u, v) = e^{-(1/4\hbar)(m\omega u^2 + v^2/m\omega)} e^{-(1/2\hbar^2)(\Sigma_p{}^2 u^2 + \Sigma_q{}^2 v^2)}.$$

This new $\tilde{\Omega}$ satisfies obviously the four conditions (a)–(d), and

$$\mathcal{F}^{-1}_{p,q}(\tilde{\Omega}(u, v)) = (1/2\pi\sigma_q\sigma_p) e^{-(1/2)(q^2/\sigma_q{}^2 + p^2/\sigma_p{}^2)},$$

with

$$\sigma_q{}^2 = \Sigma_q{}^2 + \hbar/2m\omega \quad \text{and} \quad \sigma_p{}^2 = \Sigma_p{}^2 + m\omega\hbar/2.$$

Consequently, we have the inequality $\sigma_q\sigma_p > \hbar/2$. This shows that $\tilde{\Omega}(u, v)$ cannot be the quasicharacteristic function of any state.

Indeed, coming back to the Srinivas and Wolf approach, let

$$\Delta(\hat{q}, \hat{p}, q, p) = \frac{1}{(2\pi\hbar)^2} \int_{R^2} \tilde{\Omega}(u, v) e^{(i/\hbar)[u(\hat{p} - p) + v(\hat{q} - q)]}\, du\, dv.$$

It means that $\Delta/\mathrm{Tr}(\Delta)$ is not a density operator. A simple calculation gives the following.

(i) $\mathrm{Tr}(\Delta) = (2\pi\hbar)^{-1}$.

(ii) $$\langle q'|\Delta/\mathrm{Tr}(\Delta)|q''\rangle = \sqrt{2/B}\,(2\pi\hbar)^{-1/2} e^{-(i/\hbar)(q'' - q')p} \times e^{-(1/4\hbar)A(q'' - q')^2} \times e^{-(1/\hbar B)(q - (q'' + q')/2)^2},$$

where $A = m\omega + 2\Sigma_p{}^2/\hbar$ and $B = 1/m\omega + 2\Sigma_q{}^2/\hbar$.

(iii) $$\langle q'|(\Delta/\mathrm{Tr}(\Delta))^2|q''\rangle = \sqrt{2/B}\,(2\pi\hbar)^{-1/2}(\sqrt{2}/\sqrt{1 + AB}) \times e^{-(i/\hbar)(q'' - q')p} e^{-(1/4\hbar)A(q'' - q')^2} \times e^{-(1/\hbar B)(q - (q'' + q')/2)^2} \times e^{[AB(q'' - q') - (4q - 3q'' - q')]^2/(8\hbar B(1 + AB))} \times e^{-(q - q'')^2/\hbar B}.$$

It is clear that only for $\Sigma_p = \Sigma_q = 0$ do we obtain the idempotence relation $(\Delta/\mathrm{Tr}(\Delta))^2 = (\Delta/\mathrm{Tr}(\Delta))$.

More generally, we give now a larger class of $\tilde{\Omega}(u, v)$ giving the positivity property without being associated to any wave function.

We consider an $\Omega(u, v)$ which is the quasicharacteristic function of some state θ. It guarantees the positivity of f^G_ψ. We construct another generalized Wigner distribution function [namely $\tilde{f}^G_\psi(q, p)$] associated to the new choice of $\Omega(u, v)$

$$\tilde{\Omega}(u, v) = \langle e^{(i/\hbar)(uP + vQ)}\rangle_\theta \xi(u, v), \tag{6}$$

where $\xi(u, v)$ is an arbitrary function of positive type in Bochner's sense, therefore verifying $\xi(0, 0) = 1$, $\xi(\overline{u, v}) = \xi(-u, -v)$, and $|\xi(u, v)| \leqslant |\xi(0, 0)|$. The properties (a)–(d) are consequently fulfilled. Moreover, it is easy to verify that $\tilde{f}^G_\psi(q, p)$ is everywhere positive for every state ψ, since

$$\mathcal{F}_{v,u}(f^G_\psi(q, p)) = \langle e^{(i/\hbar)(uP + vQ)}\rangle_\psi (\langle e^{(i/\hbar)(uP + vQ)}\rangle_\theta \xi(u, v)),$$

and taking the inverse Fourier transform,

$$\tilde{f}^G_\psi(q, p) = (f^W_\psi(q'', p'') \underset{q', p'}{*} f^W_\theta(q'', p'')) \underset{q, p}{*} \mathcal{F}^{-1}_{p', q'}(\xi(u, v)).$$

Consequently, $\tilde{f}^G_\psi(q, p)$ is the convolution of two positive functions $f^G_\psi(q', p')$ and $\mathcal{F}^{-1}_{p'q'}(\xi(u, v))$. Nevertheless, the previous $\tilde{\Omega}(u, v)$ as given in (6) is *not* strictly obliged to be a characteristic function of any state $\tilde{\theta}$.

The relation (6) defines a class of functions guaranteeing the positivity of the generalized Wigner distribution, larger than the class given by Srinivas and Wolf in Ref. 3.

Among the class of double Gaussian smoothing kernels, the one with $\sigma_p\,\sigma_q = \hbar/2$ is physically interesting since it is the Wigner function associated to the minimum wave packet. Finally, the need of a smoothing through a kernel of the class given by Srinivas and Wolf (may be an attractive idea for a physical interpretation) must be justified on properties beyond the mere positivity.

[1] L. Cohen, J. Math. Phys. **7**, 781 (1966).
[2] G. J. Ruggeri, Prog. Theor. Phys. **46**, 1703 (1971).
[3] M. D. Srinivas and E. Wolf, Phys. Rev. D **11**, 1477 (1975).
[4] J. P. Amiet and P. Huguenin, in *Mecaniques Classique et Quantique dans l'Espace des Phases* (Université de Neuchâtel, Neuchâtel, 1981), p. 100.
[5] J. G. Kruger and A. Poffyn, Physica A **85**, 84 (1976).
[6] M. Springborg, J. Phys. A: Math. Gen. **16**, 535 (1983).
[7] V. I. Tatarskii, Sov. Phys. Usp. **26**, 311 (1983).
[8] E. P. Wigner, in *Perspective in Quantum Theory* (M.I.T., Cambridge, MA, 1971), p. 25.
[9] C. D. Cushen and R. L. Hudson, J. Appl. Prob. **8**, 454 (1971).
[10] I. M. Guelfand and N. Y. Vilenkin, *Les Distributions T.4* (Collection Universitaire, Dunod, Paris, 1980), pp. 143.
[11] N. D. Cartwright, Physica A **83**, 210 (1976).

Spectral properties of the Dirac equation with anomalous magnetic moment

Horst Behncke
Fachbereich Mathematik/Informatik, Universität Osnabrück, Albrechtsrasse 28, 4500 Osnabrück, West Germany

(Received 31 October 1985; accepted for publication 29 March 1985)

Though the electron possesses an anomalous magnetic moment, this term has so far been treated by perturbation methods only. Here this rather singular term is treated nonperturbatively. For a general and large class of spherical potentials it is shown that the Hamiltonian is essentially self-adjoint. If the potentials vanish at infinity the essential spectrum is $\mathbb{R}\setminus(-m,m)$. These results hold for all, also strong, Coulomb-type potentials. We further establish that the eigenvalues for such Hamiltonians are stable and can be computed by a low-order perturbation theory, if $l \geq 3$. The above results can also be extended to multicenter potentials.

I. INTRODUCTION

In Refs. 1 and 2 it was shown that the Dirac Hamiltonian with an anomalous magnetic moment term is essentially self-adjoint for physically reasonable potentials. This has been demonstrated recently also by Gesztesy, Simon, and Thaller.[3] In addition it was shown for a large class of such operators that the essential spectrum is $(-m,m)^c$. In this paper these results are extended to multicenter Hamiltonians. In addition we show the stability of the eigenvalues for $l>3$ and a low-order pertubation theory for Coulomb-like Hamiltonians.

The notation of this paper is standard or is that of Refs. 1 and 2.

II. MULTICENTER HAMILTONIANS

Our first result states that Dirac Hamiltonians have well-localized properties, which permit the application of a general decomposition principle. Mathematically this result is due to the relative compactness of multiplication operators associated to compact sets.

Let M_i, $i=1,...,n$ be disjoint finite sets and let V_i be formally symmetric 4×4 matrix potentials with coefficients in $L^3_{\mathrm{loc}}(\mathbb{R}^3\setminus M_i)$. Moreover let $M_0=\cup_{i=1}^n M_i$ and $V_0=\Sigma_{i=1}^n V_i$. On $\mathscr{H}=L^2(\mathbb{R}^3)\otimes\mathbb{C}^4$ we are going to consider the Hamiltonians

$$H_{00}=\alpha\cdot(-i\nabla)+m\beta,$$
$$H_j=\alpha\cdot(-i\nabla)+m\beta+V_j,\quad j=0,1,...,n,$$

with domain $\mathscr{D}=C_0^\infty(\mathbb{R}^3\setminus M_0)$. Then the $H_j, j=0,1,...,n$, are symmetric and H_{00} is even essentially self-adjoint.

Theorem 1: Let M_j, V, and H_j be as above. Then

$$\mathrm{def}\,H=\sum_{j=1}^n \mathrm{def}\,H_j\,. \tag{1}$$

Moreover if $\|V_j(x)\|\to0$, for $\|x\|\to\infty$, and if each H_j has equal finite deficiency indices, then

$$\sigma_{\mathrm{ess}}(H_0')=\cup\sigma_{\mathrm{ess}}(H_j')\,,$$
$$(-m,m)^c\subset\sigma_{\mathrm{ess}}(H_j')\,,\quad j=0,1,...n. \tag{2}$$

Here H_j' denotes a self-adjoint extension of H_j.

Proof: (a) For $j=1,...,n$, let the O_j be open relatively compact disjoint neighborhoods of the M_j. Write $V_j=\chi_j\times V_j+U_j$, where χ_j is the characteristic function of O_j. Arguing as in the proof of Theorem 4 of Ref. 1, one sees that the U_j do not contribute to the deficiency index. If the U_j vanish at infinity, they do not even alter the essential spectrum. Thus we may assume $U_j=0$ for all j.

(b) As in the paper by Jörgens[4] one shows now

$$H_j^*u=\tilde{H}_{00}+\tilde{V}_ju\,,\quad u\in D(H_j^*)\,, \tag{3}$$

where the $\tilde{H}_j$ are unique extensions of the H_j. As in Ref. 4, we define for $\varphi\in\mathscr{C}^\infty(\mathbb{R}^3)$ the bounded operator $A(\varphi)$ by $A(\varphi)u=H_{00}\varphi u-\varphi H_{00}u$ on $\mathscr{D}_{H_{00}}$ and extend $A(\varphi)$ by continuity.

(c) For the remaining proof let φ_j, $j=1,...,n$, be a $C^\infty(\mathbb{R}^3)$ function, which vanishes off O_j and is equal to 1 on a neighborhood of M_j. Moreover, let φ be a $C^\infty(\mathbb{R}^3)$ function, which is 1 for large x and vanishes on $\cup_{j=1}^n O_j$. Then $\psi=1-\Sigma\varphi_i-\varphi$ satisfies $\psi\mathscr{D}(H_j^*)\subset\cap\mathscr{D}_j$, because of (3). Assume now $\mathrm{def}\,H_j=(m_j,n_j)$, $m_j,n_j<\infty$, for $j=1,...,n$, and let $u_{j,k}$, $k=1,...,m_j$, be independent solutions of $H_j^*\times u_{j,k}=-iu_{j,k}$. Then $\varphi_j u_{j,k}\in\mathscr{D}(H_0^*)\cap\mathscr{D}(H_j^*)$ and the $\{\varphi_j\times u_{j,k}\,|\,j=1,...,n,k=1,...,m_j\}$ are linear independent modulo $\mathscr{D}_0$. Since, moreover, $\mathrm{Im}\langle H_0^*\varphi_j u_{j,k}\,\varphi_j u_{j,k}\rangle<0$ we see $m_0\geqslant\Sigma_{j=1}^n m_j$. Similarly one shows the remaining relations $m_0\leqslant\Sigma m_j$ and $n_0\leqslant\Sigma n_j$, $n_0\geqslant\Sigma n_j$.

(d) Let $\lambda\in\sigma_{\mathrm{ess}}(H_j')$ and let $u_l\in\mathscr{D}_j$ be a λ-singular sequence. Since the operator multiplication with $(1-\varphi_j-\varphi)$ is compact, $(1-\varphi_j-\varphi)u_l\to0$ and $(H_j-\lambda)(1-\varphi_j-\varphi)u_l=(1-\varphi_j-\varphi)(H_j-\lambda)u_l-A(\varphi_j+\varphi)u_l\to0$. Thus a subsequence of $\{\varphi_j u_l\}$ or $\{\varphi u_l\}$ is a λ-singular sequence for H_0'. In the latter case it is even a λ-singular sequence for $\bar{H}_{00}$. This is only possible if $\lambda\notin(-m,m)$. Conversely let $\lambda\in\sigma_{\mathrm{ess}}(H_0')$ and let $\{u_l\}\subset\mathscr{D}_0$ be a λ-singular sequence for H_0'. Arguing as above, we see that a subsequence of $\{\varphi_1u_l\},...,\{\varphi_nu_l\}$ or $\{\varphi u_l\}$ is a λ-singular sequence of $H_j',...,H_n'$ or $\bar{H}_{00}$.

(e) For each $\lambda\notin(-m,m)$ there exists a λ-singular sequence $\{u_l\}$ for $\bar{H}_{00}$ with $\mathrm{supp}\,u_l\subset(\cup_{j=1}^n O_j)^c$. Then $\{u_l\}$ is also a λ-singular sequence for each $\bar{H}_j$.

Remark 1: The above result can be generalized in various ways. First of all, one can allow more general potentials. Off the O_j, local Kato boundedness with respect to $\bar{H}_{00}$ suffices. Second, it is possible to allow for a more general M_j. One only needs that H_{00} is essentially self-adjoint on $C_0^\infty(\mathbb{R}^3\setminus M_0)$. Finally, this theorem can be formulated in the more general setting of Jörgens,[4] who only allows ∞ as a

 0022-2488/85/102556-04$02.50

singular point. In fact a part of such a program has been carried out by Karnarski.[5] One might even attempt to formulate such decomposition techniques for higher-order operators like Schrödinger operators. The appropriate setting for these results, however, would be a Fredholm theory for differential systems.

Remark 2: The proof of the theorem shows that a λ in the essential spectrum with $|\lambda| \geqslant m$ is generally associated with a singular sequence localized near infinity. But a λ with $|\lambda| < m$ is necessarily associated to a sequence localized near one M_j.

Remark 3: Theorem 1 together with the results of Ref. 1 shows that any multicenter Coulomb–Dirac Hamiltonian with anomalous magnetic moment is essentially self-adjoint.

Remark 4: Large classes of Dirac Hamiltonians have equal deficiency indices. This holds in particular for all Hamiltonians which arise from relatively bounded perturbations of spherically symmetric Hamiltonians. Thus this is true for all operators studied here or in Refs. 1 and 2.

III. STABILITY

Because of Theorem 1 we shall consider from now on single center Hamiltonians only and assume that H has the form

$$H = \alpha \cdot (-i\nabla) + V_e + \beta V_s - i\beta\alpha \cdot (x/r) V_m + H_1 = H_0 + H_1, \tag{4}$$

with V_e, V_s, and V_m spherically symmetric. Here V_e is essentially due to the electrostatic potentials, while V_s is a scalar potential. The V_m describes the interaction due to the anomalous magnetic moment. It is the most singular term in H. We assume that H_1 is a symmetric 4×4 matrix potential, whose coefficients are linear combinations of the remaining less singular terms. In order for H to be a symmetric operator on H with domain $\mathscr{D}_H = \mathscr{C}_0^\infty(\mathbb{R}^3_+) \otimes \mathbb{C}^4$, where $\mathbb{R}^3_+ = \mathbb{R}^3 \setminus \{0\}$, we require

$$V_e, V_s, V_m \text{ are real values and in } L^3_{\text{loc}}(\mathbb{R}^3_+), \text{ the matrix elements of } H_1 \text{ are in } L^\infty_{\text{loc}}(\mathbb{R}^3). \tag{5}$$

Moreover we assume an almost Coulomb-like situation near 0 by requiring in addition

$$V_e = e/r + U_e, \quad V_s = a/r + U_s + m, \quad V_m = 1/r^2 = U_m,$$
$$U_e, U_s \text{ are twice differentiable near 0,}$$
$$U_m \text{ is differentiable near 0,} \tag{6}$$

$$\lim_{r\to 0} r^{1+i} U^{(i)}_{e,s}(r) = 0, \quad i = 0, 1, 2,$$

$$\lim_{r\to 0} r^2 U_m(r) = \lim_{r\to 0} r^3 U'_m(r) = 0,$$

$$\lim_{r\to\infty} U^{(i)}_{e,s}(r) = 0, \quad i = 0, 1, 2,$$

$$\lim_{r\to\infty} U^{(i)}_m(r) = 0, \quad i = 0, 1.$$

Lemma 1: Let V_e, V_s, V_m, and H_1 satisfy (4)–(6), as well as $\lim V_i(x) = \lim H_1(x) = 0$, for $x \to \infty$. Then $\sigma_{\text{ess}}(H) = (-m,m)^c$.

The proof of this lemma has been indicated in Ref. 2.

Remark: Under very general conditions which include all Coulomb-like potentials, Heinz[6] has shown that the separated Dirac Hamiltonian has an absolutely continuous spectrum in $[-m,m]^c$.

The lemma above shows that the eigenvalues of the Coulomb–Dirac Hamiltonian, $V_e = e/r$, $V_s = m$, and $V_e = f/r^2$ all lie in $[-m,m]$ with accumulation points $\pm m$ at most. Missing so far is a method to determine these eigenvalues. Since even the Coulomb–Dirac case cannot be solved in closed form, one has to resort to perturbation methods. The next results are devoted to these problems.

For spherically symmetric potentials it suffices to study the separated equations only. Since the perturbation with V_e and V_s is regular, the mathematically interesting and physically more important problem is the perturbation of H with V_m. Hence let $f \geqslant 0$ and write

$$K_f = \begin{pmatrix} V_e - V_s & D + l/r - fV_m \\ -D + l/r - fV_m & V_e + V_s \end{pmatrix}$$

for the separated operator. Then with $V_3 = fV_m - l/r$,

$$K_f^2 \geqslant V_3^2 - |V_3'| - 2|V_e V_3| \tag{7}$$

holds in the form sense on $\mathscr{C}^\infty(0,\infty) \otimes \mathbb{C}^2$. Thus the domain $\mathscr{D}_{\overline{K}_f}$ of $\overline{K}_f$ is independent of f and $\mathscr{D}_{\overline{K}_f} \subset D_{V_m}$ if $f \neq 0$.

If K_0 is essentially self-adjoint on $\mathscr{C}_0^\infty \otimes \mathbb{C}^2$ we even have strong resolvent convergence $K_f \to K_0$ under very weak assumptions. This, however, is not sufficient to obtain information on the behavior of the point spectrum for $f \to 0$, as is well known.

Even naive perturbation theory encounters considerable difficulties as the Coulomb–Dirac case shows.[7] In this case one has

$$|u(r)|, |v(r)| \sim r^\delta, \quad \text{with } \delta = (l^2 - e^2)^{1/2} \quad \text{for small } r,$$

for the eigenfunctions of K_0. Thus these eigenfunctions are in the domain (form domain) of $V_m = f/r^2$ only if $l^2 > \frac{9}{4} + e^2$ ($l^2 > \frac{1}{4} + e^2$).

For this reason Bethe uses in the case $l = 1$ the Pauli wave functions[7] instead of the proper Dirac eigenfunctions. In order to establish the validity of any form of perturbation theory, one has to show the stability of the eigenvalues first.

This means that the spectrum of K_f is a continuous function of f at $f = 0$.

In order to show this, we use the following comparison result, which is probably well known.

Lemma 2: (a) Assume W is locally square integrable on $(0,r_0)$ and satisfies $W(r) \geqslant 3/4r^2$. Then any square integrable solution y of $y'' = Wy$, $y(r_0) = 1$ satisfies the following.

(i) yy' is strictly monotonically increasing without zeros in $(0,r_0)$.

(ii) $\lim_{r\to 0} y(r) = \lim_{r\to 0} y'(r) = 0$.

(b) Assume $W_1(r) \geqslant W_2(r) \geqslant 3/4r^2$ and assume the W_i, $i = 1, 2$ to be locally square integrable. Let y_i be a square integrable solution of $W_i\ y_i = y_i''$ with $y_i(r_0) = 1$. Then $y_1(r) \leqslant y_2(r)$ on $(0,r_0)$.

Part (a) of this Lemma is shown by comparing y with solutions of $f''(r) = 3/4r^2 f(r)$. Part (b) is standard. In both

cases one has to manipulate $\int_r^{r_0}(W_1 - W_2)y_1 y_2$.

Theorem 2: Let V_e, V_s, and V_m as above. Moreover, assume $a + e \neq 0$ and

$$l - e^2 + a^2 > \tfrac{5}{2}, \quad \text{if} \quad l > 0,$$
$$l^2 - e^2 + a^2 > \tfrac{9}{4}, \quad \text{if} \quad l < 0.$$

Then $K_f \to K_0$ in strong resolvent convergence and the eigenvalues are stable.

The proof of this theorem is rather technical. Therefore only a sketch of important steps will be given.

Proof: (a) Let $K_f\begin{pmatrix} u \\ v \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \end{pmatrix}$. Then

$$\begin{aligned} v' &= (fV_m - l/r)v + V_s\, u - (V_e - \lambda)u, \\ u' &= -(fV_m - l/r)u + V_s\, v + (V_e - \lambda)v. \end{aligned} \tag{8}$$

Differentiation of v and v' leads with $fV_m - l/r \to V_m$ and $V_e - \lambda \to V_e$ to

$$u'' = \left[V_m^2 + V_s^2 - V_e^2 - V_m' + \frac{V_s' + V_e'}{V_s + V_e} V_m\right] u + \frac{V_s' + V_e'}{V_s + V_e} u',$$

provided $(V_s + V_e)(r) \neq 0$. This can always be achieved by restricting to a suitably small interval. This equation and $v = (V_s + V_e)^{-1}(u' + V_m u)$ are equivalent to the original system. The factorization

$$u = +[(V_e + V_s)]^{1/2}\psi_u$$

leads to

$$\psi_u'' = W_u \psi_u, \tag{9}$$

with

$$W_u = -\frac{(V_e + V_s)''}{2(V_e + V_s)} + \frac{3(V_e + V_s)'^2}{4(V_e + V_s)^2} + \left\{V_m^2 + V_s^2 - V_e^2 - V_m' + V_m \frac{(V_e + V_s)'}{V_e + V_s}\right\}.$$

By our assumptions, we can write

$$W_u = (fV_m - (l - \tfrac{1}{2})/r + U_1)^2 + 1/r^2(l - \tfrac{1}{2} - e^2 + a^2) + U_2,$$

with $\lim_{r\to 0} r\,U_1(r) = 0$ and $\lim_{r\to 0} r^2\,U_2(r) = 0$. In fact these limits may be chosen to be uniform in $0 \leqslant f \leqslant \delta$ for some $\delta > 0$ and $|\lambda| \leqslant m$. There exists $r_0, r_1 > 0$ such that for some $\epsilon > 0$ with $l - e^2 + a^2 - \epsilon > \frac{5}{2}$, if $l > 0$, resp. $l^2 - e^2 + a^2 - \epsilon > \frac{5}{2}$, if $l < 0$,

$$W_u(f,\lambda,r) \geqslant (l - e^2 + a^2 - \tfrac{1}{2} - \epsilon)/r^2, \quad 0 < r \leqslant r_0, \quad l > 0,$$

resp.

$$W_u(f,\lambda,r) \geqslant (l^2 - e^2 + a^2 - \epsilon - \tfrac{1}{4})/r^2, \quad 0 < r \leqslant r_1, \quad l < 0,$$

for $|\lambda| \leqslant m$ and $0 \leqslant f \leqslant \delta$.

In the remainder we shall only treat the case $l > 0$, since the proof for $l < 0$ is almost identical. With the aid of Lemma 2 one derives from this

$$|u(r)| \leqslant |u(r_0)|(r/r_0)^{\alpha_1}, \qquad 0 \leqslant r \leqslant r_0,$$
$$|u'(r)| \leqslant C_1|u(r_0)|(r/r_0)^{\alpha_1 - 1}, \qquad 0 \leqslant r \leqslant r_0,$$

where $\alpha_1 = (l - e^2 + a^2 - \frac{1}{4} - \epsilon)^{1/2}$ and where C_1 can be chosen independently of λ and f. This implies with (8)

$$|v(r)| \leqslant |u(r_0)|C_2(r/r_0)^{\alpha_1}, \quad 0 \leqslant r \leqslant r_0,$$

and

$$\left|\left|\frac{1}{r^2}\begin{pmatrix} u \\ v \end{pmatrix}\right|\right|^2_{[0,r_0]} \leqslant C_3 r_0^{-3}|u(r_0)|^2,$$

where C_2 and C_3 may be chosen independently of f, λ; $0 \leqslant f \leqslant \delta$ and $|\lambda| \leqslant m$.

(b) Since $\mathscr{D}_{K_0} \subset W_0^1(\mathbb{R}_+) \otimes \mathbb{C}^2$ there exist constants A, B with

$$\|u\|_\infty, \|v\|_\infty \leqslant A\left|\left|K_0\begin{pmatrix} u \\ v \end{pmatrix}\right|\right| + B\left|\left|\begin{pmatrix} u \\ v \end{pmatrix}\right|\right|, \quad \begin{pmatrix} u \\ v \end{pmatrix} \in \mathscr{D}_{K_0}.$$

For a normalized λ eigenfunction $\begin{pmatrix} u \\ v \end{pmatrix}$ of K_f this shows

$$\left|\left|K_0\begin{pmatrix} u \\ v \end{pmatrix}\right|\right| \leqslant \left|\left|K_f\begin{pmatrix} u \\ v \end{pmatrix}\right|\right| + f\left|\left|V_m\begin{pmatrix} u \\ v \end{pmatrix}\right|\right|_{[0,r_0]} + \sup_{r > r_0}|fV_m(r)|.$$

This and the first inequality finally yield, for sufficiently small $0 \leqslant f \leqslant \delta$ and $|\lambda| \leqslant m$,

$$\|u\|_\infty \|v\|_\infty \leqslant C_4.$$

(c) For any $\lambda_0 \in (-m,m)$ there exists $r_3 \geqslant r_0$ and $k > 0$ such that

$$W_u(f,\lambda,r) \geqslant k^2, \quad f \leqslant \delta$$

and $|\lambda - \lambda_0| < \epsilon$ for a suitable small $\epsilon > 0$. This implies as in (a) that any λ eigenfunction $\begin{pmatrix} u \\ v \end{pmatrix}$ of K_f satisfies

$$u(r), v(r) \sim e^{-kr}, \quad \text{for } r \geqslant r_3.$$

This together with (a) and (b) shows

$$\left|\left|\frac{1}{r^2}\begin{pmatrix} u \\ v \end{pmatrix}\right|\right| \leqslant C\left|\left|\begin{pmatrix} u \\ v \end{pmatrix}\right|\right|, \tag{10}$$

for any λ eigenfunction $\begin{pmatrix} u \\ v \end{pmatrix}$ of K_f; $|\lambda - \lambda_0| < \epsilon$, $0 \leqslant f \leqslant \delta$. We may choose C independently of f and λ.

(d) From (10) one can easily derive the stability of the eigenvalues $\lambda \in (-m,m)$ of K_0, because any normalized λ eigenfunction of K_f is a Cf-approximate λ eigenfunction for K_0.

Remark: Probably the condition $l^2 - e^2 + a^2 > \frac{5}{2}$ will suffice in general to prove stability. The case $l = 1$ however may lead to new bound states, which in general are in $[-m,m]$.

The case $e + a = 0$ can be treated with similar methods as above and leads to the condition $l > \frac{13}{2}$, respectively, $l^2 > \frac{13}{2}$.

The eigenvalue problem for K_0 with $U_e = U_s = 0$ can be solved explicitly. The eigenfunctions in this case are products of $r^\gamma\, e^{\beta r}$ and confluent hypergeometric functions, which degenerate to polynomials.

IV. PERTURBATION EXPANSION

Once one has shown stability of the eigenvalues it is not too difficult to compute these eigenvalues asymptotically by

a low-order perturbation theory. This is the content of the next theorem.

Let λ_0 be an eigenvalue of K_0 with corresponding normalized eigenvector x_0 and let $|\lambda-\lambda_0|\leqslant\delta$ with δ suitably small. Asymptotic perturbation theory is usually based on the formal expression

$$(K_f-\lambda)^{-1}=\sum_{n=0}^{N-1}(-f)^n(K_0-\lambda)^{-1}[V(K_0-\lambda)^{-1}]^n + (-f)^N(K_f-\lambda)^{-1}[V(K_0-\lambda)^{-1}]^N \tag{11}$$

(see Ref. 8 and Ref. 9, Chap. VIII). We say that K_f has an asymptotic expansion near λ_0 of order N, if (11) can be applied termwise to x_0.

Theorem 3: Assume $\gamma=(l^2+a^2-e^2)^{1/2}>N+\frac{1}{2}$ and assume the conditions of Theorem 2 hold. Then K_f has an asymptotic expansion of order N near each eigenvalue $\lambda_0\in(-m,m)$ of K_0.

The proof is based on the following lemma, which states roughly that $V(K_0-\lambda)^{-1}$ increases the singularity of the vector to which it is applied by r^{-1}. Thus

$$[V(K_0-\lambda)^{-1}]^N x_0\sim r^{\gamma-1-N}$$

near 0, because the first application of this term increases the singularity by r^{-2}.

Lemma 3: Let $\gamma-\frac{1}{2}>n$ and let $0<\alpha<\operatorname{Re}\beta$, where $\beta=(m^2-\lambda^2)^{1/2}$. Moreover let $f_{n,\alpha}$ be a smooth function which agrees with r^n on $[0,1]$ with $e^{-\alpha r}$ on $[2,\infty)$ and satisfies $f_{n,\alpha}\geqslant e^{-2\alpha}$ on $[1,2]$. Denote by $M_{n,\alpha}$ the operator multiplication by $f_{n,\alpha}$, then, for any $\epsilon>0$,

$$M^{-1}_{n+1-\epsilon,\alpha-\epsilon}(\lambda-K_0)^{-1}M_{n,\alpha}$$

is a bounded operator on $\mathscr{L}^2(0,\infty)\otimes\mathbb{C}^2$.

Proof: (a) Since K_0 is essentially self-adjoint, there are essentially unique solutions u_0 (resp. u_∞) of $(K_0-\lambda)u_{0,\infty}=0$, which are square integrable near 0 (resp. ∞). These solutions satisfy the asymptotic conditions

$$\begin{aligned}
&u_{0,i}(r)\sim r^\gamma g_i(r), && r \text{ small}, && i=1,2,\\
&u_{\infty,i}(r)\sim r^{-\gamma}\tilde{g}_i(r), && r \text{ small}, && i=1,2,\\
&u_{0,i}(r)\sim e^{\beta r}h_i(r), && r \text{ large}, && i=1,2,\\
&u_{\infty,i}(r)\sim e^{-\beta r}\tilde{h}_i(r), && r \text{ large}, && i=1,2.
\end{aligned}$$

The comparison lemma shows that g_i and $\tilde{g}_i$ can be estimated by $r^{-\epsilon}$ for any $\epsilon>0$ for small r. Similarly h_i and $\tilde{h}_i$ can be estimated by $e^{\epsilon r}$ and large r.

(b) It follows[10] that $(\lambda-K_0)^{-1}$ is an integral operator with kernel

$$K_{i,j}(r,s)=C\chi_{[0,r]}(s)u_{\infty,i}(r)\bar{u}_{0,j}(s) + C\chi_{[r,\infty)}(s)u_{0,i}(r)\bar{u}_{\infty,j}(s).$$

Then $M_{n+1-\epsilon,\alpha-\epsilon}(\lambda-K_0)^{-1}M_{n,\alpha}$ is likewise an integral operator with kernel

$$\tilde{K}_{ij}(r,s)=f^{-1}_{n+1-\epsilon,\alpha-\epsilon}(r)K_{i,j}(r,s)f_{n,\alpha}(s).$$

Its boundedness is shown by the Schur test[11] with test function ϕ

$$\phi(r)=\begin{cases}r^{-1/2} & 0<r\leqslant 1,\\ 1 & r\geqslant 1.\end{cases}$$

In the necessary computations for this one has to use that g, $\tilde{g}$, h, and $\tilde{h}$ can be estimated as stated above.

This lemma shows

$$x\in\mathscr{D}_{M^{-1}_{n,\alpha}} \quad\text{then}\quad (\lambda-K_0)^{-1}x\in\mathscr{D}_{M^{-1}_{n+1-\epsilon,\alpha-\epsilon}}.$$

Thus it permits the control of the asymptotic behavior of $(\lambda-K_0)^{-1}x$ in terms of x.

Proof of Theorem 3: For $0\leqslant n\leqslant N$, let $x_n=[V(K_0-\lambda)^{-1}]^n x_0$. By induction on n one shows $x_n\in\mathscr{D}_{M^{-1}_{\gamma-1/2-n(1+\epsilon),\alpha-n\epsilon}}$, for any $0<\alpha<\operatorname{Re}\beta$ and $\epsilon>0$ suitably small. From this the theorem follows immediately.

Remark: The lemma actually shows more than one needs for the proof of the theorem, and probably one can push the resolvent estimates higher by a factor of $f^{1/2}$ by using form techniques as in Ref. 9, Chap. VIII, Sec. 4.3.[11]

In general the computation of the spectrum of K_f will be rather complicated, because V_m is singular at 0. For the neutral particle, however, a complete solution can be given.

Theorem 4: Assume $V_s=V_e=0$ and that V_m is differentiable and more singular than r^{-1} at 0. Then $\sigma(K_f)=(-m,m)^c$.

Proof: Assume $\lambda\in\sigma(K_f)$ satisfies $\lambda^2<m^2$. Then (9) becomes

$$(-D^2+V_m^2-V_m'+m^2-\lambda^2)u=0.$$

With $F(r)=\exp(\int_r^{r_0}V_m(t)dt)$ this can be written as

$$\{-(F^{-1}DF)(FDF^{-1})+(m^2-\lambda^2)\}u=0.$$

Since the first summand is positive, this implies $m^2=\lambda^2$.

If, in addition to the above conditions, V_m is monotonic for large r, the only eigenvalues of K_f can be $\pm m$.

[1] H. Behncke, "The Dirac equation with anomalous magnetic moment," Math. Z. **174**, 213 (1980).

[2] H. Behncke, "The Dirac equation with anomalous magnetic moment. II" in *Proceedings of the 7th Conference on Ordinary and Partial Differential Equations, Lecture Notes in Math., Vol. 964* (Springer, Berlin, 1982).

[3] F. Gesztesy, B. Simon, and B. Thaller, "On the self-adjointness of Dirac operators with anomalous magnetic moment," preprint.

[4] K. Jörgens, "Perturbations of the Dirac operator," in *Conference on the Theory of Ordinary and Partial Differential Equations* (Dundee, 1972), pp. 87–102; *Lecture Notes in Mathematics, Vol. 280* (Springer, Berlin, 1972).

[5] B. Karnarski, "Generalized Dirac operators with several singularities," J. Op. Theory **13**, 171 (1985).

[6] E. Heinz, "Ober das absolut stetige Spektrum singulärer Differentialgleichungssysteme," Nachr. Akad. Wiss. Göttingen **1**, 1 (1982).

[7] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Problems* (Springer, Berlin, 1957), Sec. 19.

[8] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1978), Vol. IV, p. 40.

[9] T. Kato, *Perturbation Theory for Linear Operators* (Springer, Berlin, 1966).

[10] J. Weidmann, "Oszillationsmethoden für Systeme gewöhn-licher Differentialgleichungen," Math. Z. **119**, 349 (1971) Theorem 1.6.

[11] G. O. Okikiolu, *Aspects of the Theory of Bounded Integral Operators in Lp Spaces* (Academic, New York, 1971), Theorem 4.1.2.

The relation between the (N)- and ($N-1$)-electron atomic ground states

P. H. Briet
Centre de Physique Théorique, Centre National de la Recherche Scientifique, Luminy Case 907, F-13288 Marseille Cedex 9, France and PHYMAT-Université de Toulon et du Var, La Garde, France

(Received 17 July 1984; accepted for publication 29 March 1985)

The relations between the ground states of an N- and $(N-1)$-electron atomic system are studied. It is shown that in some directions of the configuration space, the ratio of the N-electron atomic ground state to the one-particle density is asymptotically equivalent to the $(N-1)$-electron atomic ground state.

I. INTRODUCTION

We consider the Schrödinger operator associated to an atomic system with N electrons and infinitely heavy nucleus with charge Z

$$H^N = -\sum_{i=1}^{N} \Delta_i - \sum_{i=1}^{N} \frac{Z}{|x_i|} + \sum_{i\neq j}^{N} \frac{1}{|x_i - x_j|}. \tag{1.1}$$

In the following we will not take into account the Fermi statistics of the electrons and consider only the mathematical ground state φ_0^N of H^N which is a positive function, symmetric in the electron coordinates $x = (x_1, x_2, \ldots, x_N)$. The asymptotic behavior of the ground state φ_0^N associated to H^N has been studied by many authors.[1–4] Recently Agmon[5] and Carmona–Simon[6] obtained the following bounds:

$$\forall\, \delta > 0, \quad \forall\, x \in \mathbb{R}^{3N},$$
$$Ce^{-(1+\delta)\gamma(x)} \leqslant \varphi_0^N(x) \leqslant De^{-(1-\delta)\gamma(x)}. \tag{1.2}$$

Here $\gamma(x)$ is the Agmon metric associated to (1.1); it gives a nonisotropic asymptotic behavior for φ_0^N depending on the direction along which x goes to infinity in the configuration space of electrons. Along this line of result, we want to show that for large $|x_1|$ one can obtain φ_0^N as a product of the $(N-1)$-particle ground state φ_0^{N-1} and the one-particle density

$$\rho^{1/2}(x_1) = \left(\int |\varphi_0^N|^2 d_x^{3(N-1)} \right)^{1/2}.$$

Since our analysis only works if both φ_0^{N-1} and φ_0^N are square integrable, we will assume in the following that the existence of an N-electron mathematical ground state implies the same for the $(N-1)$-electron system (although to our knowledge this is still an open problem).

More precisely, we prove the conjecture of Morgan and Hoffman–Ostenhof

$$\lim_{|x_1|\to\infty} \|\varphi_0^N/\rho^{1/2} - \varphi_0^{N-1}\|_{L^2(\mathbb{R}^{3(N-1)})} = 0. \tag{1.3}$$

We shall show also a stronger version of (1.3)

$$\lim_{|x_1|\to\infty} (\varphi_0^N/\rho^{1/2}, (H^{N-1} - E_0^{N-1})\varphi_0^N/\rho^{1/2}) = 0. \tag{1.4}$$

The formulas (1.3) and (1.4) were obtained by Combes–Hoffman-Ostenhof–Hoffman-Ostenhof[7] for heliumlike systems. For N-body systems Lieb–Simon[8] proved (1.3) in the special case where the potentials have compact support.

To prove (1.3) and (1.4) we extend the Combes–Hoffman-Ostenhof–Hoffman-Ostenhof method with the help of the bounds (1.2). An essential ingredient of this approach is the fundamental density associated to φ_0^N

$$u(x_1) = (\varphi_0^N, \varphi_0^{N-1})_{L^2(\mathbb{R}^{3(N-1)})}(x_1).$$

The function u possesses some interesting properties for our problem: on the one hand u is a lower bound asymptotically equal to $\rho^{1/2}$ and on the other hand u is the solution of a one-particle Schrödinger equation with a potential Ve of the form

$$Ve(|x_1|) \simeq [Z - (N-1)]/|x_1|, \quad \text{for large } |x_1|. \tag{1.5}$$

This provides us the detailed information required for the proof of (1.3) and (1.4). This proof involves some slight modifications of the complex boost method as used in Refs. 1 and 3. As in this last paper we will also use the so-called geometrical techniques in the operator form developed, e.g., in Refs. 2, 9, and 10.

In Sec. II, we recall the results obtained in Ref. 7 and adapt them to our problem: properties on function u are described in Sec. III. Section IV is devoted to the spectral analysis of operators that we shall need in Sec. IV to prove the formulas (1.3) and (1.4).

II. PRELIMINARIES–RESULTS

The proof of formulas (1.3) and (1.4) requires the generalization of some results obtained by Combes–Hoffman-Ostenhof–Hoffman-Ostenhof,[7] namely, Theorem 2.1 below. For its proof we need no more supplementary material than in Ref. 7 and we shall describe here only the relevant part.

Let P be the orthogonal projection operator on $L^2(\mathbb{R}^{3N})$ defined by

$$\forall\, \psi \in L^2(\mathbb{R}^{3N}),$$
$$(P\psi)(x) = (\psi, \varphi_0^{N-1})_{L^2(\mathbb{R}^{3(N-1)})}(x_1)\varphi_0^{N-1}(x_2 \cdots x_N) \tag{2.1}$$

and $Q = \mathbb{1} - P$.

If $\psi = \varphi_0^N$ we set $u = (\varphi_0^N, \varphi_0^{N-1})_{L^2(\mathbb{R}^{3(N-1)})}$. From its definitions and properties of φ_0^N (Ref. 10), u is continuous, strictly positive, and spherically symmetric.

For $\alpha \in \mathbb{R}$ consider the unitary group on x defined by

$$\forall\, \psi \in L^2(\mathbb{R}^{3N}), \quad (V(\alpha)\psi)(x) = (e^{-i\alpha \ln(u)}\psi)(x). \tag{2.2}$$

One has

$$H^N(\alpha) = V(\alpha)H^N V_{(\alpha)}^{-1}$$
$$= \left(i\nabla_1 + \alpha\frac{\nabla u}{u}\right)^2 - \sum_{i=2}^{N} \Delta_i - \sum_{i=1}^{N} \frac{Z}{|x_i|} + \sum_{i\neq j}^{N} \frac{1}{|x_i - x_j|}. \tag{2.3}$$

Under the conditions (*) on u (cf. Sec. III, Remark 3.3),

 0022-2488/85/102560-05$02.50

$\{H^N(\alpha);\ \alpha\in\mathbb{R}\}$ can be analytically continued to the whole complex plane into a self-adjoint holomorphic family of type A.

Let $W^{2,2}(\mathbb{R}^{3N})$ denote the usual Sobolev space, and Ω the unit ball in $\mathbb{C}$; we have the following result.

Theorem 2.1: $V(\alpha)Q\varphi_0^N = u^{-i\alpha}Q\varphi_0^N \in W^{2,2}(\mathbb{R}^{3N})$ as long as $E_0^N \notin \sigma_{ess}(H^N(\alpha)+\delta P)$ for some $\delta>0$ and $\alpha\in\Omega$.

Sketch of proof: By using the well-known partitioning technique[7] one gets

$$\forall\,\alpha\in\mathbb{R},\quad u^{-i\alpha}Q\varphi_0^N = (E_0^N - QH^N(\alpha)Q)^{-1} \times\left(\sum_{j>1}^{N} Q\frac{1}{|x_1-x_j|}u^{1-i\alpha}\varphi_0^{N-1}\right). \tag{2.4}$$

Now it is sufficient to show that under the conditions of Theorem 2.1, the right-hand side (rhs) of 2.4 has an $L^2(\mathbb{R}^{3N})$-continuous continuation in Ω. But this is true if $(E_0^N - QH^N(\alpha)Q)^{-1}$ is a bounded family of operators in Ω, or, in other words, if $E_0^N \notin \sigma(QH^N(\alpha)Q)$.

On the one hand $E_0^N \notin \sigma_{ess}(H^N(\alpha)+\delta P)$ by assumption; then $E_0^N \notin \sigma_{ess}(PH^N(\alpha)P+\delta P+QH^N(\alpha)Q)$ since

$$\sum_{j>1} P\frac{1}{|x_1-x_j|}Q + Q\frac{1}{|x_1-x_j|}P$$

is $H^N(\alpha)$ compact.[11]

On the other hand, by some standard analyticity arguments[1] $E_0^N \notin \sigma_d(PH^N(\alpha)P+\delta P+QH^N(\alpha)Q)$, if $E_0^N \notin \sigma_d \times(PH^NP+\delta P+QH^NQ)$.

Suppose that $E_0^N \in \sigma_d(PH^NP+\delta P)$; then $\inf_\psi(\psi, H^N\psi) \geqslant E_0^N+\delta$. This is a contradiction. Now if $E_0^N \in \sigma_d(QH^NQ)$, then $P\varphi_0^N = 0$ and $u=0$, which contradicts the positivity of u. We conclude $E_0^N \notin \sigma(PH^N(\alpha)P + QH^N(\alpha)Q+\delta P)$; therefore $E_0^N \notin \sigma(QH^N(\alpha)Q)$.

To complete this proof, we must show that

$$\sum_{j>1} Q\frac{u^{1-i\alpha}}{|x_1-x_j|}\varphi_0^N$$

is a family of $L^2(\mathbb{R}^{3N})$ vectors continuous in Ω, which can be seen in Ref. 7.

III. THE FUNDAMENTAL DENSITY ASSOCIATED TO φ_0^N

The analysis of the essential spectrum of the operator family $\{H^N(\alpha);\alpha\in\mathbb{C}\}$ is a subtle point because the operators are non-self-adjoint. It will require some results on the fundamental density u and in particular on the asymptotic behavior of $\nabla u/u$. This is the purpose of this present section.

The study of u is based on the following one-particle Schrödinger equation[4,8]:

$$[-\Delta_1 - z/|x_1| + (N-1)V_e]u = -\epsilon u. \tag{3.1}$$

Here ϵ is the ionization energy $\epsilon = E_0^{N-1} - E_0^N$ and $V_e(|x_1|)$ is a screened potential

$$V_e(|x_1|) = u^{-1}(|x_1|)\int \frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|}d^{3(N-1)}y. \tag{3.2}$$

Equation (3.1) is easily obtained from the identity $([H^N - E_0^N]\,\varphi_0^N, \varphi_0^N)_{L^2(\mathbb{R}^{3(N-1)})} = 0$. The "effective" potential $V_e(|x_1|)$ is a continuous, strictly positive, and spherically symmetric function, and we have the following.

Lemma 3.1: $V_e(|x_1|)$ satisfies the pointwise estimate

$$V_e(|x_1|)\leqslant C(N-1)/|x_1| + O(|x_1|e^{-\alpha|x_1|}),\quad \text{for } C>1 \text{ and } \alpha>0. \tag{3.3}$$

Proof: Let us recall that u is a bounded function; more precisely we have[8]

$$\forall\delta>0,\quad \exists C, D>0,\quad Ce^{-(1+\delta)\epsilon^{1/2}|x_1|}\leqslant u(x_1)\leqslant De^{-(1-\delta)\epsilon^{1/2}|x_1|}. \tag{3.4}$$

For fixed a, $1>a>0$, we split the integral

$$\int d^{3(N-1)}y\frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|} = \int_{|x_j|<a|x_1|}\frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|} + \int_{|x_j|>a|x_1|} d^{3(N-1)}y\frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|}. \tag{3.5}$$

The first part is bounded by

$$\int_{|x_j|<a|x_1|} d^{3(N-1)}y\frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|}\leqslant\frac{1}{|1-a|}\frac{u(x_1)}{|x_1|}. \tag{3.6}$$

By using the explicit bounds on φ_0^N, φ_0^{N-1} (see 1.2), respectively, we now get

$$\int_{|x_j|>a|x_1|} d^{3(N-1)}y\frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|} \leqslant C_\gamma^{-(1-\gamma)\epsilon^{1/2}|x_1|}\int_{|x_j|>a|x_1|} d^{3(N-1)}y\frac{e^{-\beta|y|}}{|x_j-x_1|}, \tag{3.7}$$

for $C_\gamma>0$ and some $\beta>0$, and with (3.4), (3.6) gives

$$\int_{|x_j|>a|x_1|} d^{3(N-1)}y\frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|} \leqslant C_\gamma e^{2\gamma|x_1|}u(x_1)\int_{|x_j|>a|x_1|} d^{3(N-1)}y\frac{e^{-\beta|y|}}{|x_j-x_1|}. \tag{3.8}$$

Now the integration of the last integral and the choice of suitable γ (3.7) lead to

$$\int_{|x_j|>a|x_1|} d^{3(N-1)}y\frac{\varphi_0^N\varphi_0^{N-1}}{|x_1-x_j|} \leqslant \text{const}(|x_1| + 1/|x_1| + 1)\,e^{-\alpha|x_1|}u(x_1), \tag{3.9}$$

for some $\alpha>0$.

Finally, by lumping together (3.5) and (3.8) we obtain (3.3).

Remark 3.1: In fact, we expect the following, more precise result:

$$V_e(|x_1|)\simeq(N-1)/|x_1| + O(1/|x_1|^\beta)\quad \text{for some } \beta>1. \tag{3.10}$$

This estimate is more difficult to obtain, but not necessary for our purpose.

We can state now the main result of this section.

Theorem 3.1:

$$\lim_{|x_1|\to\infty}(\nabla u/u) = -\epsilon^{1/2}(x_1/|x_1|). \tag{3.11}$$

Proof: By using the following transformation $f = -1/|x_1| + u'/u$ in (3.1), one gets

$$f' - f^2 + \epsilon + W_e = 0, \tag{3.12}$$

where

$$W_e(|x_1|) = -z/|x_1| + V_e(|x_1|). \tag{3.13}$$

Equation (3.12) is a Riccati equation, which we can analyze as follows: first we assume that $W_e \equiv 0$; then the solutions of (3.12) are

$$f = \epsilon^{1/2} + 2\epsilon^{1/2}\left[Ce^{-2\epsilon^{1/2}|x_1|} - 1\right], \quad \forall C \geqslant 0. \tag{3.14}$$

Equation (3.14) shows the existence of only one solution of (3.12) such that $f \to -\epsilon^{1/2}$ as $|x_1| \to \infty$; otherwise ($C \neq 0$), $f \to \epsilon^{1/2}$ as $|x_1| \to \infty$. We are going to show that this asymptotic behavior is stable under the perturbation W_e given by (3.13).

From general theorems on the stability of asymptotic behavior for nonlinear differential equations,[12] there exists some solution of (3.12) such that

$$\lim_{|x_1| \to \infty} f_0 = \epsilon^{1/2}. \tag{3.15}$$

Let f_0 be a solution of (3.12) satisfying (3.15); the linearization method for the Riccati equation gives the general solution of (3.12)

$$f = f_0 + \psi^{-2}\left[C - \int_{|x_1|}^{\infty} dt\, \psi^{-2}(t)\right]^{-1}, \quad \forall C \geqslant 0, \tag{3.16}$$

with

$$\psi(|x_1|) = \exp\left[\int_b^{|x_1|} dt f_0(t)\right], \quad \text{for some } b > 0. \tag{3.17}$$

According to the fact that $f_0'/f_0^2 \to 0$ as $|x_1| \to \infty$, the integration by parts leads to

$$f = f_0 + \psi^{-2}\left[C - [(1 + K(|x_1|))\, 2f_0]^{-1}\psi^{-2}\right]^{-1}, \tag{3.18}$$

where K is a function such that $K(|x_1|) \to 0$ as $|x_1| \to \infty$. This last expression shows that there exists only one solution such that $f \to -\epsilon^{1/2}$ as corresponding to $C = 0$, whereas for $C \neq 0$ the limit is $\epsilon^{1/2}$.

Now, since $u(|x_1|)$ is a nonincreasing function, $u'/u \to -\epsilon^{1/2}$, as the $|x_1| \to \infty$.

We close this section with remarks about the regularity of function u.

Remark 3.2:(i) Let S_0 be the domain $S_0 = \{z \in \mathbb{C} \setminus |\mathrm{Arg}(z)| < \pi/2\}$, φ_0^N, φ_0^{N-1} being analytic vectors for the group of dilatation in S_0. u has an analytic continuation in S_0 (Refs. 13 and 14).

(ii) by using the explicit kernel of $(-\Delta_r + 1)^{-1}$, where Δ_r is the radial Laplacian operator, and (3.1), one can show that u has bounded first radial derivative.[15]

Remark 3.3 [conditions (*)]: From Remark 3.2(i) and Theorem 3.1, one has the following.

(i) $\nabla u/u = u'/u \cdot x_1/|x_1|$ is bounded on $\mathbb{R}^3$.

In the other hand, it is not difficult to show the following from Eq. (3.1) and Theorem 3.1.

(ii) $\nabla(\nabla u/u) \in L^2_{\mathrm{LOC}}(\mathbb{R}^3) + L^{\infty}_{\epsilon}(\mathbb{R}^3)$.

IV. SPECTRAL ANALYSIS OF THE OPERATOR FAMILY $\{H^N(\alpha); \alpha \in \mathbb{C}\}$

This analysis is an extension of the geometrical method used in the proof of the HVZ theorem[2,9] to non-self-adjoint operators with N-body interaction.

Let us start with some definitions. In the following the fixed nucleus will be the particle labeled by "0." Let $D = \{(C_1, C_2)\}$ be the set of two-cluster decompositions; for each D we define the operator

$$H_D(\alpha) = H^N(\alpha) + \delta P - V_D, \tag{4.1}$$

where V_D is the intercluster potential and P is considered as an interaction between particles "0,2,...,N"; so P is included in V_D unless $D = (C_1, C_2)$ with $C_1 = \{1\}$, $C_2 = \{0,2,...,N\}$. The Simon partition of unity[2] is a collection of functions $J_0 \in C^{\infty}$ homogeneous of degree zero for $|x_1| > 1$, such that

$$\sum_{0 = (C_1, C_2)} J_D^2 = 1, \tag{4.2a}$$

$$\mathrm{Supp}(J_0) \subseteq \{x \in \mathbb{R}^{3N} \setminus |x_i - x_j| \geqslant d\,|x|, \quad \text{for } i \in C_1, j \in C_2 \text{ and } d > 0\}. \tag{4.2b}$$

According to these properties of J_D, one can easily show that $|\nabla J_D|$ and $J_D/|x_i - x_j|$ for i, j belonging to different clusters of D are H_0^N compact (in fact, by using the definition of J_D, it is easy to see that these functions decay as $|x|^{-1}$).

With the partition given by the J_D's we can write a localization formula for the resolvent operators[14]

$$(H^N(\alpha) + \delta P - Z)^{-1}(1 + K(Z)) = \sum_D J_D\,(H_D(\alpha) - Z)^{-1} J_D, \tag{4.3}$$

with

$$K(Z) = \sum_D [H^N(\alpha) + \delta P) J_D - J_D H_D(\alpha)] \times (H_D(\alpha) - Z)^{-1} J_D. \tag{4.4}$$

Actually, $K(Z)$ is a sum of compact operators. For the coulombic potentials and the terms $[H_0^N, J_D]$ this comes from the properties of J_D stated above. We also have to consider the term $[P, J_D]$ when $D = (C_1, C_2)$; $C_1 = \{1\}$ and $C_2 = \{0,2,...,N\}$. One has

$$[P, J_D] = P(J_D - 1) + (J_D - 1)P \tag{4.5}$$

and

$$P(J_D - 1)(H_0^N + 1)^{-1} = P\left(1 + \sum_{j=2}^{N} |x_j|\right)(1 - J_D)\left(1 + \sum_{j=2}^{N} |x_j|\right)^{-1}(H_0^N + 1)^{-1}. \tag{4.6}$$

Here,

$$(1 - J_D)\left(1 + \sum_{j=2}^{N} |x_j|\right)^{-1}(H_0^N + 1)^{-1}$$

is compact because on the support of $(1 - J_D)$, there is at least $j \neq 1$ such that $|x_j| > d\,|x|$. The exponential decay of φ_0^N implies that $P(1 + \Sigma_{j=2}^N |x_j|)$ is bounded so that $P(J_D - 1)(H_0^N + 1)^{-1}$ is compact. By standard arguments the same is true for $(J_D - 1)P(H_0^N + 1)^{-1}$.

Note that an H_0^N-compact operator is also $H_D(\alpha)$ compact since $H_D(\alpha)$ is obtained from H_0^N by adding relatively bounded perturbations with arbitrarily small bound.

Now the sectoriality of $H_D(\alpha)$, $\forall D$ implies there exists some Z such that $(1+K(Z))$ has a bounded inverse; hence by the analytic Fredholm theorem,[16] $[1+K(Z)]^{-1}$ is a meromorphic family in $\mathbb{C}$. By (4.3), $(H^N(\alpha)+\delta P-Z)^{-1}$ also is meromorphic if $Z\in\cap_D \rho(H_D(\alpha))$. [Here, $\rho(A)$ is the resolvent set for A.] Therefore, one has

$$\sigma_{\rm ess}(H^N(\alpha)+\delta P)\subset \cup_D \sigma(H_D(\alpha)). \tag{4.7}$$

Remark 4.1: The poles of $(H^N(\alpha)+\delta P-Z)^{-1}$ are the discrete eigenvalues of $H^N(\alpha)+\delta P$.

For the purpose of the inductive proof of Theorem 4.1 below, we need to consider separately the transformed family of P-electron operators

$$H(\alpha)=\left(i\nabla_1+\alpha\frac{\nabla u}{u}\right)^2-\sum_{i=2}^{P}\Delta_i+\sum_{i\neq j}^{P}\frac{1}{|x_i-x_j|}. \tag{4.8}$$

It corresponds to a cluster Hamiltonian $H_D(\alpha)$, for $N=P$ and $D=(C_1,C_2)$ with $C_1=\{0\}$ and $C_2=\{1,2,...,N\}$. The essential spectrum of $H(\alpha)$ is given by the following.

Lemma 4.1:

$$\sigma_{\rm ess}(H(\alpha))\subset\{z\in\mathbb{C}\backslash {\rm Re}(z)\geqslant -({\rm Im}\,\alpha)^2\epsilon\},$$

where $\epsilon=E_0^{N-1}-E_0^N$.

Proof: One has

$$H(\alpha)=\left(i\nabla_1+\alpha\epsilon^{1/2}\frac{x_1}{|x_1|}\right)^2-\sum_{i=2}^{P}\Delta_i+\sum_{i\neq j}^{P}\frac{1}{|x_i-x_j|}+V_{01}, \tag{4.9}$$

with

$$V_{01}=(i\nabla_1+\alpha(\nabla u/u))^2-(i\nabla_1-\alpha\epsilon^{1/2}(x_1/|x_1|)^2). \tag{4.10}$$

Here, V_{01} is an interaction between $\{1\}$ and a fictitious fixed source which is labeled by "0"; by Theorem 3.1 V_{01} is Δ_1 compact. Then

$$\sigma_{\rm ess}(H(\alpha))\subset\{z\in\mathbb{C}\backslash {\rm Re}(z)\geqslant -({\rm Im}\,\alpha)^2\epsilon\}. \tag{4.11}$$

We prove this inclusion by an HVZ analysis along the lines given above. Let us consider, for example, the Hamiltonian $\tilde{H}_D(\alpha)$ with $D=(C_1,C_2)$, $C_1=\{0\}$, and $C_2=\{1,2,...,P\}$. We have for the corresponding cluster Hamiltonian

$$\tilde{H}_D(\alpha)=\left(i\nabla_1+\alpha\epsilon^{1/2}\frac{x_1}{|x_1|}\right)^2-\sum_{i=2}^{P}\Delta_i+\sum_{i\neq j}^{P}\frac{1}{|x_i-x_j|}. \tag{4.12}$$

By Sigal's theorem[9] one gets

$$\begin{aligned}&\inf({\rm Re}(\sigma_{\rm ess}(\tilde{H}_D(\alpha))))\\ &\quad\geqslant\inf\left(\sigma_{ess}\left(-\sum_{i=1}^{P}\Delta_i+\sum_{i\neq j}^{P}\frac{1}{|x_i-x_j|}-({\rm Im}\,\alpha)^2\epsilon\right)\right)\\ &\quad\geqslant -({\rm Im}\,\alpha)^2\epsilon.\end{aligned} \tag{4.13}$$

Therefore

$$\sigma_{\rm ess}(\tilde{H}_D)\subset\{z\in\mathbb{C}\backslash {\rm Re}(z)\geqslant -({\rm Im}\,\alpha)^2\epsilon\}. \tag{4.14}$$

Finally, by standard arguments,[13] $\tilde{H}_D(\alpha)$ has no eigenvalues so that $\sigma(\tilde{H}_D(\alpha))=\sigma_{\rm ess}(\tilde{H}_D(\alpha))$. The other cases follow from this one by induction. Now we are ready to state the main result of this section, namely, $\sigma_{\rm ess}(H^N(\alpha)+\delta P)$ is included in the half-plane ${\rm Re}(z)\geqslant\Sigma-({\rm Im}\,\alpha)^2\epsilon$, where Σ is the smallest threshold of the system.

Theorem 4.1:

$$\begin{aligned}&\sigma_{\rm ess}(H^N(\alpha)+\delta P)\\ &\quad\subset\{z\in\mathbb{C}\backslash {\rm Re}(z)\geqslant E_0^{N-1}+\delta-({\rm Im}\,\alpha)^2\epsilon\},\\ &\quad 0<\delta<E_1^{N-1}-E_0^{N-1}.\end{aligned} \tag{4.15}$$

By induction, assume that Theorem 4.1 holds for all subsystems of less than $(N+1)$ particles. Assuming particle 1 is in C_1, one has for $D=(C_1,C_2)$

$$H_D(\alpha)=H_{C_1}(\alpha)\otimes 1+1\otimes H_{C_2}, \tag{4.16}$$

and by the Ichinoze lemma[10]

$$\sigma(H_D(\alpha))=\sigma(H_{C_1}(\alpha))+\sigma(H_{C_2}). \tag{4.17}$$

We have two different types of cluster decomposition. First the nucleus is in C_1; then, if N_{C_1} (N_{C_2}) denotes the number of particles in C_1 (C_2), one has by the induction hypothesis

$$\sigma_{\rm ess}(H_{C_1}(\alpha))\subset\{z\in\mathbb{C}\backslash {\rm Re}(z)\geqslant E_0^{N_{C_1}-1}-({\rm Im}\,\alpha)^2\epsilon\}, \tag{4.18}$$

and by some analyticity arguments mentioned above the isolated eigenvalues of $H_{C_1}(\alpha)$ remain independent of α as long as they are not adsorbed in the essential spectrum. With H_{C_2} being a positive operator, for such a D one gets

$$\sigma(H_D(\alpha))\subset\{z\in\mathbb{C}\backslash {\rm Re}(z)\geqslant \min(E_0^{N_{C_1}},E_0^{N_{C_1}-1}-({\rm Im}\,\alpha)^2\epsilon)\}. \tag{4.19}$$

Now, if the nucleus is in C_2, then H_{C_2} is the Hamiltonian of an atomic system with N_{C_2} electrons, and its spectrum is given by

$$\sigma(H_{C_2})=[E_0^{N_{C_2}},E_1^{N_{C_1}},...)\cup[E_0^{N_{C_2}-1},+\infty). \tag{4.20}$$

By Lemma 4.1

$$\sigma(H_{C_1}(\alpha))\subset\{z\in\mathbb{C}\backslash {\rm Re}(z)\geqslant -({\rm Im}\,\alpha)^2\epsilon\}. \tag{4.21}$$

Therefore

$$\sigma(H_D(\alpha))\subset\{z\in\mathbb{C}\backslash {\rm Re}\,z\geqslant E_0^{N_{C_2}}-({\rm Im}\,\alpha)^2\epsilon\}. \tag{4.22}$$

Finally, under the condition on δ, the minimum over D gives (4.15).

Remark 4.2: A complete characterization of $\sigma_{\rm ess}\times(H^N(\alpha)+\delta P)$ is given in Ref. 15. It consists of a union of paraboloids centered at various thresholds of the system, but (4.15) is sufficient for our investigations.

V. MAIN RESULTS

Combining Theorems 2.1 and 4.1 we have

$$u^{-i\alpha}Q\varphi_0^N\in W^{2,2}(\mathbb{R}^{3N}),\quad \forall\alpha\in\Omega. \tag{5.1}$$

From this we get the following lemma.

Lemma 5.1:

$$\tau(|x_1|)=\|Q\varphi_0^N/u\|_{L^2(\mathbb{R}^{3(N-1)})}(|x_1|)\leqslant C/|x_1|, \tag{5.2}$$

for some $C>0$ and $|x_1|\geqslant a>0$.

Proof: Let

$$\mathscr{B}(|x_1|)=\{r\in\mathbb{R}^+\backslash |r-|x_1||\leqslant 1\} \tag{5.3}$$

and $\varphi \in W^{2,1}(\mathscr{B}(|x_1|);\ dr)$ ($W^{2,1}$ denotes the usual Sobelev space). By the Sobelev embedding theorem[17] one gets

$$|\varphi(|x_1|)| \leqslant D \|\varphi\|_{W^{2,1}(\mathscr{B}(|x_1|))},$$

for some $D>0$. (5.4)

Setting $\varphi(|x_1|) = |x_1|\tau^2(|x_1|)$ one has

$$||x_1|\tau^2(|x_1|)| \leqslant D \int_{\mathscr{B}(|x_1|)} \left(\left| \frac{d^2 r\, \tau^2}{dr^2} \right| + |r\tau^2| \right) dr. \quad (5.5)$$

Then

$$||x_1|\tau^2(|x_1|)| \leqslant \frac{D}{1+|x_1|} \int_{\mathscr{B}(|x_1|)} (|\Delta_r \tau^2| + |\tau^2|)\, r^2\, dr. \quad (5.6)$$

Now, by using the definition of τ^2, (5.6) leads to

$$|\tau(x_1)| \leqslant (C/|x_1|)\, \|Q\varphi_0^N/u\|_{W^{2,2}(\mathbb{R}^{3N})}, \quad (5.7)$$

for some $C>0$ and $|x_1| \geqslant a > 0$.

Remark 5.1: By integration of $\|Q\varphi_0^N/u\|_{L^2(\mathbb{R}^{3(N-1)})}$ one has

$$\|Q\varphi_0^N/u\|^2 = |1 - u^2/\rho|. \quad (5.8)$$

Equation (5.8) and Lemma 5.1 imply

$$u(|x_1|) = \rho^{1/2}(|x_1|)(1 + O(1/|x_1|^2)). \quad (5.9)$$

We can now prove our main result.

Theorem 5.1:

$$\|\varphi_0^N/\rho^{1/2} - \varphi_0^{N-1}\|_{L^2(\mathbb{R}^{3(N-1)})} = O(1/|x_1|), \quad (5.10a)$$

$$\left(\frac{\varphi_0^N}{\rho^{1/2}} (H^{N-1} - E_0^{N-1}) \frac{\varphi_0^N}{\rho^{1/2}} \right)_{L^2(\mathbb{R}^{3(N-1)})} = O\left(\frac{1}{|x_1|^2} \right). \quad (5.10b)$$

Proof: One has

$$\|\varphi_0^N/\rho^{1/2} - \varphi_0^{N-1}\|_{L^2(\mathbb{R}^{(2N-1)})} \leqslant \|P\varphi_0^N/\rho^{1/2} - \varphi_0^{N-1}\|_{L^2()} + \|Q\varphi_0^N/\rho^{1/2}\|_{L^2()}. \quad (5.11)$$

Now

$$\|Q\varphi_0^N/\rho^{1/2}\|_{L^2()} \leqslant \|Q\varphi_0^N/u\|_{L^2()},$$

and by Lemma 5.1

$$\|Q\varphi_0^N/\rho^{1/2}\|_{L^2(\mathbb{R}^{3(N-1)})} \leqslant C/|x_1|, \quad \text{for some } C>0. \quad (5.12)$$

For the second term in the rhs of (5.11), one gets

$$\|P\varphi_0^N/\rho^{1/2} - \varphi_0^{N-1}\|_{L^2()} = |1 - u/\rho^{1/2}|. \quad (5.13)$$

Remark 5.1 shows that this term decays like $1/|x_1|^2$. This proves (5.10a). To prove (5.10b) we proceed analogously

$$\left(\frac{\varphi_0^N}{\rho^{1/2}}, (H^{N-1} - E_0^{N-1}) \frac{\varphi_0^N}{\rho^{1/2}} \right) \leqslant \left(\frac{Q\varphi_0^N}{u}, (H^{N-1} - E_0^{N-1}) \frac{Q\varphi_0^N}{u} \right). \quad (5.14)$$

Let us set

$$\xi(|x_1|) = \left(\frac{Q\varphi_0^N}{u}, (H^{N-1} - E_0^{N-1}) \frac{Q\varphi_0^N}{u} \right);$$

then $\xi(|x_1|)$ is spherically symmetric, and by the same arguments as in the proof of Lemma 5.1 one has

$$||x_1|\xi(|x_1|)| \leqslant \frac{D}{1+|x_1|} \int_{\mathbb{R}^3} (|\Delta_r \xi| + |\xi|)\, d^3x. \quad (5.15)$$

Now, the Cauchy–Schwarz inequality leads to

$$\int_{\mathbb{R}^3} d^3x |\Delta_r \xi| \leqslant D \{ \|\Delta_1 (Q\varphi_0^N/u)\|_{L^2(\mathbb{R}^{3N})} \times \|(H^{N-1} - E_0^{N-1})(Q\varphi_0^N/u)\|_{L^2(\mathbb{R}^{3N})} + \|(H^{N-1} - E_0^{N-1})^{1/2} \nabla_1 (Q\varphi_0^N/u)\| L^2(\mathbb{R}^{3N}) \}, \quad (5.16)$$

for some $D>0$.

Noting that the last term can be estimated by

$$\|(H^{N-1} - E_0^{N-1})^{1/2} \nabla_1 (Q\varphi_0^N/u)\| \leqslant \|(-\Delta_1 + H^{N-1} - E_0^{N-1})(Q\varphi_0^N/u)\|^2, \quad (5.17)$$

we see that the integral in the left-hand side of (5.16) is finite; therefore by (5.15)

$$\xi(|x_1|) \leqslant C/|x_1|^2, \quad (5.18)$$

for some $C>0$ and $|x_1| \geqslant a > 0$.

ACKNOWLEDGMENTS

I thank J. M. Combes and E. Mourre for fruitful discussions.

[1] J. M. Combes and L. Thomas, Commun. Math. Phys. **34**, 251 (1973).
[2] B. Simon, Commun. Math. Phys. **55**, 259 (1977).
[3] P. Deift, W. Hunziker, B. Simon, and E. Vock, Commun. Math. Phys. **64**, 1 (1978).
[4] R. Ahlrichs, M. Hoffman-Ostenhof, T. Hoffman-Ostenhof, and J. Morgan, III, Phys. Rev. A **23**, 2106 (1981).
[5] S. Agmon, *Mathematical Notes* (Princeton U.P., Princeton, NJ, 1982).
[6] R. Carmona and B. Simon, Commun. Math. Phys. **80**, 59 (1981).
[7] J. M. Combes, M. Hoffman-Ostenhof, and T. Hoffman-Ostenhof, J. Math. Phys. **22**, 1299 (1981).
[8] E. H. Lieb and B. Simon, Adv. Appl. Math. **7**, 324 (1980).
[9] I. Sigal, Commun. Math. Phys. **85**, 309 (1982).
[10] J. M. Combes, *Rigorous Atomic and Molecular Physics*, edited by Wightman and Velo (Plenum, New York, 1981).
[11] J. M. Combes, Commun. Math. Phys. **12**, 283 (1969).
[12] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (Mc Graw–Hill, New York, 1955).
[13] M. Reed and B. Simon, *Scattering Theory* (Academic, New York, 1978), Vol. III.
[14] D. Babbitt and E. Balslev, J. Funct. Anal. **18**, 1 (1975).
[15] P. Briet, thèse 3ème cycle, Université d'Aix-Marseille II, 1983.
[16] M. Reed and B. Simon, *Functional Analysis* (Academic, New York, 1975), Vol. I.
[17] R. A. Adams, *Sobolev Spaces* (Academic, New York, 1975).

The quantum normal form and its equivalents

M. K. Ali
Department of Physics, University of Lethbridge, Lethbridge, Alberta, Canada T1K 3M4

(Received 25 March 1985; accepted for publication 7 June 1985)

A quantum analog, called the quantum normal form, of the classical Birkhoff–Gustavson normal form is presented. The algebraic relationship between the quantum and Birkhoff–Gustavson normal forms has been established by developing the latter using Lie transforms. It is shown that the Birkhoff–Gustavson normal form can be obtained from the quantum normal form. Using an anharmonic oscillator and a Henon–Heiles system as test cases, the equivalence between the quantum normal form and the Rayleigh–Schrödinger perturbation method is shown. This equivalence provides an algebraic connection between the Birkhoff–Gustavson normal form and the Rayleigh–Schrödinger perturbation approach. The question of Weyl and torus quantizations of the Birkhoff–Gustavson normal form is discussed in the light of the quantum normal form.

I. INTRODUCTION

In recent years, the Birkhoff–Gustavson normal form[1,2] (BGNF) has received much attention[3–5] in the description of dynamical systems. A quantum description via the BGNF involves two steps: (1) finding the normal form by canonical transformations of coordinates and momenta, and (2) quantizing the normal form. The first step is rather straightforward but quantization of the normal form addresses the fundamental question[3,4,6] of how to quantize a classical problem in which the coordinates and momenta do not appear in a simple manner (the BGNF consists of terms that are products of high powers of coordinates and momenta). A number of investigators[4] have used the method of torus quantization and recently Robnik[3a] and Uzer and Marcus[3b] have used the linear Weyl quantization for the BGNF. In general, these methods yield different quantum results for a given BGNF. As pointed out by Robnik,[3a] the Weyl quantization does not preserve the Poisson structure in all coordinates. As an illustration we consider the Hamiltonian $K = 1/Q^2 + P^2Q^4$, where Q and P are canonical. Using the Weyl quantization gives $\mathbf{K} = 1/\mathbf{Q}^2 + \frac{1}{4}(\mathbf{Q}^4\mathbf{P}^2 + 2\mathbf{P}\mathbf{Q}^4\mathbf{P} + \mathbf{P}^2\mathbf{Q}^4)$. If, however, we introduce the canonical transformation $x = 1/Q$ and $p = -PQ^2$, then K is converted into the harmonic oscillator Hamiltonian $H = p^2 + x^2$. Working quantization rules for H are known and we obtain $\mathbf{H} = \mathbf{p}^2 + \mathbf{x}^2$, where $[\mathbf{x},\mathbf{p}] = i\hbar$. If it is also required that $[\mathbf{Q},\mathbf{P}] = i\hbar$, then the energy spectra of $\mathbf{H}$ and $\mathbf{K}$ become different. When the Weyl quantization is applied to the BGNF the Poisson structure is not conserved, and no unique solution to the problem is known at the present time. Nonetheless, for those limiting cases that have been tested, the energies obtained by the quantization of the BGNF are found to be quite good. To have a better understanding of the situation, we have adopted a method that does not depend on the Weyl quantization but does produce the BGNF in the classical limit. We call this method the quantum normal form (QNF). The main purpose of this paper is (1) to compare the results of the QNF and the quantized BGNF, (2) to show the algebraic similarity between the QNF and the BGNF, and (3) to show the equivalence between the QNF and Rayleigh–Schrödinger perturbation theory[7] (RSP). The equivalence between the QNF and RSP and the algebraic similarity between the BGNF and QNF establish an algebraic connection between the BGNF and RSP.

In Sec. II we derive the BGNF of an anharmonic oscillator and a two-dimensional Henon–Heiles system using the Lie transform technique. The general algorithm developed in Sec. II for the Lie transform is adapted in Sec. III for obtaining the QNF of the above-mentioned systems. In Sec. IV we show that the QNF yields the BGNF and RSP results exactly. In Sec. IV we also compare the energies obtained by the QNF, Weyl quantization, and torus quantization. Concluding remarks are given in Sec. V.

II. THE LIE TRANSFORM FOR THE BGNF

In this section we employ the elegant technique of the Lie transform[8] for deriving the BGNF of a one-dimensional anharmonic oscillator and a two-dimensional Henon–Heiles system. The algorithm developed here will be adapted in Sec. III for finding the QNF of these systems. Consider the Hamiltonian $H(x,\epsilon)$ that depends on the $2n$-vector $x = x_1,x_2,\ldots,x_n,p_1,p_2,\ldots,p_n$ and the formal parameter ϵ. We want to transform the vector x to the $2n$-vector $y = \xi_1,\xi_2,\ldots,\xi_n,\eta_1,\eta_2,\ldots,\eta_n$ and the Hamiltonian $H(x,\epsilon)$ to $K(y,\epsilon)$ so that K is in the normal form. The transformations can be carried out by the method of Lie transform, for which various algorithms are available in the literature.[9–11] We have used the algorithm that was proposed by Deprit[9] and rederived by Kamel.[10] A brief description of it is given below.

Let us assume that the transformation $(x \to y)$ can be expressed in a series as

$$x = y + \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} y^{(n)}(y) . \tag{1}$$

Equation (1) is in fact a Taylor series expansion of x in powers of ϵ about $\epsilon = 0$ with $x = y$ at $\epsilon = 0$ and $(d^n x/d\epsilon^n)_0 = y^{(n)}(y)$. Now consider any differentiable scalar function $f(x,\epsilon)$, which can be expanded in the Taylor series

$$f(x,\epsilon) = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} f_n(x) , \tag{2}$$

where

 0022-2488/85/102565-08$02.50

$$f_n(x,\epsilon)=\left(\frac{\partial^n}{\partial\epsilon^n}f(x,\epsilon)\right)_{\epsilon=0}.$$

By expressing x in terms of y, the expansion (2) can be cast in the form

$$f(x,\epsilon)=\sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}f^{(n)}(y), \tag{3}$$

where

$$f^{(n)}(y)=\left(\frac{d^n}{d\epsilon^n}f(x,\epsilon)\right)_{\epsilon=0,x=y}.$$

In order to express $f(x,\epsilon)$ in terms of y one needs to know the $f^{(n)}(y)$. Let us now discuss how to calculate the $f^{(n)}(y)$. Using the rules of differentiation we have

$$\frac{d}{d\epsilon}f(x,\epsilon)=\frac{\partial}{\partial\epsilon}f(x,\epsilon)+\left[\frac{\partial}{\partial x}f(x,\epsilon)\right]\cdot\frac{dx}{d\epsilon}, \tag{4}$$

where $[(\partial/\partial x)f(x,\epsilon)]\cdot(dx/d\epsilon)$ means the dot product of the gradient of $f(x,\epsilon)$ with respect to x and the vector $dx/d\epsilon$. The Lie transform is now defined by the equations

$$\begin{aligned}\frac{dx_i}{d\epsilon}&=\frac{\partial}{\partial p_i}w(x,\epsilon),\\ \frac{dp_i}{d\epsilon}&=-\frac{\partial}{\partial x_i}w(x,\epsilon),\qquad i=1,2,\ldots,n.\end{aligned} \tag{5}$$

Substituting (5) in (4) gives

$$\frac{d}{d\epsilon}f(x,\epsilon)=\frac{\partial}{\partial\epsilon}f(x,\epsilon)+L_w f(x,\epsilon), \tag{6}$$

where $L_w f(x,\epsilon)$ is the Lie derivative of f generated by w and is given by

$$L_w f(x,\epsilon)=\{f,w\}=\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}\frac{\partial w}{\partial p_i}-\frac{\partial f}{\partial p_i}\frac{\partial w}{\partial x_i}\right). \tag{7}$$

In Eq. (7), $\{f,w\}$ is the Poisson bracket of f and w. The function w is also assumed to have a series expansion

$$w(x,\epsilon)=\sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}w_{n+1}(x). \tag{8}$$

Using Eqs. (2) and (6)–(8) we obtain

$$\frac{d}{d\epsilon}f(x,\epsilon)=\sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}f_n^{(1)}(x), \tag{9}$$

where

$$f_n^{(1)}(x)=f_{n+1}(x)+\sum_{m=0}^{n}\binom{n}{m}L_{m+1}f_{n-m}(x)$$

and

$$L_j f(x)=L_{w_j}=\{f(x),w_j(x)\}. \tag{10}$$

The k th derivative of $f(x,\epsilon)$ with respect to ϵ can be written (by induction) as

$$\frac{d^k}{d\epsilon^k}f(x,\epsilon)=\sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}f_n^{(k)}(x), \tag{11}$$

where

$$f_n^{(k)}(x)=f_{n+1}^{(k-1)}(x)+\sum_{m=0}^{\infty}\binom{n}{m}L_{m+1}f_{n-m}^{(k-1)}(x),$$
$$k\geqslant 1,\quad m\geqslant 0. \tag{12}$$

Equation (12) is the desired recursive algorithm for finding $f_0^{(n)}(x)$. Denoting $f_n^{(0)}(x)=f_n(x)$, $f_0^{(n)}(x)=f^{(n)}(x)$, and using Eq. (3) we obtain

$$f^{(k)}(y)=(f^{(k)}(x))_{x=y}.$$

Thus, the $f^{(n)}(y)$ of Eq. (3) are found and the expansion of $f(x,\epsilon)$ in terms of y is accomplished. For our Hamiltonian problem

$$f(x,\epsilon)=H(x,\epsilon)=\sum_{n=0}\frac{\epsilon^n}{n!}H_n(x) \tag{13}$$

and

$$f(x(y,\epsilon),\epsilon)=K(y,\epsilon)=\sum_{n=0}\frac{\epsilon^n}{n!}K_n(y). \tag{14}$$

Comparison of Eqs. (14) and (3) gives

$$K_n(y)=f^{(n)}(y), \tag{15}$$

which provides the canonical transformation of $H(x,\epsilon)$ to $K(y,\epsilon)$. Working out a few low-order terms $[K_i(y)=f^{(i)}(y)]$ gives

$$\begin{aligned}K_0&=H_0,\\ K_1&=H_1+\{H_0,w_1\},\\ K_2&=H_2+2\{H_1,w_1\}+\{H_0,w_2\}+\{\{H_0,w_1\},w_1\}.\end{aligned} \tag{16}$$

Although $K(y,\epsilon)$ is the canonical transform of $H(x,\epsilon)$, it is not yet in the BGNF. To be in the BGNF, it has to satisfy the additional condition

$$\{K_0(y),K_m(y)\}=0,\quad \forall m, \tag{17}$$

where the Poisson bracket

$$\{K_0(y),K_m(y)\}=\sum_{i=1}^{n}\left(\frac{\partial K_0}{\partial\xi_i}\frac{\partial K_m}{\partial\eta_i}-\frac{\partial K_0}{\partial\eta_i}\frac{\partial K_m}{\partial\xi_i}\right). \tag{18}$$

Suppose that $K_0,K_1,\ldots,K_{m-1}$ have been found in the normal form and we want to compute the K_m term. To do this, it is realized that Eq. (12) reduces to the form

$$K_m=f^{(m)}=\{K_0,w_m\}+H_m+T_m, \tag{19}$$

where H_m is the mth-order term of the known Hamiltonian H and T_m is known because it contains $H_0,H_1,\ldots,H_{m-1}$ and $w_1,w_2,\ldots,w_{m-1}$. Since the condition (17) needs to be satisfied, the known sum H_m+T_m is split[2] into two parts

$$H_m+T_m=N_m+R_m, \tag{20}$$

so that

$$\{K_0,N_m\}=0 \tag{21}$$

and hence

$$K_m=N_m. \tag{22}$$

We then find w_m by solving the differential equation

$$\{K_0,w_m\}=-R_m. \tag{23}$$

The process is then continued to higher m values. Using the method outlined above, we have generated the normal forms for a one-dimensional anharmonic oscillator and a two-dimensional Henon–Heiles system. For both cases, the symbolic language MAPLE[12] was used for algebraic manipulations.

A. The BGNF of the anharmonic oscillator

The anharmonic oscillator that we have studied is described by the Hamiltonian

$$H = H_0 + H_1 + H_2/2\,, \tag{24}$$

$$H_0 = (\omega/2)(p^2 + x^2)\,, \tag{25}$$

$$H_1 = bx^3\,, \tag{26}$$

$$H_2 = gx^4\,. \tag{27}$$

The parameters ω, b, and g can be varied to have one or two real minima for the potential $V(x) = (\omega/2)x^2 + bx^3 + gx^4/2$. For certain values of these parameters, the oscillator has been studied by quantum mechanical approximate methods.[13] However, there seems to exist no approximate classical solution of the dynamical problem described by Eqs. (24)–(27). In our present study of the normal forms we develop classical as well as quantum perturbation series for the oscillator. To obtain the BGNF, we expanded K and w in powers of ξ and η with K_i and w_i being homogeneous polynomials of order $i+2$. The following preliminary transformation was found useful:

$$Q = (1/\sqrt{2})(\xi - i\eta), \quad P = (1/\sqrt{2})(-i\xi + \eta)\,. \tag{28}$$

The transformation in Eq. (28) is canonical and hence it preserves Poisson brackets. In terms of P and Q we have

$$K_0 = i\omega QP$$

and for any function $A(Q,P)$

$$\{K_0, A\} = i\omega dA\,, \tag{29}$$

where

$$d = P\frac{\partial}{\partial P} - Q\frac{\partial}{\partial Q}\,. \tag{30}$$

From Eqs. (29) and (30) it is clear that a monomial $C_{\mu\nu}Q^\mu P^\nu$ of A is an eigenfunction of the operator d with the eigenvalue $\nu - \mu$. This practical advantage of the transformation (28) is used in evaluating the Poisson brackets involving K_0. For example, in generating N_m from Eq. (20) one needs to compute the Poisson bracket $\{K_0, H_m + T_m\}$, where the sum $H_m + T_m$ is a polynomial in Q and P. A monomial of $H_m + T_m$ belongs to N_m if $\nu - \mu = 0$, otherwise it belongs to R_m. In Eq. (23), w_m and R_m are homogeneous polynomials of order $m+2$ in Q and P and the equation is solved by comparing the coefficients of individual monomials. Following the mentioned steps we have generated the normal form terms $K_0, K_1, \ldots, K_{16}$ for the anharmonic oscillator. A few of these are given below:

$$K_0 = \omega\tau\,, \tag{31}$$

$$K_2 = (\tfrac{3}{2}g - (15/2\omega)b^2)\tau^2\,, \tag{32}$$

$$K_4 = \left(675\frac{gb^2}{\omega^2} - \frac{51}{2\omega}g^2 - \frac{2115}{2\omega^3}b^4\right)\tau^3, \quad \tau = \frac{\xi^2 + \eta^2}{2}\,, \tag{33}$$

$$K_j = 0, \quad \text{for odd } j\,,$$

$$K \simeq K_0 + K_2/2! + K_4/4!\,. \tag{34}$$

The one-dimensional oscillator, Eqs. (24)–(27), has one constant of motion (the Hamiltonian itself) and hence it is integrable for real time (integrability and its meaning in the complex time plane require further investigations[14]). Equations (31)–(34) show that the transformed Hamiltonian K also depends on only one constant K_0, since K is a series in K_0. The convergence of the BGNF series for this integrable system will be dealt with elsewhere. In Sec. IV it will be shown that the BGNF results of Eq. (31)–(34) are the classical limits of the corresponding QNF results.

B. The BGNF of the Henon–Heiles system

The system of Henon–Heiles Hamiltonians can be described by the equations

$$H = H_0 + H_1\,, \tag{35}$$

$$H_0 = (\omega_1/2)(p_1^2 + x_1^2) + (\omega_2/2)(p_2^2 + x_2^2)\,, \tag{36}$$

$$H_1 = \beta x_1^2 x_2 + \gamma x_2^3\,, \tag{37}$$

$$H_j = 0, \quad \text{for } j \geqslant 2\,.$$

The original Henon–Heiles Hamiltonian[15] corresponds to the parameter values $\beta = 1$ and $\gamma = -\frac{1}{3}$ and the BGNF of this extensively studied Hamiltonian has been reported earlier.[4,5] Using our Lie transform method we have generated the terms $K_0, K_1, \ldots, K_{10}$ for the 1:1 resonant case. The *resonant* and *nonresonant* cases are defined by the equation

$$\mu\omega_1 + \nu\omega_2 = 0\,. \tag{38}$$

If ω_1 and ω_2 are such that two nonzero integers μ and ν can be found to satisfy (38), then it is a resonant case, otherwise it is a nonresonant problem. Normal forms of resonant and nonresonant cases need to be treated separately.

In the *nonresonant* case the normal form becomes a function of the Hamiltonians of the two uncoupled harmonic oscillators of Eq. (36). The results, up to and including the K_2 term, are given by

$$K_0(\text{nonres}) = \omega_1\tau_1 + \omega_2\tau_2\,, \tag{39}$$

$$K_2(\text{nonres}) = \frac{\beta^2}{4}\left[\frac{2\omega_2\tau_1^2 - 16\omega_1\tau_1\tau_2}{4\omega_1^2 - \omega_2^2} - \frac{4\tau_1^2}{\omega_2}\right] - \frac{6\beta\gamma}{\omega_2}\tau_1\tau_2 - \frac{15}{2}\frac{\gamma^2}{\omega_2}\tau_2^2\,, \tag{40}$$

where

$$\tau_1 = (\eta_1^2 + \xi_1^2)/2\,, \quad \tau_2 = (\eta_2^2 + \xi_2^2)/2\,.$$

In this work we have considered the *1:1 resonant case* and the results are given by

$$K_0(\text{res}) = K_0\begin{pmatrix}\text{nonres}\\ \omega_1 = \omega_2 = 1\end{pmatrix}, \tag{41}$$

$$K_2(\text{res}) = K_2\begin{pmatrix}\text{nonres}\\ \omega_1 = \omega_2 = 1\end{pmatrix} + \tfrac{1}{2}(\beta\gamma/2 - \beta^2)\{(\xi_1^2 - \eta_1^2)(\xi_2^2 - \eta_2^2) + 4\xi_1\xi_2\eta_1\eta_2\}\,. \tag{42}$$

It will be shown in Sec. IV that the above BGNF results for the resonant and nonresonant cases, Eqs. (39)–(42), are the classical limits of the corresponding degenerate and nondegenerate QNF results. Quantized energies of these results will be discussed in Sec. IV B. We consider now the QNF expansions.

III. THE QNF EXPANSIONS

Th first step in obtaining the QNF is to select a suitable coordinate system. In the chosen coordinate system, the classical Hamiltonian should assume a form that is quantizable by rules having empirical support. In the following we assume that the x-coordinate system is such a selection. The basic algorithm for obtaining the QNF has been developed in Sec. II. The modifications that need to be made in that algorithm are the following: (1) replace all the observables by their quantum operators, and (2) replace all Poisson brackets by the corresponding commutators. It is clear that replacing the Poisson brackets by commutators establishes the algebraic similarities between the BGNF and QNF. Undoubtedly, the Lie series expansion of the BGNF has made this quantization simple. Thus Eqs. (13), (14), (8), (12), (16), (17), and (19)–(23) assume the forms

$$\mathbf{H}(\mathbf{x},\epsilon)=\sum_{m=0}^{\infty}\frac{\epsilon^m}{m!}\mathbf{H}_m(\mathbf{x}), \tag{43}$$

$$\mathbf{K}(\mathbf{y},\epsilon)=\sum_{m=0}^{\infty}\frac{\epsilon^m}{m!}\mathbf{K}_m(\mathbf{y}), \tag{44}$$

$$\mathbf{W}(\mathbf{x},\epsilon)=\sum_{m=0}^{\infty}\frac{\epsilon^m}{m!}\mathbf{W}_{m+1}(\mathbf{x}), \tag{45}$$

$$\mathbf{f}_n^{(k)}(\mathbf{x})=\mathbf{f}_{n+1}^{(k-1)}(\mathbf{x})+\sum_{m=0}^{n}\binom{n}{m}\ \mathbf{L}_{m+1}\mathbf{f}_{n-m}^{(k-1)}(\mathbf{x}), \tag{46}$$

$$\mathbf{L}_m\mathbf{f}=[\mathbf{f},\mathbf{W}_m],$$

$$\mathbf{K}_0=\mathbf{H}_0,$$

$$\mathbf{K}_1=\mathbf{H}_1+[\mathbf{H}_0,\mathbf{W}_1],$$

$$\mathbf{K}_2=\mathbf{H}_2+[\mathbf{H}_0,\mathbf{W}_2]+2[\mathbf{H}_1,\mathbf{W}_1]+[[\mathbf{H}_0,\mathbf{W}_1],\mathbf{W}_1], \tag{47}$$

$$[\mathbf{K}_0,\mathbf{K}_m]=0, \tag{48}$$

$$\mathbf{K}_m=[\mathbf{H}_0,\mathbf{W}_m]+\mathbf{H}_m+\mathbf{T}_m, \tag{49}$$

$$\mathbf{H}_m+\mathbf{T}_m=\mathbf{N}_m+\mathbf{R}_m, \tag{50}$$

$$[\mathbf{K}_0,\mathbf{N}_m]=0, \tag{51}$$

$$\mathbf{K}_m=\mathbf{N}_m, \tag{52}$$

$$[\mathbf{K}_0,\mathbf{W}_m]=-\mathbf{R}_m. \tag{53}$$

The requirement of the $\mathbf{K}_m$ to be in the QNF is described by Eq. (48). We now discuss the QNF of the anharmonic oscillator and the Henon–Heiles system of Sec. II.

A. The QNF of the anharmonic oscillator

The Hamiltonian of the oscillator, Eqs. (24)–(27), has a simple form in the chosen coordinate system. It is quadratic in the momentum, and working quantization rules for such forms of Hamiltonians are known. We convert the initial Hamiltonian H into a quantum operator by using the creation and annihilation operators of the harmonic oscillator Eq. (25). The operators for a set of harmonic oscillators[16] are given by

$$a_j=(1/\sqrt{2\hbar})(\mathbf{x}_j+i\mathbf{p}_j), \tag{54}$$

$$a_j^\dagger=(1/\sqrt{2\hbar})(\mathbf{x}_j-i\mathbf{p}_j), \tag{55}$$

where

$$[a_j,a_k^\dagger]=\delta_{jk}, \tag{56}$$

$$a_j|n_j\rangle=\sqrt{n_j}\,|n_j-1\rangle, \tag{57}$$

$$a_j^\dagger|n_j\rangle=\sqrt{n_j+1}\,|n_j+1\rangle. \tag{58}$$

In Eqs. (57) and (58), $|n_j\rangle$ represents the nth eigenket of the jth harmonic oscillator. For the one-dimensional problem of Eqs. (24)–(27), $j,k=1$. Writing $a_1=a$ and $a_1^\dagger=a^\dagger$, we have the Hamiltonian operators

$$\mathbf{H}_0=(\hbar\omega/2)(aa^\dagger+a^\dagger a), \tag{59}$$

$$\mathbf{H}_1=b(\hbar/2)^{3/2}(a^3+3a^\dagger aa^\dagger+3aa^\dagger a+a^{\dagger 3}), \tag{60}$$

$$\mathbf{H}_2=(g\hbar^2/4)[a^4+2(a^2a^\dagger a+aa^\dagger a^2+a^{\dagger 2}aa^\dagger+a^\dagger aa^{\dagger 2}) + 3(aa^\dagger aa^\dagger+a^\dagger aa^\dagger a)+a^{\dagger 4}]. \tag{61}$$

The operator $\mathbf{H}=\mathbf{H}_0+\mathbf{H}_1+\mathbf{H}_2/2$ is transformed into $\mathbf{K}$, the QNF, for computational purposes. The steps that one follows to find the $\mathbf{K}_m$ are similar to those for the classical case. The generator $\mathbf{W}$ is assumed to have an expansion in a and $a^\dagger$ (all possible permutations of powers of a and $a^\dagger$ are considered) while the $\mathbf{W}_m$ of Eq. (45) are homogeneous polynomials of order $m+2$. For solving the recursion equations, such as Eq. (47), we notice that each commutator reduces by 2 the combined order of the operators involved. This behavior of a commutator is similar to that of a Poisson bracket and is due to the basic relation (56). In addition, the commutator of K_0 and any other operator is easy to evaluate because

$$[a^\dagger a,aa^\dagger]=0,$$

$$[a^\dagger a,a^{\dagger\mu}a^\nu]=C_1\,a^{\dagger\mu}a^\nu,$$

$$[aa^\dagger,a^{\dagger\mu}a^\nu]=C_2\,a^{\dagger\mu}a^\nu,$$

where μ,ν are any two integers and C_1,C_2 are constants. The results for the anharmonic oscillator through $\mathbf{K}_4$ are given below:

$$\mathbf{K}_0=(\hbar\omega/2)(aa^\dagger+a^\dagger a), \tag{62}$$

$$\mathbf{K}_2=\tfrac{3}{4}g\hbar^2[aa^\dagger aa^\dagger+a^\dagger aa^\dagger a] - (b^2\hbar^2/4\omega)[a^2a^{\dagger 2}+a^{\dagger 2}a^2] - (5b^2\hbar^2/4\omega)[a^\dagger a^2a^\dagger+aa^{\dagger 2}a] \tag{63}$$

$$\mathbf{K}_4=-(g^2\hbar^3/\omega)[\tfrac{3}{8}T_1+\tfrac{3}{16}T_2+3T_3+\tfrac{15}{8}T_4+\tfrac{9}{4}T_5+\tfrac{3}{4}T_6] + (gb^2\hbar^3/\omega^2)[-\tfrac{19}{16}T_1+\tfrac{259}{32}T_2+\tfrac{243}{2}T_3+\tfrac{631}{32}T_4 + 56T_5+\tfrac{253}{16}T_6+\tfrac{135}{4}T_7] - (b^4\hbar^3/\omega^3)[-\tfrac{3}{2}T_1+\tfrac{177}{8}T_2+\tfrac{495}{2}T_3+\tfrac{369}{16}T_4 + \tfrac{347}{4}T_5+\tfrac{63}{4}T_6+\tfrac{385}{8}T_7]$$

$$T_1=a^3a^{\dagger 3}+a^{\dagger 3}a^3,$$

$$T_2=a^\dagger a^3a^{\dagger 2}+a^{\dagger 2}a^3a^\dagger+a^2a^{\dagger 3}a+aa^{\dagger 3}a^2,$$

$$T_3=aa^\dagger aa^\dagger aa^\dagger+a^\dagger aa^\dagger aa^\dagger a,$$

$$T_4=a^2a^{\dagger 2}aa^\dagger+a^{\dagger 2}a^2a^\dagger a+aa^\dagger a^2a^{\dagger 2}+a^\dagger aa^{\dagger 2}a^2,$$

$$T_5=a^\dagger aa^\dagger a^2a^\dagger+a^\dagger a^2a^\dagger aa^\dagger+aa^{\dagger 2}aa^\dagger a+aa^\dagger aa^{\dagger 2}a,$$

$$T_6=a^2a^\dagger aa^{\dagger 2}+a^{\dagger 2}aa^\dagger a^2,$$

$$T_7=aa^{\dagger 2}a^2a^\dagger+a^\dagger a^2a^{\dagger 2}a. \tag{64}$$

The expression for $\mathbf{K}$ is obtained from Eq. (44) after setting $\epsilon=1$.

B. The QNF of the Henon–Heiles system

The classical Hamiltonian of the Henon–Heiles system given in Eqs. (35)–(37) has a simple form in the chosen x-coordinate system. We convert it into a quantum operator by using the creation and annihilation operators of Eqs. (54)–(58) with $j,k = 1,2$. Thus we have

$$\mathbf{H}_0 = (\hbar\omega_1/2)(a_1 a_1^\dagger + a_1^\dagger a_1) + (\hbar\omega_2/2)(a_2 a_2^\dagger + a_2^\dagger a_2)\,, \tag{65}$$

$$\mathbf{H}_1 = \beta(\hbar/2)^{3/2}(a_1^2 + a_1 a_1^\dagger + a_1^\dagger a_1 + a_1^{\dagger 2})(a_2 + a_2^\dagger) + \gamma(\hbar/2)^{3/2}(a_2^3 + 3(a_2 a_2^\dagger a_2 + a_2^\dagger a_2 a_2^\dagger) + a_2^{\dagger 3})\,, \tag{66}$$

$$\mathbf{H}_j = 0, \quad \text{for } j \geqslant 2\,.$$

The classical resonant and nonresonant conditions, Eq. (38), give rise to degenerate and nondegenerate states of the operator $\mathbf{H}_0$. As in the classical case, these degenerate and nondegenerate cases need to be treated separately. The basic computational steps to follow are the same as those used for the oscillator. The main difference is that here we have $a_1, a_1^\dagger, a_2$, and $a_2^\dagger$, whereas for the oscillator we had only a and $a^\dagger$. For the present case the operators $\mathbf{K}_m, \mathbf{W}_m, \mathbf{N}_m, \mathbf{R}_m$, etc. are expanded in series of $a_1, a_1^\dagger, a_2$, and $a_2^\dagger$. Thus, for the *nondegenerate case* we obtain

$$\mathbf{K}_0(\text{nondeg}) = \mathbf{H}_0\,, \tag{67}$$

$$\mathbf{K}_2(\text{nondeg}) = \frac{\beta^2\hbar^2}{4}\left[\frac{\omega_2\mathbf{C} - 4\omega_1\mathbf{AB}}{4\omega_1^2 - \omega_2^2} - \frac{\mathbf{A}^2}{\omega_2}\right] - \frac{3\beta\gamma\hbar^2}{2\omega_2}\mathbf{AB} - \frac{\gamma^2\hbar^2}{4\omega_2}[a_2^2 a_2^{\dagger 2} + a_2^{\dagger 2} a_2^2 + 5(a_2^\dagger a_2^2 a_2^\dagger + a_2 a_2^{\dagger 2} a_2) + 9(a_2 a_2^\dagger a_2 a_2^\dagger + a_2^\dagger a_2 a_2^\dagger a_2)]\,,$$

$$\mathbf{A} = a_1 a_1^\dagger + a_1^\dagger a_1\,,$$

$$\mathbf{B} = a_2 a_2^\dagger + a_2^\dagger a_2\,,$$

$$\mathbf{C} = a_1^2 a_1^{\dagger 2} + a_1^{\dagger 2} a_1^2\,, \tag{68}$$

$$\mathbf{K}_j = 0, \quad \text{for odd values of } j\,.$$

The results for the *1:1 degenerate case* ($\omega_1 = \omega_2 = 1$) are found to be

$$\mathbf{K}_0(\text{deg}) = \mathbf{K}_0\begin{pmatrix}\text{nondeg}\\ \omega_1 = \omega_2 = 1\end{pmatrix}, \tag{69}$$

$$\mathbf{K}_2(\text{deg}) = \mathbf{K}_2\begin{pmatrix}\text{nondeg}\\ \omega_1 = \omega_2 = 1\end{pmatrix} + \left(\frac{\beta\gamma}{2} - \beta^2\right)\hbar^2\mathbf{D}\,,$$

$$\mathbf{D} = a_1^2 a_2^{\dagger 2} + a_1^{\dagger 2} a_2^2\,. \tag{70}$$

Equations (67)–(70) are the quantum equivalents of the classical results of Eqs. (39)–(42). A comparison of the results of Eqs. (39)–(42) and (67)–(70) shows how various classical normal form terms have assumed their quantized forms. Further discussions of these results are given in the next section.

IV. CLASSICAL AND QUANTUM EQUIVALENTS OF THE QNF

In this section we show that the QNF yields the BGNF, and that the Rayleigh–Schrödinger perturbation (RSP) results are identical to those of the QNF. Here we also compare the QNF energies with the energies obtained by algebraic and torus quantizations.

A. The BGN yields the BGNF

The BGNF is obtained from the QNF by replacing the operators a_j and $a_j^\dagger$ by their "classical values"

$$a_j = (1/\sqrt{2\hbar})(\xi_j + i\eta_j), \quad a_j^\dagger = (1/\sqrt{2\hbar})(\xi_j - i\eta_j)\,,$$

where ξ_j and η_j are the classical (commuting) coordinates and momenta. It, however, should be remembered that if the QNF obtained in Sec. III need to be transformed into different forms by using the commutation rules of the a_j and $a_j^\dagger$, Eq. (56), then care must be taken not to change the homogeneity or the order of the $\mathbf{K}_m$. The reason for this can be illustrated by considering the two expressions $a_j^\dagger a_j + \frac{1}{2}$ and $\frac{1}{2}(a_j a_j^\dagger + a_j^\dagger a_j)$. Although these expressions are quantum mechanically equivalent, they produce the classical results $\frac{1}{2}(\xi_j^2 + \eta_j^2) + \frac{1}{2}$ and $\frac{1}{2}(\xi_j^2 + \eta_j^2)$, respectively. The homogeneous form gives the correct classical Hamiltonian. Thus, by preserving the order and homogeneity of the QNF one obtains exactly the BGNF equations [(31)–(34) and (39)–(42)] from the QNF equations (62)–(64) and (67)–(70)].

B. The QNF energies

By construction, the QNF's are diagonal (block diagonal in the degenerate case) in the harmonic oscillator basis sets, and hence evaluation of the QNF energies is a simple matter. One needs to use the basic properties, Eqs. (56)–(58), of the creation and annihilation operators. The energy of the *anharmonic oscillator* is given by

$$E(\mathbf{K}_0) = \langle n|\mathbf{K}_0|n\rangle = \hbar\omega(n + \tfrac{1}{2})\,, \tag{71}$$

$$E(\mathbf{K}_2) = \left\langle n\left|\frac{\mathbf{K}_2}{2}\right|n\right\rangle = -\frac{15}{4}\frac{b^2\hbar^2}{\omega}\left(n^2 + n + \frac{11}{30}\right) + \frac{3}{4}g\hbar^2\left(n^2 + n + \frac{1}{2}\right), \tag{72}$$

$$E(\mathbf{K}_4) = \left\langle n\left|\frac{\mathbf{K}_4}{24}\right|n\right\rangle, = -(g^2\hbar^3/96\omega)[102n^3 + 153n^2 + 177n + 63] + (gb^2\hbar^3/48\omega^2)[1350n^3 + 2025n^2 + 1701n + 513] - (b^4\hbar^3/96\omega^3)[4230n^3 + 6345n^2 + 4905n + 1395]\,, \tag{73}$$

$$E(\mathbf{K}_j) = 0, \quad \text{for odd } j\,,$$

$$E \simeq E(\mathbf{K}_0) + E(\mathbf{K}_2) + E(\mathbf{K}_4)\,. \tag{74}$$

We now compare E of Eqs. (71)–(74) with E_{torus}, the energy obtained by substituting $(n + \frac{1}{2})$ for τ in the BGNF equations [(31)–(34)]. The difference between E and E_{torus} is of the order of $\hbar^2$ and is given by

$$E - E_{\text{torus}} = \frac{\hbar^2}{16}\left(3g - \frac{7b^2}{\omega}\right) - (2n+1)\hbar^3 \times\left[\frac{67g^2}{128\omega} - \frac{459gb^2}{64\omega^2} + \frac{1155b^4}{128\omega^3}\right]. \quad (75)$$

Comparing Eqs. (71)–(74) and (75), it is noticed that, *for each order of approximation*, the first two terms in E and E_{torus}, are the same. The substitution of $(n+\frac{1}{2})^m$ for τ^m means quantization of τ^m by $(a^\dagger a + \frac{1}{2})^m$. This "second quantization" preserves the Poisson and BGNF structures. On the other hand, the QNF also preserves the Poisson and BGNF structures. For $m>1$, the QNF corresponds to "higher-order quantizations" (through the successive replacements of Poisson brackets by commutators). An example of the effect of higher-order quantizations is given in Eq. (75), which indicates that E_{torus} can be called an approximation of E. Equation (75) also shows that the difference between E and E_{torus} is more pronounced at low n values as compared to high n values. This implies that at and near the ground level E_{torus} is expected to be less accurate than E for a given order of approximation. Since $E_{\text{torus}} \to E$ as n becomes large, one may use E_{torus} instead of E for large n. An illustration of the relative behavior of E and E_{torus} is given by the numerical data of Table I. The "exact" values in Table I were found by diagonalizing a 200×200 matrix formed by the elements of the operator $\mathbf{H}$ of Eqs. (59)–(61) in the harmonic oscillator basis set. The data of Table I bear out the expected behaviors of E and E_{torus}. It can be seen that both E and E_{torus} deviate more from the exact energy as n increases. Our experience shows that these low-order one-center approximate results (E and E_{torus}) cannot account for the quantum mechanical barrier penetration that becomes important when b is significant. Therefore, the practical use of these results is limited to low-lying energy levels under small perturbations.

The QNF energies of the Henon–Heiles system have been computed through $\mathbf{K}_2$. For the *nondegenerate case* we have

$$E_{\mathbf{K}_0}(\text{nondeg}) = \langle n_2 n_1|\mathbf{K}_0(\text{nondeg})|n_1 n_2\rangle = \hbar\omega_1(n_1+\tfrac{1}{2}) + \hbar\omega_2(n_2+\tfrac{1}{2}), \quad (76)$$

$$E_{\mathbf{K}_2}(\text{nondeg}) = \left\langle n_2 n_1\left|\frac{\mathbf{K}_2(\text{nondeg})}{2}\right|n_1 n_2\right\rangle = \frac{\beta^2\hbar^2}{8}\left[\frac{2\omega_2(n_1^2+n_1+1) - 4\omega_1(2n_1+1)(2n_2+1)}{4\omega_1^2-\omega_2^2} - \frac{(2n_1+1)^2}{\omega_2}\right] - \frac{3\beta\gamma\hbar^2}{4\omega_2}(2n_1+1)(2n_2+1) - \frac{15\gamma^2\hbar^2}{4\omega_2}\left(n_2^2+n_2+\frac{11}{30}\right), \quad (77)$$

$$E \simeq E_{\mathbf{K}_0}(\text{nondeg}) + E_{\mathbf{K}_2}(\text{nondeg}). \quad (78)$$

Comparing the energy E, Eqs. (76)–(78), with E_{torus} obtained by replacing $\tau_1 = n_1 + \frac{1}{2}$ and $\tau_2 = n_2 + \frac{1}{2}$ in Eqs. (39) and (40), we find the difference

$$E - E_{\text{torus}} = \frac{\hbar^2}{16}\left(\frac{3\beta^2\omega_2}{4\omega_1^2-\omega_2^2} - \frac{7\gamma^2}{\omega_2}\right), \quad \forall n_1, n_2. \quad (79)$$

As in the anharmonic oscillator problem discussed above, E_{torus} for the Henon–Heiles system is expected to be less accurate than E at and near the ground level. For higher energy levels E_{torus} may be used for E.

Unlike the previous two cases, Eqs. (71)–(74) and (76)–(79), the operator $\mathbf{K}_2$ for the degenerate states is block diagonal (except when $n_1+n_2<2$). The nth eigenstate, $n = n_1 + n_2 \geq 2$, of the degenerate state is constructed from a linear combination of the unperturbed states $|n_1 n_2\rangle$. For calculating the energy one needs to diagonalize the matrix $\mathbf{M}$ having elements $M_{ij} = \langle ji|\mathbf{K}_2/2|ij\rangle$. It may be pointed out that the difference between our $\mathbf{M}$ for $n_1+n_2=2$ and the corresponding matrix of Refs. 3(a) and 3(b) is only in the diagonal elements and this difference is given by Eq. (79). For the 1:1 degeneracy and $n_1+n_2=2$, the eigenvalues e_i of the matrix $\mathbf{M}$ are given by

$$e_1 = -(\hbar^2/8)(27\beta^2 + 54\beta\gamma + 71\gamma^2),$$
$$e_2 = -(\hbar^2/2)(A + B^{1/2}), \quad e_3 = -(\hbar^2/2)(A - B^{1/2}),$$

where

$$A = -\tfrac{1}{4}(17\beta^2 + 30\beta\gamma + 101\gamma^2),$$
$$B = -\tfrac{1}{4}(41\beta^4 - 446\beta^2\gamma^2 + 2025\gamma^4 - 16\beta^3\gamma).$$

TABLE I. Comparison of energies of the anharmonic oscillator. Here E_{QNF} is obtained by using Eqs. (71)–(74), E_{torus} is obtained by substituting $(n+\frac{1}{2})$ for τ in the BGNF equations [(31)–(34)], and E_{exact} is obtained by diagonalizing a 200×200 matrix formed by the matrix elements of the operator $\mathbf{H}$ of Eqs. (59)–(61) in the harmonic oscillator basis with variable exponent. The E_{QNF}, E_{torus}, and E_{exact} correspond to $\omega=1$, $b=0$, and $g=0.01$, while E^a_{QNF}, E^a_{torus} and E^a_{exact} correspond to $\omega=1$, $b=-0.02$, and $g=0.02$.

n	E_{QNF}	E_{torus}	E_{exact}	E^a_{QNF}	E^a_{torus}	E^a_{exact}
0	0.507 24	0.5037	0.507 26	0.5068	0.5033	0.5068
1	1.535 4	1.5323	1.535 6	1.5328	1.5297	1.5328
2	2.589 8	2.5871	2.590 8	2.5839	2.5811	2.5840
3	3.667 8	3.6655	3.671 1	3.6590	3.6565	3.6591
4	4.766 9	4.7650	4.774 9	4.7567	4.7545	4.7570
5	5.884 5	5.8830	5.901 0	5.8758	5.8739	5.8766
6	7.018 1	7.0170	7.048 3	7.0151	7.0135	7.0170

For the special case ($\hbar = 1$, $\beta = 1/\sqrt{80}$, $\gamma = -1/3\sqrt{80}$), considered by Refs. 3(a) and 3(b), we have the six lowest energy levels

$$E_1 = 1 + \langle 00|\mathbf{K}_2/2|00\rangle = \tfrac{719}{720}, \tag{80}$$

$$E_2 = E_3 = 2 + \langle 01|\mathbf{K}_2/2|10\rangle = 2 + \langle 10|\mathbf{K}_2/2|01\rangle = \tfrac{1433}{720}, \tag{81}$$

$$E_4 = 3 + e_3 = \tfrac{2129}{720}, \quad E_5 = E_6 = 3 + e_1 = 3 + e_2 = \tfrac{215}{72}. \tag{82}$$

The results of Eqs. (80)–(82) are compared in Table II with those reported by Robnik.[3(a)] We believe that there is a small error in Eq. (3.17) of Ref. 3(b) [the number -2 should read as -4 for Eq. (3.17) to be consistent with Eq. (2.4) of Ref. 3(b)]. It is clear from Table II that the difference between the QNF and quantized BGNF results is more pronounced for $n = 0$. This difference depends, Eqs. (76)–(79), on the values of ω_1, ω_2, β, and γ. It may be remarked that the choice $\tau_i^2 = (n_i + \frac{1}{2})^2$ is not unique since other choices of the quantum operators for τ_i^2 also preserve the Poisson and BGNF structures. In the QNF approach, the quantization of τ_i^2 is achieved in a consistent way by successive replacements of the Poisson brackets that are used to generate the BGNF. It is clear from the contents of Secs. III and IV A, that the QNF results preserve the Poisson and BGNF structures. A study of the data in Table II also shows the slight edge of the QNF results over other quantum calculations at and near the ground state.

C. The QNF is identical to the RSP

We calculated the Rayleigh–Schrödinger perturbation (RSP) energies through fourth order for the anharmonic oscillator and second order for the resonant and nonresonant Henon–Heiles system. These perturbation energies were found identical to the corresponding QNF results given in Eqs. (62)–(64) and (71)–(74). It is not hard to see the reason for the agreement between these results. The RSP expression for the energy of the nth perturbed state can be written (for simplicity, assume nondegenerate states) as

$$H = H_0 + V, \quad H|\psi\rangle = E|\psi\rangle, \quad H_0|\phi_m\rangle = E_m|\phi_m\rangle,$$
$$|\phi_n\rangle = P|\psi\rangle, \quad Q = I - P,$$
$$P = |\phi_n\rangle\langle\phi_n|, \quad Q = \sum_{m\neq n} |\phi_m\rangle\langle\phi_m|, \quad \langle\phi_n|\psi\rangle = 1,$$
$$\Omega = V\sum_m \left\{\frac{Q}{E_n - H_0}(E_n - E + V)\right\}^m.$$

If Ω_p^n is the projection of Ω on the space of $|\phi_n\rangle$, then $\langle\phi_n|\Omega|\psi\rangle = \langle\phi_n|\Omega_p^n|\phi_n\rangle$, where Ω_p^n is diagonal (block diagonal in the degenerate case) in $|\phi_n\rangle$. In the QNF approach, on the other hand, the transformations invoked by Eqs. (43)–(53) yield the operators $\mathbf{K}_{i>0}$ that commute with $\mathbf{K}_0$. Thus $\mathbf{K}_{i>0}$ is diagonal (block diagonal in the degenerate case) in the eigenbasis $|\phi_n\rangle$ and an agreement between QNF and RSP is expected. In our work we have found that, for arbitrary n, the operators $\mathbf{K}_{i>0}$ and Ω_p^n are identical to the same order of approximation. Therefore, the RSP and QNF are two alternate routes to the same goal.

We have also seen in Sec. IV A that the QNF generates the BGNF exactly. In view of the above discussion, the BGNF can therefore be called the classical analog of the RSP. It may be possible to obtain the BGNF from the RSP by taking appropriate classical limits.

V. SUMMARY

By employing the Lie transform, we have presented the BGNF's of a one-dimensional anharmonic oscillator and the two-dimensional Henon–Heiles system. The QNF's of these systems have been obtained by replacing with commutators the Poisson brackets that are present in the Lie series for the BGNF's. This trading of Poisson brackets for commutators has provided an algebraic connection between the BGNF and QNF. Our comparison of the energies, obtained by the QNF and Weyl quantization of the BGNF, shows that the main difference between these two sets of quantization results occurs at and near the ground energy level. The QNF, which in effect corresponds to higher-order quantizations of the BGNF, provides better results for the ground energy. We have demonstrated how to obtain the BGNF from the QNF by replacing the quantum operators by their classical values. We have also shown the equivalence between the QNF and RSP. Thus it may be possible, at least in principle, to obtain the BGNF from the RSP by taking appropriate classical limits. Although our work does not solve the fundamental problem of quantization of classical Hamiltonian systems, it does provide a better understanding of why the Weyl and torus quantizations of the BGNF yield reasonable results, and the algebraic connection between the BGNF and RSP is a new result.

TABLE II. Comparison of 1:1 degenerate energies of the Henon–Heiles system. The energies in columns $E_{\text{exact}}, E_{\text{QNF}}, E^a_{\text{AQ}}, E^b_{\text{AQ}}, E_{\text{SC}}$, and E_{US}, respectively, correspond to (1) exact quantum mechanical variational calculations of Refs. 4(e) and 4(f), (2) QNF results obtained from Eqs. (80)–(82), (3) algebraic quantization results of Ref. 3(a), (4) algebraic quantization results of Ref. 3(b), (5) classical trajectory data with semiclassical quantization of Refs. 4(e) and 4(f), and (6) uniform semiclassical results up to and including K_6 of Refs. 4(a).

States	E_{exact}	E_{QNF}	E^a_{AQ}	E^b_{AQ}	E_{SC}	E_{US}
E_1	0.9986	0.9986	0.9984	0.9948	0.9947	0.9966
$E_2 = E_3$	1.9901	1.9903	1.9901	1.9863	1.9863	1.9868
E_4	2.9562	2.9569	2.9568	2.9532	2.9506	2.9558
$E_5 = E_6$	2.9853	2.9861	2.9859	2.9823	2.9815	2.9812

ACKNOWLEDGMENTS

I would like to thank Arlene McMurray for her patient typing of the manuscript using suitable codes for the laser printer. I would also like to express my thanks to Dr. G. Cardini and Dr. S. F. O'Shea for helpful comments, to Dr. R. L. Somorjai for drawing my attention to the technique of Lie transform, and to Dr. J. S. Devitt for introducing me to the new symbolic computer language MAPLE. I am grateful to the University of Lethbridge for providing excellent computing facilities.

This work has been supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

[1]G. D. Birkhoff, *Dynamical Systems* (Am. Math. Soc., New York, 1927), Vol. IX; V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1980); C. L. Siegel, Ann. Math. **42**, 806 (1941).

[2]F. G. Gustavson, Astron. J. **71**, 670 (1966).

[3](a) M. Robnik, J. Phys. A **17**, 109 (1984); (b) T. Uzer and R. A. Marcus, J. Chem. Phys. **81**, 5013 (1984).

[4](a) C. Jaffé and W. P. Reinhardt, J. Chem. Phys. **77**, 5191 (1982); (b) R. B. Shirts and W. P. Reinhardt, J. Chem. Phys. **77**, 5204 (1982); (c) W. P. Reinhardt and D. Farelly, J. Phys. Paris Colloq. **43**, C2 (1982); (d) R. T. Swimm and J. B. Delos, J. Chem. Phys. **71**, 1706 (1979); (e) D. W. Noid and R. A. Marcus, J. Chem. Phys. **67**, 559 (1977); (f) D. W. Noid, M. L. Koszykowski, M. Tabor, and R. A. Marcus, J. Chem. Phys. **73**, 6169 (1980).

[5]R. C. Churchill, G. Pecelli, and D. L. Rod, in *Lecture Notes in Physics*, Vol. 93 (Springer, Berlin, 1978); R. C. Churchill, M. Kummer, and D. L. Rod (private communication, 1981).

[6]A. Einstein, Verh. Dtsch. Phys. Ges. Berlin **19**, 82 (1917) [English translation by C. Jaffé, is available as JILA Tech. Rep. No. 116 (University of Colorado, Boulder, Sept. 1980)]; M. V. Berry, "Semiclassical mechanics of regular and irregular motion," presented at the Les Houches Summer School, 1981; M. Springborg, J. Phys. A **16**, 536 (1983); R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Benjamin, Reading, MA, 1978); J. Hietarinta, Phys. Lett. A **93**, 55 (1982); I. C. Percival, Adv. Chem. Phys. **36**, 1 (1977); M. C. Gutzwiller, in *Path Integrals*, edited by J. T. Devreese (Plenum, New York, 1978). (Many more references can be traced through the above cited articles.)

[7]See, for example, A. Messiah, Quantum Mechanics (North-Holland, Amsterdam, 1966), Vol. II.

[8]S. Lie, *Theorie der Transformationgruppen I* (Teubner, Leipzig, 1888), Sec. 12.

[9]A. Deprit, Celes. Mech. **1**, 12 (1969).

[10]A. A. Kamel, Celes. Mech. **3**, 90 (1970).

[11]J. Henrard, Celes. Mech. **3**, 107 (1970); R. A. Howland, Jr., *ibid.* **19**, 95 (1979); W. A. Mersman, *ibid.* **3**, 81 (1970); M. Rapaport, *ibid.* **21**, 177 (1980); J. R. Cary, "Hamiltonian perturbation theory using Lie transform," Lawrence Berkeley Laboratory, Berkeley, California, LBL-6310, 1977; G. Hori, Pub. Astron. Soc. Jpn. **18**, 287 (1966); D. M. Greig and M. A. Abdel-Naby, Numer. Math. **34**, 1 (1980).

[12]K. O. Geddes, G. H. Gonnet, and B. W. Char, MAPLE *User's Manual* (Department of Computer Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, 1982).

[13]B. Simon, Ann. Phys. **58**, 76 (1970); G. S. Sharma and L. K. Sharma, J. Math. Phys. **25**, 2947 (1984); F. T. Hioe and E. W. Montroll, *ibid.* **16**, 1945 (1975); P. M. Mathews, M. Seetharaman, S. Raghavan, and V. T. A. Bhargava, Phys. Lett. A **83**, 118 (1981); A. G. Arteca, M. F. Fernández, and A. E. Castro, J. Math. Phys. **25**, 3492 (1984).

[14]M. D. Kruskal, International Conference on Solitons and Coherent Structures, held at the Institute for Theoretical Physics, Santa Barbara, California, 11–16 Jan. 1985.

[15]M. Henon and C. Heiles, Astron. J. **69**, 73 (1964).

[16]See, for example, R. Shankar, *Principles of Quantum Mechanics* (Plenum, New York, 1980).

Regular maximal slicings in asymptotically flat space-times

Mauro Carfora[a)]
Dipartimento di Matematica, "G. Castelnuovo," Università di Roma, Città Universitaria, 00100 Roma, Italy and Istituto Nazionale di Física Nucleare, Sezione di Pavia, Pavia, Italy

(Received 31 December 1984; accepted for publication 31 May 1985)

An estimate is provided for the lapse function defining asymptotically Euclidean maximal slicings in asymptotically flat space-times. This estimate is found to be in agreement with a similar estimate suggested, on heuristic grounds, by Smarr and York. It is also shown that in vacuum space-times the scalar curvature of maximal slices remains uniformly bounded in time provided that suitable conditions on the rate of growth of the (negative) lower bound of the Ricci curvature of the slices are satisfied.

I. INTRODUCTION

Asymptotically Euclidean maximal slicings play a key role in discussing, either analytically or numerically, the dynamics of asymptotically flat space-times. For instance, within the context of the initial value problem they permit a separation of the Hamiltonian and of the momentum constraints. Also, much of the work in numerical relativity relies on the nice kinematical properties of such slicings. Recently, Bartnik[1] (see also Refs. 2 and 3 for a different approach) has proven that only rather mild conditions are necessary in order to ensure the existence of maximal slicings in a given asymptotically flat space-time. This result further emphasizes the importance of discussing the dynamics of maximal slicings as a mirror image of the dynamics of the ambient space-time. An important point in this context is to discuss how much of a globally hyperbolic space-time admitting a maximal slice can be foliated by the one-parameter family of maximal slices generated starting from the given slice. In this connection, a number of well-known examples, among which is the extended maximal slicing of the Schwarzschild–Kruskal space-time found by Estabrook *et al.*,[4] suggest that maximal slices tend to lose their spacelike character, by turning null-like, in regions of strong curvature. As a consequence, in such regions the dynamics will be halted, and the given family of maximal slices will no more be able to cover the whole Cauchy development of the initial data. Evidence for such behavior is also supported by various numerical examples such as the colliding black holes solution of Smarr[5] and Eppley,[6] or by solutions describing imploding gravitational waves.[7] In order to get a better understanding of the relation between the mass–energy content of the slices and their dynamics, Smarr and York[8] discussed a particularly simple model. In this way, they were able to argue that the lapse function N_t defining a maximal slicing $\{S_t\}$ tends to drop to zero (and S_t turns null-like on two-spheres, consequently) in strong field regions according to

$$N_t \simeq \exp(-X_0(t)), \tag{1.1}$$

where the dimensionless parameter $X_0(t)$ is a measure of the strength of the scalar curvature on the slice considered. This conjecture proved to be in remarkable agreement with the actual behavior of the lapse function defining the Estabrook maximal slicing. Estimate (1.1) also allowed us to understand in what circumstances a maximal slicing can turn null-like before hitting a possible singularity or whether it can allow the field equations to evolve the data until the singularity (if any) is reached.

The main purpose of this paper is to provide some general results concerning such topics. Without referring to heuristic models or to particular examples, we show that an estimate similar in structure to (1.1) holds true for the lapse function defining the generic asymptotically Euclidean maximal slicing. Such an estimate reproduces, in a more precise way, the results of Smarr and York. Namely, it shows that maximal slices tend to turn null-like in strong field regions if the product of the square root of a suitable average of the scalar curvature times the "radius" of the region, from which the main contributions to the curvature come, grows sufficiently fast with the coordinate time labeling the slices. A new feature of our estimate is the significant role played in it by the negative part of the Ricci curvature of the slices. This role was not evident in the original analysis of Smarr and York, since that analysis was carried out on a heuristic flat space model. From a geometrical and function-theoretical point of view, such a role cannot be surprising, for the lapse function associated with an asymptotically Euclidean maximal slicing $\{S_t\}$ is, on each S_t, a subharmonic function (i.e., $\Delta N_t \geqslant 0$, where Δ is the Laplace–Beltrami operator on the given S_t). And, as is well known, the existence of subharmonic functions on complete Riemannian manifolds is strictly conditioned by sign restrictions on the Ricci curvature of the manifold (e.g., see Ref. 9). According to our estimate, if the ratio between the (negative) lower bound of the Ricci curvature and the scalar curvature of the slices grows with the time function labeling the slices (which reduces to proper time at spatial infinity) at a sufficiently large rate, then we can avoid a breakdown in the uniform spacelike character of the slicing. This behavior is to be expected, for a growing (negative part of the) Ricci curvature implies that the geometry of the slices is developing large inhomogeneities. And, as already argued by Smarr and York (on the basis of the analysis of marginally bound dust collapse in Bondi–Tolman space-times as discussed by Eardley and Smarr,[10] such inhomogeneities can preserve the uniform spacelike

a) Permanent address: Dipartimento di Fisica Nucleare e Teorica, Università degli Studi di Pavia, Via Bassi 6, 27100 Pavia, Italy.

 0022-2488/85/102573-08$02.50

character of a maximal slicing.

The estimate for the lapse function so obtained yields interesting results when applied to maximal slicings in vacuum space-times. A rather simple argument based on such an estimate and on the evolutive part of Einstein's equations implies that the scalar curvature of asymptotically Euclidean maximal slices stays uniformly bounded above during the evolution, provided that the (negative) lower bound of the Ricci curvature does not grow too fast with time. This result appears to be quite satisfactory, for it provides an analytical form to the well-known singularity-avoidance features of most maximal slicings.

[Notation: for any smooth symmetric bilinear form $\mathbf{A}$, we set $\mathrm{tr}\,\mathbf{A} \equiv A^{i}{}_{i}$, $|\mathbf{A}| \equiv |A_{ik}A^{ik}|^{\frac{1}{2}}$, $(\nabla\cdot\mathbf{A})_k \equiv \nabla^i A_{ik}$, where ∇ denotes the Riemannian connection, and where $i, k, \ldots = 1,2,3$. Riemann tensor sign conventions are fixed by $(\nabla_i\nabla_k - \nabla_k\nabla_i)v^j = R_{ikm}{}^{j}v^m$ with $R_{ik} \equiv R^{m}{}_{imk}$; physical units are taken so that $G = c = 1$.]

II. PRELIMINARIES

Let $(v^4, \mathbf{g})$ be a space-time manifold, a solution of the Einstein field equations $G_{\mu\nu} = 8\pi T_{\mu\nu}$, describing a given gravitating system. According to the (3 + 1) dynamical formulation of general relativity, we may regard $(V^4 \overset{\varphi}{\simeq} S_0 \times I,\ \mathbf{g})$ as the Cauchy development of some regular initial data set $(S_0, \mathbf{h}_0, \mathbf{K}_0)$, where S_0 is a three-manifold carrier of the initial data, φ is a diffeomorphism mapping V^4 onto the product $S_0 \times I$ (I being a suitable subset of $\mathbb{R}$), and $\mathbf{h}_0$, $\mathbf{K}_0$ are tensor fields on S_0 representing, respectively, the first and the second fundamental form associated with the embedding of S in the final space-time.

We assume that the data are evolved from a slice $S_t = \varphi^{-1}(S_0 \times \{t\})$ to the nearby slice $S_{t+\delta t}$ along a nonspacelike congruence Γ. Correspondingly we denote by $\mathbf{n}$ the unit timelike future pointing normal vector field associated with the slicing $\{S_t\}$, by N_t, the lapse function, which is the proper-time normal separation between two nearby slices of $\{S_t\}$, and by $\mathbf{t}$, the nonspacelike future pointing vector field tangent to the t-parametrized lines $C_x = \varphi^{-1}(\{x\} \times I)$ defining the congruence Γ.

As is known (e.g., see Ref. 11), $\mathbf{n}$, n_t, and $\mathbf{t}$ are related to each other via

$$\mathbf{t} = N_t\mathbf{n} + \boldsymbol{\beta}_t\,, \tag{2.1}$$

where $\boldsymbol{\beta}_t : S_t \to TS_t$ is the shift vector field. Notice that as long as N_t is strictly positive, $\mathbf{t}$ is nowhere tangent to S_t.

We denote by $\mathbf{h}_t$ the Riemannian three-metric on a given slice S_t, by $\mathbf{K}_t$ the rate-of-deformation tensor associated with $\mathbf{n}$ (the second fundamental form corresponding to the embedding of S_t in V^4), by $\tilde{\mathbf{K}}_t$ its trace-free part (the shear tensor), and by $k_t = \mathrm{tr}(\mathbf{K}_t)$ the mean extrinsic curvature of S_t in V^4. Locally, in terms of the fields N_t, $\mathbf{n}$, $\mathbf{t}$, $\boldsymbol{\beta}_t$ introduced above, we can write

$$\mathbf{g} = (\mathbf{h}_t)_{ik}\,dx^i\,dx^k + 2(\boldsymbol{\beta}_t)_i\,dx^i\,dt - (N_t^2 - |\boldsymbol{\beta}_t|^2)dt^2\,, \tag{2.2}$$

$$\begin{aligned}\mathbf{K}_t &= -\tfrac{1}{2}L_{\mathbf{n}}(\mathbf{h}_t)_{ik}\,dx^i\,dx^k \\ &= -\tfrac{1}{2}N_t^{-1}(L_{\mathbf{t}}(\mathbf{h}_t)_{ik} - L_{\boldsymbol{\beta}_t}(\mathbf{h}_t)_{ik})dx^i\,dx^k\,,\end{aligned} \tag{2.3}$$

$$k_t = N_t^{-1}(\nabla_t\cdot\boldsymbol{\beta}_t - L_{\mathbf{t}}(\log(\det\mathbf{h}_t)^{1/2}))\,, \tag{2.4}$$

where L denotes Lie differentiation along the vector field indicated.

We assume the usual falloff conditions on $\mathbf{h}_t$, $\mathbf{K}_t$, and the matter fields associated with the boundedness of the total four-momentum of the system described by $(V^4, \mathbf{g})$. Assuming (for the moment) that each slice S_t is diffeomorphic to $\mathbb{R}^3$ and denoting by r the Euclidean distance in the region external to some compact set, we require that there exist in S_t asymptotic Cartesian coordinate charts in which

$$\begin{aligned}&D^a((\mathbf{h}_t)_{ik} - e_{ik}) = O(r^{-1-|a|})\,,\\ &D^a((\mathbf{K}_t)_{ik}) = O(r^{-2-|a|})\,,\quad D^a\mu_t = O(r^{-4-|a|})\,,\\ &D^a(\mathbf{J}_t)^i = O(r^{-4-|a|})\,,\quad D^a(\mathbf{s}_t)_{ik} = O(r^{-4-|a|})\,,\end{aligned} \tag{2.5}$$

where a is a multi-index [i.e., a triple of non-negative integers (a_1, a_2, a_3), with $|a| = a_1 + a_2 + a_3$, and $D^a = \partial_1^{a_1}\partial_2^{a_2}\partial_3^{a_3}$]. The tensor fields $\mathbf{e}$, μ_t, $\mathbf{J}_t$, and $\mathbf{s}_t$ introduced in (2.5) are, respectively, the Euclidean metric, the energy density, the density of momentum, and the density of momentum current of the external sources relative to a system of ∞^3 observers momentarily at rest on the given S_t. We also identify the parameter t with the proper time at spatial infinity, and thus we assume that

$$D^a(N_t - 1) = O(r^{-1-|a|})\,, \tag{2.6}$$

together with

$$D^a(\boldsymbol{\beta}_t)^k = O(r^{-1-|a|})\,. \tag{2.7}$$

Finally notice that by suitably utilizing the freedom in choosing the diffeomorphism φ mapping V^4 onto $S_0 \times I$, we can always assume that $D^a k_t = O(r^{-3-|a|})$ (see Ref. 2). (For more details on this point as well as for the formulation of the above asymptotic conditions either in terms of the weighted Sobolev spaces of Niremberg–Walker–Cantor, or in terms of the weighted Hölder spaces introduced by Choquet-Bruhat, see Refs. 12 and 2, respectively.)

III. MAXIMAL SLICINGS

As a basis for subsequent discussions we now assume that the initial submanifold S_0 is maximal, i.e., $k_0 = 0$, and examine under what conditions such maximal character can be maintained in time.

As is known (see, e.g., Ref. 2), the rate of variation of the mean extrinsic curvature k_t in passing from a slice S_t to the nearby slice $S_{t+\delta t}$, along the lines of Γ, is provided, independently of the field equations, by

$$L_{\mathbf{t}}k_t = -\Delta_t N_t + N_t(|\mathbf{K}_t|^2 + {}^{(4)}\mathrm{Ric}(\mathbf{n},\mathbf{n})) + L_{\boldsymbol{\beta}_t}k_t\,, \tag{3.1}$$

where ${}^{(4)}\mathrm{Ric}(\,,\,)$ is the Ricci curvature tensor of the enveloping space-time, and where Δ_t is the Laplace–Beltrami operator in the given S_t [in a local coordinate neighborhood $\Delta_t = (\det\mathbf{h}_t)^{-1/2}\partial_i((\det\mathbf{h}_t)^{1/2} \times (\mathbf{h}_t)^{ij}\partial_j)$].

Taking into account the field equations, (3.1) reduces to

$$L_{\mathbf{t}}k_t = -\Delta_t N_t + N_t(|\mathbf{K}_t|^2 + \tfrac{1}{2}(\mu_t + \mathrm{tr}(\mathbf{s}_t))) + L_{\boldsymbol{\beta}_t}k_t\,.$$

Thus, if $k_0 = 0$, we can propagate the maximal character of the embedding along Γ provided that the lapse function N_t defining the slicing $\{S_t\}$ is a solution, on each slice S_t, of the second-order elliptic partial differential equation

$$\Delta_t N_t - N_t(|\mathbf{K}_t|^2 + \tfrac{1}{2}(\mu_t + \mathrm{tr}(s_t))) = 0\,, \tag{3.2}$$

with the condition $N_t > 0$.

In order to provide a physical interpretation to (3.2), we may consider the maximal foliation $\{S_t\}$ resulting from solving (3.2) on each S_t, as defining an irrotational fluid of reference $\mathscr{R}$. Then, we can associate with $\{S_t\}$ the scalar field $U_t \equiv -\log N_t$, and the spacelike vector field $\mathbf{G} \equiv \nabla U_t$, in which we easily recognize the relative scalar acceleration potential and the relative dragging gravitational field with respect to $\mathscr{R}$, respectively. In terms of U_t, Eq. (3.2) can be written as

$$\Delta_t U_t + H_t - |\nabla U_t|^2 = 0\,, \tag{3.3}$$

where we have set for convenience $H_t \equiv (|\mathbf{K}_t|^2 + \tfrac{1}{2}(\mu_t + \mathrm{tr}(\mathbf{s}_t)))$.

Equation (3.3) simply tells us what instantaneous acceleration ($\mathbf{A} = -\mathbf{G}$) must be provided to the observers of $\mathscr{R}$ in order to balance the focusing effect of matter and gravitation in such a way as to maintain, in time, the incompressibility condition $k_t = 0$. One should note that, according to Eq. (3.3), such an acceleration results from an instantaneous balance between the focusing inducing terms $(\mu_t + \mathrm{tr}(s_t))$ and $|K_t|^2$, and the "energy density" $|\nabla U_t|^2$ associated with the defocusing acceleration field itself.

It is quite simple to provide conditions ensuring the solvability of problem (3.2) on each given S_t. To this effect, we can apply a general theorem proved by Fischer-Colbrie and Schoen,[13] concerning the properties of the operator $(\Delta - W)$, with W a given smooth function, on complete, noncompact, manifolds (equivalently, we could apply a theorem by Cantor and Brill,[14] more specifically concerned with the Laplacian on asymptotically Euclidean manifolds). In this way, we obtain that (3.2) admits, in the given S_t, a positive solution satisfying the given asymptotic conditions if and only if (iff) the first eigenvalue of the linear operator $L \equiv \Delta_t - H_t$ is non-negative on each compact domain D of S_t. Namely iff

$$\inf\left(\int_D (H_t f^2 + |\nabla f|^2) dv_t\colon \mathrm{supp}\, f \subset D\,,\ \int_D f^2\, dv_t = 1\right) = \lambda_1(D) \geqslant 0\,,$$

or equivalently iff

$$-\int_{S_t} H_t f^2\, dv_t \leqslant \int_{S_t} |\nabla f|^2\, dv_t\,, \tag{3.4}$$

for any function $f \not\equiv 0$, with compact support.

Condition (3.4) is trivially satisfied if the external sources are such that the strong energy condition

$$\mu_t + \mathrm{tr}(\mathbf{s}_t) \geqslant 0 \tag{3.5}$$

holds true. However, it is clear that (3.4) does not require the pointwise validity of the strong energy condition in order that problem (3.2) admits a well-behaved solution. Rather, what we need is only the validity of (3.5) in an averaged sense. To be more precise, by an argument similar to that in Ref. 14, it can be shown that condition (3.4) is satisfied if

$$\left(\int_{S_t^-} |H_t|^{3/2}\, dv_t\right)^{2/3} < C\,,$$

where C is a large positive constant (that can be estimated), and where S_t^- denotes the union of those regions of S_t where $H_t < 0$. This latter circumstance is physically significant, as it implies that the maximal slicing condition also can be propagated in space-time regions where (3.5) does not hold true. This can be the case, for instance, when the external sources are represented by a massive scalar field. Similarly, if we consider high-density regimes in spontaneously broken gauge theories, it is possible that the strong energy condition is violated (as Tipler[15] argued, such violations are likely to be only local, with the strong energy condition holding in an averaged sense). However interesting it may be, the analysis of the behavior of maximal slicings in such regions is quite difficult [essentially because if (3.5) does not hold true pointwise, an N_t solution of (3.2), does not obey the maximum principle]. Thus, we shall assume in the following that the strong energy condition is satisfied.

From the above analysis it follows that by imposing the condition $k_0 = 0$, $L_t k_t = 0$, on the initial and on all subsequent slices, respectively, and by solving the associated problem (3.2), we obtain a one-parameter family, (N_t), of scalar functions determining a maximal slicing $\{S_t\}$ in a suitable neighborhood of the initial slice S_0. However, this procedure (which can be numerically implemented on a computer), cannot, in general, provide a well-behaved slicing $\{S_t\}$ as t grows larger and larger, since we are not ensured, at least not *a priori*, that as t grows, the slices S_t remain uniformly spacelike or singularity-free. Leaving aside, for the moment, the question of the occurrence of singularities, let us fix our attention on the maintenance in time of the uniform spacelike character of the maximal slices generated by (3.2).

As is known (see, e.g., Ref. 16) a one-parameter family of Riemannian three-metrics $\mathbf{h}_t$, induced on S_0 by a Lorentzian metric $\mathbf{g}$ by means of a smooth curve of embeddings $i_t\colon S \to V^4$, is said to be uniformly spacelike if for any uniformly timelike vector field $\mathbf{v}\colon V^4 \to TV^4$ [i.e., for any $\mathbf{v}$ such that $\mathbf{g}(\mathbf{v},\mathbf{v}) = -\gamma$, with $\gamma \geqslant a > 0$], we have $\mathbf{g}(\mathbf{v},\mathbf{n}) = -\alpha$ with α uniformly bounded above. In our case $\mathbf{g}(\mathbf{v},\mathbf{n}) = -N_t^{-1}(\mathbf{v}\cdot(\mathbf{t} - \boldsymbol{\beta}_t))$, and for $\mathbf{v}$ uniformly timelike and $(\mathbf{t} - \boldsymbol{\beta}_t)$ nonspacelike, the scalar product $(\mathbf{v}\cdot(\mathbf{t} - \boldsymbol{\beta}_t))$ is always uniformly bounded away from zero. Thus, the uniform spacelike character of the slicing $\{(S_t, \mathbf{h}_t)\}$ is ensured as long as the corresponding lapse function N_t remains uniformly bounded away from zero as the evolution goes on. Namely if, associated with the family of functions (N_t), solutions of the corresponding problem (3.2) on the successive S_t, there is a positive constant C such that

$$N_t(x) \geqslant C > 0\,, \tag{3.6}$$

for all $t \in \mathbf{R}^+$, and for all $x \in S_t$.

Geometrically, condition (3.6) implies that as t grows, the successive slices S_t remain uniformly separated. It follows, in such a case, that the description of the space-time geometry from the frame $\mathscr{R}$ associated with $\{S_t\}$, has an infinite proper-time duration (if the evolution process is not halted by the occurrence of singularities). Conversely, if condition (3.6) does not hold true, namely if $(\inf N_t(x))_{x \in S_t}$ decreases sufficiently fast with t, then, regardless of the possible occurrence of singularities, the description of the space-time geometry from $\mathscr{R}$ may have a finite proper-time

duration.

In order to check if condition (3.6) holds true or not, it is necessary to provide an estimate for $(\inf N_t)_{x\in S_t}$ in terms of the instantaneous values of the fields $\mathbf{K}_t, \mu_t$, and $\mathbf{s}_t$. This task will occupy us for most of the remaining part of this paper. We shall adapt to our particular situation an elegant technique devised by Yau.[17,18] The estimate for $\inf N_t$ will be derived in three steps. First, we use Yau's results in order to derive bounds for the magnitude of $|\nabla U_t| = |\nabla N_t|/N_t$ (the acceleration in the frame $\mathscr{R}$), in terms of the instantaneous values of the fields $\mathbf{K}_t, \mu_t$, and $\mathbf{s}_t$. Then, such estimates are used to associate with any solution of problem (3.3) a gradient bound of the form $|\nabla U_t| \leqslant \alpha_t^{-1} U_t$ with the constant α_t explicitly provided. This gradient bound together with some easy manipulation of (3.3) will finally yield the required estimate for $\inf N_t$.

IV. AN ESTIMATE FOR THE LAPSE FUNCTION

Let f be a smooth nonconstant positive function defined on a complete Riemannian three-manifold $(S,\mathbf{h})$. Correspondingly let us consider the function

$$F \equiv |\nabla \log f|^2 .$$

A direct calculation (see Ref. 17) yields

$$\nabla F = 2(\nabla f \cdot \mathrm{Hess} f) f^{-2} - 2(|\nabla f|^2/f^3)\nabla f, \tag{4.1}$$

$$\Delta F = 2f^{-2}|\mathrm{Hess} f|^2 + 2f^{-2}\nabla f\cdot\nabla(\Delta f) + 2f^{-2}\mathrm{Ric}(\nabla f,\nabla f) - 2f^{-2}|\nabla f|^2(\Delta f/f) - 4f^{-1}(\nabla F\cdot\nabla f) - 2F^2 , \tag{4.2}$$

where $(\mathrm{Hess} f)_{ik} \equiv \nabla_i\nabla_k f$, and $\mathrm{Ric}(\ ,\ )$ denote, respectively, the Hessian of f and the Ricci curvature associated with $(S,\mathbf{h})$. The philosophy behind (4.1) and (4.2) is that they allow us to estimate F provided that we have suitable information on Hess f, $\nabla(\Delta f)$, and on the Ricci curvature of the ambient manifold. [Notice that (4.2) is essentially the well-known formula of Bochner–Lichnerowicz–Weitzenböck, e.g., see Ref. 19.]

First step: Since the strong energy condition (3.5) has been assumed to hold true, it easily follows from the maximum principle[20] that on each S_t, an N_t solution of problem (3.2) is such that

$$0 < c_t \leqslant N_t \leqslant 1 \tag{4.3}$$

[notice, again, that the lower boundedness in (4.3), a consequence of the maximum principle, is not, *a priori*, uniform in time]. Thus N_t is a smooth nonconstant positive function on each $(S_t,\ \mathbf{h}_t)$, and we can consider (4.1) and (4.2) for $f = N_t$ and $(S,\mathbf{h}) = (S_t,\mathbf{h}_t)$. [In this case $F = |\nabla U_t|^2$ reduces to the squared magnitude of the four-acceleration in the frame $\mathscr{R}$ associated with $\{S_t\}$.]

If we take into account Eq. (3.2), then (4.2) yields in this particular case

$$\Delta F = 2N_t^{-2}|\mathrm{Hess}\, N_t|^2 + 2N_t^{-2}\,\mathrm{Ric}(\nabla N_t,\nabla N_t) + 2N_t^{-1}(\nabla N_t\cdot\nabla H_t) - 4N_t^{-1}(\nabla F\cdot\nabla N_t) - 2F^2 . \tag{4.4}$$

The term in (4.4) involving the Hessian of N_t can be estimated according to the following argument. Let ξ_i, $i = 1,2,3$, denote the eigenvalues of the Hessian at the generic point $x\in S_t$, and let ξ_1 be the largest (positive) eigenvalue [that $\xi_1 \geqslant 0$, follows by observing that $\mathrm{tr}(\mathrm{Hess}\, N_t) = \Sigma_i\xi_i = \Delta N_t = H_t N_t \geqslant 0$]. Since $(\mathrm{Hess}\, N_t(\nabla N_t,\nabla N_t)) = \Sigma_i\xi_i|\nabla^{(i)}N_t|^2 \leqslant \xi_1|\nabla N_t|^2$, we have

$$\xi_1 \geqslant |\nabla N_t|^{-2}\mathrm{Hess}(\nabla N_t,\nabla N_t) .$$

Hence, taking into account (4.1),

$$\xi_1 \geqslant N_t^{-1}|\nabla N_t|^2 + \tfrac{1}{2}N_t^2|\nabla N_t|^{-2}(\nabla F\cdot\nabla N_t) . \tag{4.5}$$

On the other hand, since $|\xi_2|_+ |\xi_3| \geqslant \xi_1 - \mathrm{tr}(\mathrm{Hess}\, N_t)$, we have the following chain of inequalities:

$$(\xi_2^2 + \xi_3^2)^{1/2} \geqslant (|\xi_2| + |\xi_3|)/\sqrt{2} \geqslant (\xi_1 - \mathrm{tr}(\mathrm{Hess}\, N_t))/\sqrt{2}.$$

If $\xi_1 \geqslant \mathrm{tr}(\mathrm{Hess}\, N_t)$, this implies

$$|\mathrm{Hess}\, N_t|^2 = \xi_1^2 + \xi_2^2 + \xi_3^2 \geqslant \tfrac{1}{2}(\xi_1 - H_t N_t)^2 + \xi_1^2 \geqslant \tfrac{3}{2}\xi_1^2 - \tfrac{1}{2}(H_t^2 N_t^2) .$$

This inequality is satisfied also if $\mathrm{tr}(\mathrm{Hess}\, N_t) > \xi_1$, since we can write

$$|\mathrm{Hess}\, N_t|^2 = \tfrac{1}{2}\xi_1^2 + (\xi_2^2 + \xi_3^2) - \tfrac{1}{2}\xi_1^2 \geqslant \tfrac{3}{2}\xi_1^2 - \tfrac{1}{2}(\mathrm{tr}(\mathrm{Hess}\, N_t))^2 .$$

Thus, taking into account (4.5), we get

$$|\mathrm{Hess}\, N_t|^2 \geqslant \tfrac{3}{2}(N_t^{-2}|\nabla N_t|^4 + \tfrac{1}{4}N_t^4|\nabla N_t|^{-4}(\nabla F\cdot\nabla N_t)^2 + N_t(\nabla F\cdot\nabla N_t)) - \tfrac{1}{2}(H_t^2 N_t^2) .$$

Hence

$$N_t^{-2}|\mathrm{Hess}\, N_t|^2 \geqslant \tfrac{3}{2}F^2 + \tfrac{3}{2}N_t^{-1}(\nabla F\cdot\nabla N_t) - \tfrac{1}{2}(H_t^2) ,$$

which, when introduced in (4.4), yields

$$\Delta F \geqslant F^2 - H_t^2 + 2N_t^{-2}\mathrm{Ric}(\nabla N_t,\nabla N_t) + 2N_t^{-1}(\nabla N_t\cdot\nabla H_t) - N_t^{-1}(\nabla N_t\cdot\nabla f) . \tag{4.6}$$

This result allows us to get either an integral or a pointwise estimate for $|\nabla U_t|$.

We are not going to make much use of the pointwise estimate in what follows, however, it may be worthwhile to spend a few lines in deriving it, since it is useful in getting an idea of the magnitude that the upper bound of the four-acceleration $|\nabla U_t|$ can attain in $\{S_t\}$ as the evolution goes on.

Let us start by noticing that if q is a point where N_t attains, in the given S_t, an extremum, then $\lim_{x\to q} F(x) = 0$. Furthermore, according to the asymptotic behavior of N_t [see (2.6)], we also have $\lim_{x\to\infty} F(x) = 0$. Thus, F has to attain at least one positive maximum (and possibly more) at some interior point $p\in S_t$; that is $\nabla F(p) = \mathbf{0}, \Delta F(p) \leqslant 0$. Under these conditions, estimate (4.6) yields

$$(F^2 - H_t^2 + 2F^{1/2}(\mathbf{a}\cdot\nabla H_t) + 2F\,\mathrm{Ric}(\mathbf{a},\mathbf{a}))(p) \leqslant 0 ,$$

where $\mathbf{a} \equiv (\nabla N_t/|\nabla N_t|)(p)$. This necessarily implies that

$$F^2(p) - 2(\sup \Lambda_t)F(p) - 2(\sup|\nabla H_t|)F^{1/2}(p) - (\sup H_t)^2 \leqslant 0 ,$$

where Λ_t is a non-negative function such that $(-\Lambda_t)$ is the lower bound of the Ricci curvature on the given S_t. Thus we get

$$F(p) = \sup|\nabla U_t|^2 \leqslant A_t^2 , \tag{4.7}$$

where the constant A_t is the positive solution of the quartic equation

$$z^4 - 2(\sup \Lambda_t)z^2 - 2(\sup|\nabla H_t|)z - (\sup H_t)^2 = 0 \tag{4.8}$$

(one easily verifies that, unless H_t vanishes identically, such an equation has one positive solution and at most three negative solutions).

The fact that (4.8) involves taking the supremums of H_t, $|\nabla H_t|$, and μ_t makes estimate (4.7) too sensitive to the detailed local behavior of the above fields. (Imagine, for instance, that μ_t has compact support. In such a case $\sup|\nabla\mu_t|$ and correspondingly $\sup|\nabla H_t|$ and A_t can grow very large.) On the other hand, Eq. (3.3) controlling U_t is elliptic, so we expect that U_t (and $|\nabla U_t|$) is not really so sensitive to the local behavior of H_t and $|\nabla H_t|$ as (4.7) and (4.8) suggest. What we need is a sort of distributional analog of (4.7). This can be obtained from (4.6) by following verbatim a proof similar to that in Ref. 18, pp. 15 and 16. In this way, it follows that for any geodesic ball $B(R)$ in S_t, of radius R, centered at any given point $x_0\in S_t$, there is for $|\nabla U_t|^2$ an L^p estimate of the form

$$\left[\int_{B((1-\beta)R)}|\nabla U_t|^{2p}\,dv_t\right]^{1/p} \leqslant C\left(\int_{B(R)}A_t^p\,dv_t\right)^{1/p}+Cp^{1/2}\left(\int_{B(R)}H_t^p\,dv_t\right)^{1/p} + C\beta^{-2}pR^{-2}(\mathrm{Vol}\,B(R))^{1/p}, \tag{4.9}$$

for any $0<\beta<1$, and $p\geqslant 3$, and where C is a given numerical constant.

Second step: Since $0<N_t\leqslant 1$, the acceleration potential $U_t=-\log N_t$ is a smooth function, strictly positive at any interior point. [That is, $U_t(x)>0$, $\forall x\in S_t$, $x\notin S_t\cap\mathcal{S}^2$, where $\mathcal{S}^2$ is the two-sphere intercepted at spatial infinity by the given S_t. This easily follows by observing that if $U_t(x)=0$ with x an interior point, then $N_t(x)=1$, implying that such an x is a maximum point for N_t. But, according to the maximum principle and (3.2), N_t cannot attain a maximum at any interior point.] Thus the function

$$Q\equiv|\nabla\log U_t|^2 \tag{4.10}$$

is bounded and globally well-defined on S_t. As for F, it is easily verified that such Q has to attain at least one positive maximum at some interior point. Let q denote any such point, namely $(\nabla Q)(q)=0$, $(\Delta Q)(q)\leqslant 0$. From (4.1) and (4.2) (evaluated for $f=U_t$) we get

$$(\Delta Q)(q)=(2U_t^{-2}|\mathrm{Hess}\,U_t|^2-2Q^2+2U_t^{-3}|\nabla U_t|^4 + 2U_t^{-3}(H_t|\nabla U_t|^2)-2U_t^{-2}(\mathbf{b}\cdot\nabla H_t)|\nabla U_t| + 2U_t^{-2}\,\mathrm{Ric}(\mathbf{b},\mathbf{b})|\nabla U_t|^2)(q), \tag{4.11}$$

where $\mathbf{b}\equiv((\nabla U_t)/(|\nabla U_t|))(q)$. The term in (4.11) involving the Hessian can be estimated as $|\mathrm{Hess}\,U_t|^2(q)=\Sigma_j\lambda_j^2\geqslant\lambda_1^2$, where the λ_j are the eigenvalues of Hess U_t at q. From (4.1) (for $f=u_t$), we easily get that one of such eigenvalues, say λ_1, is provided by $(|\nabla U_t|^2/U_t)(q)$. Thus

$$(|\mathrm{Hess}\,U_t|^2/U_t^2)(q)\geqslant(|\nabla U_t|^4/U_t^4)(q)=Q^2(q).$$

Introducing this in (4.11), we have

$$(U_t^{-1}|\nabla U_t|^3+H_tU_t^{-1}|\nabla U_t|-(\mathbf{b}\cdot\nabla H_t) + \mathrm{Ric}(\mathbf{b},\mathbf{b})|\nabla U_t|)(q)\leqslant 0.$$

This necessarily implies that in a suitable neighborhood $\Omega(q)$ of q,

$$Q^{1/2}(q)=\sup(|\nabla\log U_t|) \leqslant(|\nabla\log H_t|)(x)+(H_t^{-1}A_t|\nabla U_t|)(x), \tag{4.12}$$

for any $x\in\Omega(q)\subset S_t$, and where $\sup(|\nabla\log U_t|)$ denotes the (local) maximum of $|\nabla\log U_t|$ in $\Omega(q)$.

Notice that if there are m maximum points q_i $(i=1\text{–}m)$ for Q, an estimate like (4.12) will hold true in a suitable neighborhood of any such points.

In order to get nontrivial information from (4.12), we need to eliminate from it either the term $|\nabla U_t|$ or the specification of the region(s) $\Omega(q)$, on the location and extension of which we know nothing. To this end, let $y\in\Omega(q)$ (e.g., $y=q$), and let $B(r)$ and $B(R)$ denote any two geodesic balls, centered at y, and such that $B(r)\subset\Omega(q)\subset B(R)$. If we integrate (4.12) over $B(r)$ we get

$$\sup(|\nabla\log U_t|) \leqslant\langle|\nabla\log H_t|\rangle_{B(r)}+\langle(H_t^{-1}A_t|\nabla U_t|)\rangle_{B(r)}, \tag{4.13}$$

where

$$\langle f\rangle_{B(r)}\equiv(\mathrm{Vol}\,B(r))^{-1}\left(\int_{B(r)}f\,dv_t\right) \tag{4.14}$$

denotes the average value, over $B(r)$, of the quantity enclosed within the angular brackets. On applying Schwarz's inequality (in its average value form) to the second average appearing in (4.13), we get

$$\sup(|\nabla\log U_t|) \leqslant\langle|\nabla\log H_t|\rangle_{B(r)} + (\langle(A_t/H_t)^2\rangle_{B(r)})^{1/2}(\langle|\nabla U_t|^2\rangle_{B(r)})^{1/2}.$$

Since $(\langle|f|^m\rangle_{B(r)})^{1/m}\leqslant(\langle|f|^s\rangle_{B(r)})^{1/s}$ for $s>m$, the above result yields

$$\sup(|\nabla\log U_t|) \leqslant\langle|\nabla\log H_t|\rangle_{B(r)} + (\langle(A_t/H_t)^2\rangle_{B(r)})^{1/2}(\langle|\nabla U_t|^{2p}\rangle_{B(r)})^{1/2p},$$

for any $p\geqslant 1$. If we choose $p\geqslant 3$, rewrite r as $r=R(1-\beta)$, with $0<\beta\equiv((R-r)/R)<1$, and take into account the L^p estimate (4.9), then we have

$$\sup(|\nabla\log U_t|) \leqslant\langle|\nabla\log H_t|\rangle_{B(r)} + C(\mathrm{Vol}\,B(r))^{-1/2p}(\langle(A_t/H_t)^2\rangle_{B(r)})^{1/2} \times\left[\left(\int_{B(R)}A_t^p\,dv_t\right)^{1/p}+p^{1/2}\left(\int_{B(R)}H_t^p\,dv_t\right)^{1/p} + p(R-r)^{-2}(\mathrm{Vol}\,B(R))^{1/p}\right]^{1/2}. \tag{4.15}$$

In this latter expression, we can take the limit for R tending to infinity, while keeping r fixed. Since $p\geqslant 3$, and $(\mathrm{Vol}\,B(R))\leqslant M(R^3\exp(\sqrt{2}\Lambda_t R))$ for some constant M (see Refs. 18 and 21) with $\Lambda_t=O(R^{-3})$, we have $\lim_{R\to\infty}[(\mathrm{Vol}\,B(R))^{1/p}/(R-r)^2]=0$, and (4.15) reduces to

$$\sup(|\nabla \log U_t|) \leqslant \langle |\nabla \log H_t| \rangle_{B(r)} + C(\mathrm{Vol}\, B(r))^{-1/2p} (\langle (\Lambda_t/H_t)^2 \rangle_{B(r)})^{1/2} \times \left[\left(\int_{S_t} \Lambda_t^p \, dv_t \right)^{1/p} + p^{1/2} \left(\int_{S_t} H_t^p \, dv_t \right)^{1/p} \right]^{1/2}. \quad (4.16)$$

This holds true for any $B(r) \subset \Omega(q)$.

Estimate (4.16) still depends on the specification of the region $\Omega(q)$. In order to get rid of this last constraint we can proceed as follows. Let $\bar{q}_i$ be the points in S_t where $|\nabla \log H_t|$ attains its maxima (if some $\bar{q}_i$ is degenerate, replace $|\nabla \log H_t|$ by a smooth approximating function without degenerate critical points; notice, however, that the following argument can be readily generalized to the case where some $\bar{q}_i$ are degenerate). We can easily verify that each of such $\bar{q}_i$ belongs to, at least, one of the regions $\Omega(q_i)$, where (4.16) holds true. This is trivially the case if $\Lambda_t = 0$. If $\Lambda_t \neq 0$, and $\bar{q}_i \notin \Omega(q_i)$, then there must exist a neighborhood of the given $\bar{q}_i$, Σ, such that for all $B(r) \subset \Sigma$, (4.16) holds true with the inequality sign reversed. This yields a contradiction, since by choosing a $B(r) \subset \Sigma$, with r sufficiently small, we can make the right-side member of (4.16) arbitrarily large, while, by standard elliptic theory, $\sup(|\nabla \log U_t|)$ is bounded. Thus, estimate (4.16) holds true with $B(r)$ replaced by a suitably small geodesic ball centered at $\bar{q}_i$. More in general, we get that corresponding to the m local maxima $\bar{q}_i$ of $|\nabla \log H_t|$, there are m estimates of the form

$$\sup_i(|\nabla \log U_t|) \leqslant \langle |\nabla \log H_t| \rangle_{B(r_i)} + C(\mathrm{Vol}\, B(r_i))^{-1/2p} (\langle (\Lambda_t/H_t)^2 \rangle_{B(r_i)})^{1/2} \times \left[\left(\int_{S_t} \Lambda_t^p \, dv_t \right)^{1/p} + p^{1/2} \left(\int_{S_t} H_t^p \, dv_t \right)^{1/p} \right]^{1/2}, \quad (4.17)$$

where the various r_i depend on the given point $\bar{q}_i$. Each one of such estimates holds true in a suitable region $\Sigma_i \supseteq B(r_i)$, centered around the corresponding $\bar{q}_i$.

In an actual calculation, when evaluating the right side of (4.17), the optimal choice for the size of the balls $B(r_i)$ and of the regions Σ_i is naturally suggested by the gravitational configuration under study. And the calculations are much helped by noticing that for p large enough, we may approximate to a very great accuracy the right side of the above estimate with

$$\langle |\nabla \log H_t| \rangle_{B(r_i)} + C(\langle (\Lambda_t/H_t)^2 \rangle_{B(r_i)})^{1/2} \times (\langle \Lambda_t^p \rangle_{B(r_i)}^{1/p} + p^{1/2} \langle H_t^p \rangle_{B(r_i)}^{1/p})^{1/2},$$

where the average values can be conveniently interpreted in a distributional sense. For our actual purposes, the given form (4.17) is more suitable than other forms, and since $(|\nabla \log U_t|)(x) \leqslant \sup_i(|\nabla \log U_t|)$, $\forall x \in \Sigma_i$, we finally get that corresponding to any solution U_t of problem (3.3), we have, in each given Σ_i, the gradient estimate

$$|\nabla U_t|(x) \leqslant [(1 + \rho_t(i)\sigma_t(i))/\rho_t(i)] U_t(x), \quad (4.18)$$

where we have set for convenience

$$\rho_t(i) \equiv (\langle |\nabla \log H_t| \rangle_{B(r_i)})^{-1}, \quad (4.19)$$

and

$$\sigma_t(i) \equiv C(\mathrm{Vol}\, B(r_i))^{-1/2p} (\langle (\Lambda_t/H_t)^2 \rangle_{B(r_i)})^{1/2} \times \left[\left(\int_{S_t} \Lambda_t^p \, dv_t \right)^{1/p} + p^{1/2} \left(\int_{S_t} H_t^p \, dv_t \right)^{1/p} \right]^{1/2}. \quad (4.20)$$

Notice that $\rho_t(i)$ is homogeneous to a length. It provides the length scale of the region, in Σ_i, from where the main contribution to H_t comes (i.e., the inhomogeneity scale of the "source field" H_t). As follows from (4.20), $\sigma_t(i)$ is essentially an L^p average, over $B(r_i)$, of the total (viz. S_t-integrated) source field H_t and of the total negative part of the Ricci curvature. Such an average is weighted by the average value, over the same region, of the ratio (Λ_t/H_t), measuring the relative importance of the "background" negative Ricci curvature with respect to the source field H_t in the given Σ_i. In particular, this implies that whenever $(\Lambda_t/H_t) \ll 1$, in a neighborhood of $\bar{q}_i$, we may correspondingly replace (4.18) by the simpler estimate

$$|\nabla U_t(x)| \leqslant U_t(x)/\rho_t(i),$$

as expected. However, if (Λ_t/H_t) is sufficiently large [i.e., if large anisotropies develop in the underlying geometry of $(S_t, \mathbf{h}_t)$] then we are forced to use (4.18).

Third step: With the above premises, let B_t denote a geodesic ball centered at $\bar{x}$ (the point where H_t attains its local maximum in Σ_i), chosen in such a way as to satisfy the condition

$$(\mathrm{Vol}\, B_t)/(\mathrm{Vol}(\partial B_t)) = \rho_t(i), \quad (4.21)$$

where $\rho_t(i)$ is provided by (4.19). According to the above remarks, such B_t (possibly after a slight homotopic deformation) can be regarded as the region in Σ_i from where the main contribution to H_t comes. Let us apply Gauss' theorem to (3.3) over such B_t

$$-\oint_{\partial B_t} \nabla U_t \cdot ds_t = \int_{B_t} H_t \, dv_t - \int_{B_t} |\nabla U_t|^2 \, dv_t. \quad (4.22)$$

On applying Schwarz's inequality to the surface integral, and by taking into account the gradient estimate (4.18), we get from (4.22)

$$[(1 + \rho_t(i)\sigma_t(i))/\rho_t(i)](\mathrm{Vol}(\partial B_t))\sup_i U_t \geqslant \int_{B_t} H_t \, dv_t - \left[\frac{1 + \rho_t(i)\sigma_t(i)}{\rho_t(i)} \right]^2 (\mathrm{Vol}\, B_t)(\sup_i U_t)^2,$$

that is, by (4.21),

$$\sup_i U_t \geqslant \tfrac{1}{2} [(1 + 4\rho_t^2(i)\langle H_t \rangle_{B_t})^{1/2} - 1]/(1 + \rho_t(i)\sigma_t(i)).$$

Hence, observing that $\inf_i N_t = \exp(-\sup_i U_t)$, we get that, corresponding to any solution N_t of problem (3.2), we have in each Σ_i the pointwise estimate

$$\inf_i N_t \leqslant \exp\{ -\tfrac{1}{2} [(1 + 4\rho_t^2 \langle H_t \rangle_{B_t})^{1/2} - 1]/(1 + \rho_t(i)\sigma_t(i)) \}. \quad (4.23)$$

The above estimate is the result we were looking for. It implies that if the rate of growth with t (which is proper time at spatial infinity) of $(1 + 4\rho_t^2 \langle H_t \rangle_{B_t})^{1/2}/(1 + \rho_t(i)\sigma_t(i))$ is sufficiently large, then N_t cannot remain uniformly bounded away from zero in Σ_i as t increases. By adapting to our situation a remark due to York,[11] we can easily determine the critical rate of growth in order that $\inf_i N_t \to 0$ in a finite amount of proper time. To this end, let $\gamma_t(i)$ denote the

world-line of the point y_t, where the local minimum of N_t, in Σ_i, is attained as t varies [since $(\nabla N_t)(y_t)=0$, such a $\gamma_t(i)$ is the history of an observer of $\mathscr{R}$ in free-fall]. The amount of proper time elapsed from the initial slice S_0 to the slice S_t, evaluated along $\gamma_t(i)$, is provided by

$$T(\gamma_t(i))=\int_{\gamma_t}^{t}{}_{0} N_{t'}(y_{t'})dt' .$$

From (4.23) we get

$$T(\gamma_t(i))\leq\int_{\gamma_t}^{t}{}_{0}\exp\left\{\frac{-\frac{1}{2}\left[(1+4\rho_t^2\langle H_t\rangle_{B_t})^{1/2}-1\right]}{1+\rho_t(i)\sigma_t(i)}\right\}dt' .$$

This is finite as $t\to+\infty$ whenever the exponent appearing in the integrand grows with t as $(1+\epsilon)\log t, \epsilon>0$, or any faster. In such a case, the slices S_t, even if uniformly separated in a neighborhood of spatial infinity (where $N_t\to 1$), tend to come closer and closer to each other, eventually collapsing in Σ_i in a finite amount of proper time. Stated in more geometrical words, we can say that in the above case, the family of embeddings $i_t:(S_t,\mathbf{h}_t)\to(V^4,\mathbf{g})$, obtained by deforming $(S_0,\mathbf{h}_0)$ according to the maximal slicing condition, does not remain uniformly spacelike, as $t\to+\infty$, for the enveloping Lorentzian geometry $(V^4,\mathbf{g})$.

Hence, according to the above remarks, the problem of determining whether or not a maximal slice can turn null-like in a given Σ_i is reduced to the analysis of the rate of growth with time of H_t, Λ_t, $\rho_t(i)$, and $\sigma_t(i)$. When external sources are present, such analysis can be carried out only case by case, when the gravitational configuration under study has been explicitly provided. The best we can do in this situation is to get some qualitative results from (4.23). This latter tells us that the approach to the null cone for a maximal slice is governed either by the growth with time of $(\langle H_t\rangle_{B_t})^{1/2}$, or by a corresponding growth of the dimensions of the region contributing mostly to $\langle H_t\rangle_{B_t}$ [viz. $\rho_t(i)$], or by a combination of both such mechanisms. This is in complete agreement with the results obtained by the heuristic analysis of Smarr and York.[8] Incidentally, notice that a growth of $\rho_t(i)$ on a curved manifold does not necessarily correspond to an expansion of the main support of H_t. Rather, what happens is that the geodesic ball B_t, providing the main support to H_t, will have, as time goes on, increasingly small perimeter but increasing volume; in other words, space gets "sucked into" B_t. As is known,[4,22] this behavior is quite typical of maximal slices: the matter content of the slice, at sufficiently late times, ends up at the bottom of a deep cylindrical valley. The proper radius of such a cylinder tends to be constant, while the depth of the valley increases with time. Estimate (4.23) further suggests that if large anisotropies develop [i.e., if (Λ_t/H_t) and correspondingly $\sigma_t(i)$ steadily increases with time at sufficiently large rates], then N_t may remain bounded away from zero and the slices may be able to reach a possible singularity. This last behavior is confirmed again by some examples (due to Eardley and Smarr[10]) describing (highly inhomogeneous) marginally bound dust collapses in Bondi–Tolman space-times.

V. THE VACUUM CASE

The above results confirm how far-reaching was the simple analysis of Smarr and York. However, such results do not provide us with more specific information concerning the actual rate of growth with time of $\langle H_t\rangle_{B_t}$, $\rho_t(i)$, and $\sigma_t(i)$, since the large variety of possible external sources makes any such effort practically hopeless. Luckily, the situation is not so bad in the vacuum case. In this case H_t reduces to $|\mathbf{K}_t|^2$, and according to some remarks of Marsden and Tipler (see Ref. 3, especially the remarks associated with Theorem 5), it is quite unlikely that $|\mathbf{K}_t|^2$, for maximal slicings, can grow unboundedly large. We rather expect that $|\mathbf{K}_t|^2$ remains uniformly bounded in time. This conjuncture is strongly supported by a well-known result of Cheng and Yau,[9] stating that for minimal hypersurfaces in E_n a uniform bound for the squared second fundamental form of the embedding is automatically satisfied. The evidence that this result can be somehow extended to maximal slices in Lorentzian manifolds is also supported by the dynamics of the Schwarzschild–Kruskal space-time as described by the frame of reference associated with the extended maximal slicing of Estabrook *et al.*[4] As is known, such maximal slicing is generated, starting from the time-symmetry slice in the Kruskal diagram, by solving (3.2) with the inner (Neumann-type) boundary condition $\nabla_r N_t=0$ on the throat ($r=2M$, r being the standard Schwarzschild radial coordinate, and M being the mass associated with the solution). In this case, $|\mathbf{K}_t|^2$ is found to grow from zero (on the inital slice) to a maximum value of $(8M^2/9)$ at late times. Actually, by making use of estimate (4.23) it is quite simple to prove that $|\mathbf{K}_t|^2$ remains uniformly bounded in time. To see this, let us explicitly write down the evolutive part of Einstein equations in the absence of external sources. Assuming zero shift, we have

$$L_t(\mathbf{h}_t)_{ab}=-2N_t(\mathbf{K}_t)_{ab} , \tag{5.1}$$

$$L_t(\mathbf{K}_t)_{ab}=-(\mathrm{Hess}\,N_t)_{ab}+N_t\left[(\mathrm{Ric}(\mathbf{h}_t))_{ab}-2(\mathbf{K}_t)_{ac}(\mathbf{K}_t)^c_b\right] . \tag{5.2}$$

From (5.1) and (5.2) we get

$$L_t(|\mathbf{K}_t|^2)=-2(\mathbf{K}_t)_{ab}(\mathrm{Hess}\,N_t)^{ab}+2N_t\left[(\mathbf{K}_t)_{ab}(\mathrm{Ric}(\mathbf{h}_t))^{ab}-6(\mathbf{K}_t)^{ab}(\mathbf{K}_t)_{ac}(\mathbf{K}_t)^c_b\right] . \tag{5.3}$$

Consider (5.3) along the world-line, $\gamma_t(i)$, of the point y_t where the minimum of N_t, in Σ_i, is attained as t varies. Our philosophy is to show that along $\gamma_t(i)$, $|\mathbf{K}_t|^2$ remains uniformly bounded above as the evolution goes on. By continuity, this will imply the uniform boundedness of $|\mathbf{K}_t|^2$ in a suitably small neighborhood of y_t, and on applying a simple argument, this will finally yield the boundedness of $|\mathbf{K}_t|^2$.

In order to estimate, along $\gamma_t(i)$, the terms $(\mathrm{Hess}\,N_t)\cdot\mathbf{K}_t$, $(\mathrm{Ric}(\mathbf{h}_t))\cdot\mathbf{K}_t$, and $(\mathbf{K}_t\times\mathbf{K}_t)\cdot\mathbf{K}_t$, appearing in the right member of (5.3), we can proceed as follows. Let A_{ik} be a shorthand for any of the bilinear forms Hess N_t, Ric($\mathbf{h}_t$), and ($\mathbf{K}_t\times\mathbf{K}_t$). Let $\zeta_{(j)}$ and $\mathbf{E}_j$ $(j=1,2,3)$ denote the eigenvalues and the normalized eigenvectors of A_{ik} at y_t, respectively. Thus

$$A^{ab}(\mathbf{K}_t)_{ab}=\Sigma_j\zeta_{(j)}(\mathbf{K}_t)_{ab}E^a_jE^b_j .$$

As is easily verified, we can write

$$(\mathbf{K}_t)_{ab}E^a_j E^b_j \leqslant (\tfrac{3}{4}(\mathbf{K}_t)_{ab}(\mathbf{K}_t)^{ab})^{1/2}. \tag{5.4}$$

[Actually this is a particular case of a useful estimate holding true for any symmetric bilinear form B_{ik} in $\mathbb{R}^n$, with $\mathrm{tr}(B)\geqslant 0$, namely,

$$a^i a^k B_{ik} \leqslant (n/n+1)^{1/2}(\mathrm{tr}(B))(\mathbf{a})^2 + (n/n+1)^{1/2}(B_{ik}B^{ik})^{1/2}(\mathbf{a})^2,$$

for any $\mathbf{a}\in\mathbb{R}^n\setminus\{0\}$, see Ref. 17.] Hence

$$A^{ab}(\mathbf{K}_t)_{ab} \leqslant \sqrt{\tfrac{3}{4}}|\mathbf{K}_t|(\zeta_{(1)}+\zeta_{(2)}+\zeta_{(3)}). \tag{5.5}$$

Since either $(\mathbf{K}_t \times \mathbf{K}_t)$ or $(\mathrm{Hess}\, N_t)(y_t)$ is a non-negative bilinear form (y_t is a minimum point for N_t), on applying (5.5) we get

$$(\mathrm{Hess}\, N_t)^{ab}(\mathbf{K}_t)_{ab}(y_t) \leqslant \sqrt{\tfrac{3}{4}}\,(|\mathbf{K}_t|\Delta N_t)(y_t) = \sqrt{\tfrac{3}{4}}\,|\mathbf{K}|^3 N_t(y_t), \tag{5.6}$$

and

$$(\mathbf{K}_t)^{ab}(\mathbf{K}_t)_{ac}(\mathbf{K}_t)^c_b(y_t) \leqslant \sqrt{\tfrac{3}{4}}\,|\mathbf{K}_t|^3(y_t), \tag{5.7}$$

where [in (5.6)] we have explicitly taken into account Eq. (3.2). The term $(\mathrm{Ric}(\mathbf{h}_t))\cdot\mathbf{K}_t$ is not so easily estimated, since, in general, $(\mathrm{Ric}(\mathbf{h}_t))(y_t)$ has both positive and negative eigenvalues. Let us denote by $\alpha_{(i)}$ $(i=1,2,3)$ such eigenvalues. By taking into account the Hamiltonian constraint we get

$$\alpha_{(1)}+\alpha_{(2)}+\alpha_{(3)} = R(y_t) = |\mathbf{K}_t|^2(y_t),$$

so that we can order the eigenvalues in such a way as to have $\alpha_{(1)}\geqslant\alpha_{(2)}\geqslant\alpha_{(3)}$, where $\alpha_{(1)}$ is non-negative while $\alpha_{(2)}$ and $\alpha_{(3)}$ can be negative. Hence, by recalling that Λ_t is the lower bound of the Ricci tensor, we get the following chain of inequalities:

$$|\alpha_{(3)}| \leqslant \Lambda_t, \quad |\alpha_{(2)}| \leqslant |\alpha_{(3)}| \leqslant \Lambda_t,$$
$$\alpha_{(1)} = \mathbf{K}_t^2 - \alpha_{(2)} - \alpha_{(3)} \leqslant \mathbf{K}_t^2 + 2\Lambda_t. \tag{5.8}$$

Since $(\mathrm{Ric}(\mathbf{h}_t))^{ab}(\mathbf{K}_t)_{ab} = \Sigma_j \alpha_{(j)}(\mathbf{K}_t)_{ab}E^a_j E^b_j$, by taking into account (5.4) we get

$$(\mathrm{Ric}(\mathbf{h}_t))^{ab}(\mathbf{K}_t)_{ab} \leqslant \sqrt{\tfrac{3}{4}}\,|\mathbf{K}_t|(\alpha_{(1)}+|\alpha_{(2)}|+|\alpha_{(3)}|) \leqslant \sqrt{\tfrac{3}{4}}\,|\mathbf{K}_t|(|\mathbf{K}_t|^2+4\Lambda_t).$$

Hence we can bound, along $\gamma_t(i)$, the right side of (5.3) as

$$\begin{aligned}(L_t(|\mathbf{K}_t|^2))_{\gamma_t(i)} &\leqslant \sqrt{3}(|\mathbf{K}_t|^3 N_t)(y_t) \\ &\quad + \sqrt{3}N_t(y_t)[\,|\mathbf{K}_t|(\mathbf{K}_t^2+4\Lambda_t)+|\mathbf{K}_t|^3\,](y_t) \\ &= \sqrt{3}N_t(y_t)[\,3\mathbf{K}_t^3+4\Lambda_t\mathbf{K}_t\,](y_t).\end{aligned}$$

By hypothesis $N_t(y_t) = \inf_i N_t$, so that, by taking into account estimate (4.23), we finally get

$$\begin{aligned}(L_t(|\mathbf{K}_t|^2))_{\gamma_t(i)} &\leqslant \sqrt{3}\,[\,3|\mathbf{K}_t|^3+4\Lambda_t\mathbf{K}_t\,](y_t) \\ &\quad \times \exp\left\{\frac{-\frac{1}{2}[(1+4\rho_t^2(i)\langle|\mathbf{K}_t|^2\rangle_{B_t})^{1/2}-1]}{1+\rho_t(i)\sigma_t(i)}\right\}.\end{aligned} \tag{5.9}$$

This estimate immediately implies that unless Λ_t blows up, during the evolution, at an exponential rate, $|\mathbf{K}_t|^2$ remains uniformly bounded above along $\gamma_t(i)$. In order to show that this result implies the uniform boundedness of $|\mathbf{K}|^2$ on the whole $(S_t, \mathbf{h}_t)$, we can use the following argument. Let N_t, $0<N_t\leqslant 1$, be the solution of problem (3.2), so that we can formally write $|\mathbf{K}_t|^2 = (\Delta N_t)/N_t$. Since by standard elliptic theory we known that ΔN_t is bounded above, and $0<N_t\leqslant 1$, we see that the only troublesome points, where we must check the uniform boundedness of $|\mathbf{K}_t|^2$, are those where N_t may tend to vanish, namely points like y_t. But estimate (5.9) shows that along any such point $|\mathbf{K}_t|^2$ remains uniformly bounded above, and the above claim immediately follows.

The above results imply that asymptotically Euclidean maximal slices with $S_t \simeq \mathbb{R}^3$, in an asymptotically flat vacuum space-time avoid scalar curvature singularities provided that the (negative) lower bound of the Ricci curvature of the slices does not grow too fast as the evolution goes on. Estimates sharper than (5.9) (which, for the way it has been obtained, is an overkill) can probably provide more precise results concerning the critical rate of growth for Λ_t. In any case, it appears quite unlikely that one could obtain results of the above kind without imposing any restriction on the Ricci curvature of the slices.

ACKNOWLEDGMENTS

The author gratefully acknowledges the friendly environment of the Department of Nuclear and Theoretical Physics of the University of Pavia where this work was carried out.

This work has been partially supported by the Gruppo Nazionale di Fisica Matematica (GNFM) of the Consiglio Nazionale delle Richerche (CNR).

[1] R. Bartnik, Ph. D. thesis, Princeton, 1983; see also R. Bartnik, in *Asymptotic Behavior of Mass and Spacetime Geometry, Springer Lecture Notes in Physics* (Springer, New York, 1984); Commun. Math. Phys. **94**, 155 (1984).
[2] Y. Choquet-Bruhat, A. E. Fischer, and J. E. Marsden, in *Isolated Gravitating System in General Relativity*, edited by J. Ehlers (North-Holland, Amsterdam, 1979), p. 396. See also, Y. Choquet-Bruhat and J. W. York, Jr., in *General Relativity and Gravitation*, edited by A. Held (Plenum, New York, 1980), Vol. I, p. 158.
[3] J. E. Marsden and F. J. Tipler, Phys. Rep. **66**, 109 (1980).
[4] F. Estabrook, H. Wahlquist, S. Christiansen, B. Dewitt, L. Smarr, and E. Tsiang, Phys. Rev. D **7**, 2814 (1973).
[5] L. Smarr, Ann. N. Y. Acad. Sci. **302**, 569 (1977).
[6] K. Eppley, Phys. Rev. D **16**, 1609 (1977).
[7] K. Eppley, in *Sources of Gravitational Radiation*, edited by L. Smarr (Cambridge U. P., Cambridge, 1979), p. 275.
[8] L. Smarr and J. W. York, Jr., Phys. Rev. D **17**, 2529 (1979).
[9] S. Y. Cheng and S. T. Yau, Commun. Pure Appl. Math. **28**, 333 (1975); Ann. Math. **104**, 407 (1976).
[10] D. M. Eardley and L. Smarr, Phys. Rev. D **19**, 2239 (1979).
[11] J. W. York, Jr., in Ref. 7, p. 83.
[12] M. Cantor, Indiana Univ. Math. J. **24**, 897 (1975).
[13] D. Fischer-Colbrie and R. Schoen, Commun. Pure Appl. Math. **33**, 199 (1980).
[14] M. Cantor and D. Brill, Compos. Math. **43**, 317 (1981).
[15] F. Tipler, Phys. Rev. D **17**, 2521 (1978).
[16] Y. Choquet-Bruhat, D. Christodoulou, and M. Francaviglia, Ann. Inst. H. Poincaré, **31** 399 (1979).
[17] S. T. Yau, Commun. Pure Appl. Math. **28**, 201 (1975).
[18] S. T. Yau, *Seminar on Differential Geometry, Annals of Mathematical Studies*, No. 102 (Princeton U. P, Princeton, NJ, 1982).
[19] M. Berger, P. Gauduchon, and E. Mazet, *Le Spectre d'une Variété Riemannienne, Springer Lecture Notes*, Vol. 194 (Springer, Berlin, 1971).
[20] M. Protter, and H. Weinberger, *Maximum Principles in Differential Equations* (Prentice–Hall, Englewood Cliffs, NJ, 1967).
[21] R. L. Bishop and R. J. Crittenden, *Geometry of Manifolds* (Academic, New York, 1964).
[22] L. Smarr, in *General Relativity and Gravitation*, edited by B. Bertotti, F. De Felice, and A. Pascolini (Reidel, Dordrecht, 1984), p. 163.

Is perturbation theory misleading in general relativity?

Robert Geroch and Lee Lindblom
Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637

(Received 10 April 1985; accepted for publication 7 June 1985)

Two senses in which the perturbation equations of general relativity can be misleading are explored. (i) Under certain circumstances there exist solutions of the perturbation equations that appear to be gauge, in that the metric perturbation is the symmetrized derivative of a vector field, but which nonetheless are not true gauge. (ii) Under certain circumstances there exist solutions of the perturbation equations that cannot, even locally, be extended to higher order in perturbation theory. The latter is a local version of the well-known phenomenon of "linearization instability."

I. INTRODUCTION

Einstein's equation, a system of nonlinear second-order partial differential equations, is sufficiently complex that, for most situations of physical interest, there are available no corresponding exact solutions. Consequently, much of our insight into the physical implications of general relativity has come from the study of approximate solutions. The most common approximation method is perturbation theory: One introduces a background space-time—an exact solution of Einstein's equation—and then considers deviations, to first or higher orders, from this background.

To what extent do such approximate solutions correspond to exact solutions? The question has both a quantitative and a qualitative aspect. The quantitative question asks for some numerical measure of the extent to which an approximate solution corresponds to some exact solution. Consider, for example, use of the quadrupole formula to compute the amount of gravitational radiation emitted by a system. By how much does the result of this computation differ from the correct answer—that obtained from the full Einstein equation?[1] The qualitative question, on the other hand, asks whether the predictions of the approximate solutions agree, even in their broad, overall features, with those of exact solutions.

It was first noted by Brill[2] that, under certain circumstances, there is not even qualitative agreement between the linearized and full solutions of Einstein's vacuum equation. These circumstances are that the background space-time possess both a compact Cauchy surface and a Killing field.[3–6] One first writes down a certain integral over the Cauchy surface, where the integrand involves the Killing field and an arbitrary solution of the first-order perturbation equation. One then shows (i) that, by virtue of the second-order perturbation equation, this integral must vanish, and (ii) that there exist solutions of the first-order perturbation equation for which the integral is nonzero. Then these first-order perturbations, since they cannot even be extended to second order, certainly cannot come from any family of solutions of the full Einstein equation.

We consider here a somewhat different class of circumstances under which perturbation theory is qualitatively incorrect. As in the result above, we require of the background space-time that it possess symmetries. But, in contrast to that result, (i) the perturbation fields are required to respect the symmetries, rather than being allowed to break those symmetries, and (ii) the arguments are purely local, rather than global.

An example will illustrate what we have in mind. Consider the external gravitational field of a static, plane-symmetric "sheet" of matter. We expect that the space-time appropriate to this situation will have three orthogonal, commuting Killing fields: t^a (timelike, giving the static character), and x^a and y^a (spacelike, giving the plane symmetry).

Consider first the limit in which the stress energy of the sheet is small. It should then be appropriate to treat the gravitational field as a linear perturbation off Minkowski space-time. Denote by g_{ab} the Minkowski metric, by ∇_a its associated derivative operator, and by t^a, x^a, y^a three orthonormal translations in this space-time. The first-order perturbation of the metric, denoted h_{ab}, must satisfy the linearized Einstein equation

$$\nabla^m\nabla_m h_{ab} - 2\nabla^m\nabla_{(a}h_{b)m} + \nabla_a\nabla_b h^m{}_m = 0, \tag{1}$$

and must respect the symmetries

$$\mathcal{L}_t h_{ab} = \mathcal{L}_x h_{ab} = \mathcal{L}_y h_{ab} = 0. \tag{2}$$

We now claim the following: *the most general solution of Eqs.* (1) *and* (2) *is given by*

$$h_{ab} = 2\nabla_{(a}\tau_{b)}, \tag{3}$$

where τ^a is any vector field such that each of $\mathcal{L}_t\tau^a$, $\mathcal{L}_x\tau^a$, and $\mathcal{L}_y\tau^a$ is a Killing field. Clearly, any such h_{ab} does indeed satisfy (1) and (2). To prove the converse, let h_{ab} satisfy (1) and (2). Equation (2) is the statement that $z_{[a}\nabla_{b]}h_{cd} = 0$, where z^a is the unit translation in the background orthogonal to the other three. But this in turn implies

$$z_{[a}R_{bc]de} = 0, \tag{4}$$

where

$$R_{abcd} = -2\nabla_{[a}\nabla_{|[c}h_{d]|b]} \tag{5}$$

is the linearized Riemann tensor. Equation (1) is the statement that all traces of R_{abcd} vanish. But this implies, taking the double dual of (4), that $z^aR_{abcd} = 0$. Contracting (4) with z^a and using this last equation, we obtain that $R_{abcd} = 0$, i.e., that h_{ab} is of the form (3) for *some* τ^a. Finally, for h_{ab} of this form, Eq. (2) is precisely that statement that each of $\mathcal{L}_t\tau^a$, $\mathcal{L}_x\tau^a$, and $\mathcal{L}_y\tau^a$ is a Killing field, completing the proof.

We thus conclude that every linear perturbation appro-

 0022-2488/85/102581-08$02.50

priate to this problem is "pure gauge." This conclusion is not surprising, for one might expect a static, plane-symmetric sheet of matter to result in a "uniform gravitational field," i.e., in flat space-time.

For this particular problem, however, we can compare the linearized approximation with the full theory, for all vacuum solutions of the full Einstein equation having three commuting Killing fields are known.[7,8] The most general such solution can be represented as

$$ds^2 = -(\lambda z+1)^{2p_1}dt^2+(\lambda z+1)^{2p_2}dx^2 + (\lambda z+1)^{2p_3}dy^2+dz^2, \tag{6}$$

where p_1, p_2, p_3 are any three numbers satisfying $p_1+p_2+p_3=p_1{}^2+p_2{}^2+p_3{}^2=1$, and λ is any number. [The parameter λ, introduced here for later convenience, has the effect of merely rescaling the coordinates. So, the family of essentially different solutions forms a circle (intersection of a plane and a sphere in the space of the p_i).] For λ nonzero, these solutions (except those having one of the p_i one, the others zero) are nonflat; while, for λ zero, all the solutions are flat. Thus, there are exact, nonflat solutions to the problem arbitrarily close to Minkowski space-time. Yet all linear perturbations off Minkowski space-time are pure gauge! It would seem that first-order perturbation theory provides too few solutions to reflect adequately the full theory.

There is of course a direct way to see what is going on in this particular example: Take the family of metrics in (6), and linearize it in λ about $\lambda=0$. There results a linearized metric h_{ab} of the form (3), but with a particular choice of τ^a, namely

$$\tau^a = -p_1 z t t^a + p_2 z x x^a + p_3 z y y^a + \tfrac{1}{2}(p_1 t^2 - p_2 x^2 - p_3 y^2) z^a. \tag{7}$$

This h_{ab} does indeed satisfy Eqs. (1) and (2), for each of

$$\begin{aligned} \mathcal{L}_t \tau^a &= -p_1(z t^a - t z^a), \\ \mathcal{L}_x \tau^a &= p_2(z x^a - x z^a), \\ \mathcal{L}_y \tau^a &= p_3(z y^a - y z^a) \end{aligned} \tag{8}$$

is a Killing field, as required. Thus, the situation is the following. First, it is indeed true that all linearized solutions in this plane-symmetric problem are symmetrized derivatives of vectors, as in Eq. (3). But this is the "right answer," for the linearization about Minkowski space-time of our family of exact solutions also yields a metric perturbation the symmerized derivative of a vector field. It is only at higher order in perturbation theory that the nonflat character of these exact solutions becomes apparent. What is happening, then, is that a first-order perturbation of the form $h_{ab}=2\nabla_{(a}\tau_{b)}$ is necessary in order to provide access, at higher order in perturbation theory, to the nonflat exact solutions. Such first-order perturbations, then, should not in this example be regarded as "gauge."

This same example illustrates a second potential difficulty with perturbation theory. We saw above that the first-order theory appears to provide too few solutions to reflect the full theory, an appearance that resulted from a too-broad application of the term gauge. The second difficulty is that the first-order theory *does*—not only in appearance, but in fact—provide too *many* solutions to reflect the full theory. To see this, consider the linearized solution h_{ab} given by Eq. (3) with τ^a given by Eq. (7)—but with the constants p_i now chosen so that $(p_1+p_2+p_3)^2$ is not equal to $p_1{}^2+p_2{}^2+p_3{}^3$. This h_{ab} does not indeed satisfy the first-order linear equations (1) and (2). But, we claim, this particular linearized solution comes from no family of exact solutions. This claim (which will be discussed in more detail later) can be seen directly at this point by taking all linearizations about flat space-time of families of solutions taken from Eq. (6) (possibly allowing the p_i to depend on the parameter λ; possibly applying λ-dependent, Killing-field-preserving diffeomorphisms). This second difficulty is analogous to the "linearization instability" discussed earlier by Brill[2] and others.[3–6] But there are several significant differences. The earlier work required a compact Cauchy surface (while here there is none), dealt with arbitrary first-order perturbations (while here they must respect the symmetries), and involved a global argument (while here local).

This example illustrates the issues with which we shall be concerned in this paper. Under what circumstances can "apparent gauge" arise? What is the "correct" notion of gauge? Under what circumstances does perturbation theory yield solutions extraneous to the full theory? How can these extraneous solutions be identified?

II. PERTURBATION THEORY

In this section, we set up the general framework, consisting of a few definitions and their basic properties, for perturbation theory.

Fix a manifold $\mathcal{M}$. Specify a list of the types of fields to be considered on $\mathcal{M}$, as well as a list of the equations to be satisfied by those fields. In our earlier example, $\mathcal{M}$ was the manifold R^4; the fields were a symmetric, Lorentz-signature metric g_{ab} and three vectors fields, t^a, x^a, and y^a; the equations were Einstein's equation, Killing's equation for each of t^a, x^a, and y^a, and the equations asserting that all Lie brackets of these vector fields vanish.

Next, let the fields under consideration be divided into two classes: the *passive* fields and the *dynamic* fields. The passive fields will be fixed throughout, while the dynamic fields will be subject to perturbation. In our example, the three vector fields t^a, x^a, and y^a were passive, while the metric g_{ab} was dynamic.

Finally, fix a solution of the system. That is, fix actual fields, satisfying all the given equations, on $\mathcal{M}$. This solution will be called the *background*. In our example, the background consisted of a flat metric, together with three unit, orthogonal translations, on $\mathcal{M}$.

This, then, is the arrangement we contemplate for perturbation theory: a manifold $\mathcal{M}$, a list of fields and the equations they are to satisfy, a division of the fields into passive and dynamic, and a background solution of the system.

Now consider a one-parameter family of solutions of this system of equations, jointly smooth on $\mathcal{M}$ and in the parameter λ. Let this family be such that all passive fields are independent of λ, and such that the fields in the family reduce, for $\lambda=0$, to the background fields. Taking the first derivatives, with respect to λ, of the dynamic fields in this

family, and evaluating at $\lambda = 0$, we obtain the first-order perturbed fields. Taking the first derivatives, with respect to λ, of the equations applied to our family of fields, and evaluating at $\lambda = 0$, we obtain the *first-order perturbation equations*. This is a system of linear equations on the first-order perturbed fields. The first-order perturbation equations, once derived in this way, are then regarded as equations in their own right; one is free to consider solutions of this system of equations without reference to a one-parameter family—or even to whether such a family exists—giving rise to that solution. In our earlier example, the only first-order perturbed field was the perturbed metric h_{ab}, and the first-order perturbation equations were precisely Eqs. (1) and (2). (Nothing new results from the λ derivatives of the commutation relations, since these involve only the passive fields.) More generally, taking all derivatives up to the nth, with respect to λ, of the equations, and evaluating to $\lambda = 0$, we obtain the nth-*order perturbation equations*. These equations involve the background fields and the first n derivatives, with respect to λ, of the dynamic fields. The nth derivatives of the dynamic fields always appear linearly, but the other derivatives in general do not. Any solution of the nth-order perturbation equations yields immediately a solution of the mth-order equations (for $m < n$), by simply omitting all λ derivatives of order higher than the mth.

There is available a particularly simple class of solutions of the nth-order perturbation equations. Consider a one-parameter family $\mathscr{D}(\lambda)$ of diffeomorphisms on $\mathscr{M}$, jointly smooth on $\mathscr{M}$ and in λ. Let this family be such that all passive background fields are invariant under all $\mathscr{D}(\lambda)$, and such that $\mathscr{D}(0)$ is the identity diffeomorphism on $\mathscr{M}$. Applying these $\mathscr{D}(\lambda)$ to each of the dynamic background fields, we obtain a one-parameter family of solutions of the system. This family satisfies, by virtue of the conditions just imposed on $\mathscr{D}(\lambda)$, the conditions of the previous paragraph. Hence, taking all derivatives up to the nth with respect to λ, of the dynamic fields in this family, and evaluating at $\lambda = 0$, we obtain a solution of the nth-order perturbation equations. Solutions of the form so obtained will be called gauge solutions, reflecting the fact that the result of applying $\mathscr{D}(\lambda)$ to the background has exactly the same physical content as the background itself. Thus, the general gauge solution of the first-order perturbation equations is determined by a vector field τ^a on $\mathscr{M}$ [reflecting the first derivative of $\mathscr{D}(\lambda)$ at $\lambda = 0$], with respect to which the Lie derivatives of the passive background fields vanish [reflecting the condition that the passive background fields be invariant under the $\mathscr{D}(\lambda)$]. The first-order perturbed fields are those obtained by applying $\mathscr{L}_\tau$ to each of the dynamic background fields. The general gauge solution of the second-order perturbation equations is determined by two vector fields τ^a and σ^a on $\mathscr{M}$ [reflecting the first two derivatives $\mathscr{D}(\lambda)$ at $\lambda = 0$], with respect to which both of the Lie derivatives of the passive background fields vanish. The perturbed fields are those obtained by applying each of $\mathscr{L}_\tau$ and $\mathscr{L}_\tau\mathscr{L}_\tau + \mathscr{L}_\sigma$ to each of the dynamic background fields.

Fix a manifold $\mathscr{M}$, lists of the types of passive and dynamic fields to be considered on $\mathscr{M}$, a list of the equations to be satisfied by these fields, and a background solution of the system. We have seen above that certain one-parameter families of solutions of this system lead to solutions of the perturbation equations, and that certain one-parameter families of diffeomorphisms on $\mathscr{M}$ lead to gauge solutions of the perturbation equations. We now consider the extent to which these processes can be reversed. Does a given solution of the perturbation equations arise from some one-parameter family of solutions of the original system? Does a given gauge solution arise from some one-parameter family of diffeomorphisms?

The second question is easy to answer. We claim the following: Given any gauge solution of the nth-order perturbation equations, there exists a one-parameter family $\mathscr{D}(\lambda)$ of diffeomorphisms giving rise to that solution. To prove this, we must show that, given n vector fields τ^a, σ^a,...,κ^a on $\mathscr{M}$, there exists a family $\mathscr{D}(\lambda)$ of diffeomorphisms whose first n derivatives with respect to λ, at $\lambda = 0$, are characterized by these vector fields. But for $n = 1$, a suitable family[9] is given by $\mathscr{D}_\tau(\lambda)$, the diffeomorphisms generated by the vector field τ^a itself; for $n = 2$, by $\mathscr{D}_\tau(\lambda) \circ \mathscr{D}_\sigma(\lambda^2/2)$; and similarly for other n.

The question of whether a given solution of the perturbation equations arises from some one-parameter family of exact solutions of the original system is more difficult. The source of the difficulty is that one does not in general have easy access to the exact solutions of the system. Indeed, it is the lack of such access that causes one to turn to perturbation theory in the first place. Fortunately, there is a notion closely related to "come from a family of exact solutions," but far easier to work with. We say that the nth-order perturbation equations are *reliable* if every solution of those equations can be extended [by some choice of the $(n+1)$st-order perturbed fields] to a solution of the $(n+1)$st-order perturbation equations. The condition, then, is that the solutions can be extended to one higher order in perturbation theory. The advantage of this definition is that it deals only with the perturbation equations and their solutions, with no reference to exact solutions of the full system. Should it happen that every solution of the nth-order perturbation equations comes from some family of exact solutions, then the perturbation equations must certainly be reliable; extend any perturbed solution from nth to $(n+1)$st order using the family of exact solutions. But there is no guarantee that, conversely, reliability implies that all perturbed solutions must come from families of exact solutions.

III. GAUGE

We now return to the example of Sec. I: the external gravitational field of a static, plane-symmetric sheet of matter. We found in that example that the general solution of the first-order perturbation equations (1) and (2) is

$$h_{ab} = 2\nabla_{(a}\tau_{b)}\,, \tag{9}$$

where τ^a is any vector field such that each of $\mathscr{L}_t\tau^a$, $\mathscr{L}_x\tau^a$, and $\mathscr{L}_y\tau^a$ is Killing field. We initially interpreted these solutions, in light of Eq. (9), as gauge; they give, for example, vanishing curvature tensor to first order. But this interpretation was found to be unacceptable, for there are in this example exact, nonflat solutions of the system arbitrarily close to

the Minkowski background. A one-parameter family of such exact solutions does give rise to a first-order perturbed metric of the form (9), with τ^a given by Eq. (7).

In Sec. II, we introduced a general framework for perturbation theory, applicable to virtually any system of equations on fields. Within that framework, gauge solutions of the perturbation equations were defined quite generally as arising from certain families of diffeomorphisms preserving the passive fields. Applying the general definition to this example, we obtain the following: The gauge solutions of the first-order perturbation equations (1) and (2) are those of the form (9), but with τ^a now a vector field such that each of $\mathcal{L}_t\tau^a$, $\mathcal{L}_x\tau^a$, and $\mathcal{L}_y\tau^a$ vanishes. Thus, we correctly exclude from being gauge the solution with τ^a given by Eq. (7)—the solution which gives access to the nonflat exact solutions. While this solution arises from *some* diffeomorphisms, it does not arise from those preserving the passive fields. In short, gauge must be defined in this more restrictive way in the presence of passive fields.

This phenomenon—the existence of nongauge perturbed metrics that are nonetheless symmetrized derivatives of vector fields—is not just a special feature of the static, plane symmetric case. It is rather, as the following example shows, pervasive for Einstein's equation in the presence of symmetries. Fix a manifold $\mathcal{M}$. Let the dymanic field be a Lorentz-signature metric g_{ab}, and the passive fields n vector fields $\underset{i}{\xi}{}^a$. Let the equations be Einstein's equation, Killing's equation for each of the $\underset{i}{\xi}{}^a$, and a set of commutation relations

$$\mathcal{L}_{\underset{i}{\xi}}\,\underset{j}{\xi}{}^a = C^k{}_{ij}\underset{k}{\xi}{}^a, \tag{10}$$

where the $C^i{}_{jk}$ are fixed constants satisfying the Jacobi relation $C^m{}_{[ij}C^l{}_{k]m} = 0$. Fix a background solution g_{ab}, $\underset{i}{\xi}{}^a$. To simplify the discussion, we suppose that the background g_{ab} admits no Killing fields other than linear combinations, with constants coefficients, of the $\underset{i}{\xi}{}^a$. The first-order perturbation equations in this case are those analogous to Eqs. (1) and (2). One class of solutions of these equations is that with h_{ab} given by Eq. (9), where τ^a is any vector field such that each of the $\mathcal{L}_{\underset{i}{\xi}}\tau^a$ is a Killing field

$$\mathcal{L}_{\underset{i}{\xi}}\tau^a = U_i{}^k\underset{k}{\xi}{}^a, \tag{11}$$

where $U_i{}^k$ are constants. Applying $\mathcal{L}_{\underset{j}{\xi}}$ to Eq. (11) and antisymmetrizing over i and j, we find that $U_i{}^k$ must satisfy the further condition

$$U_i{}^m C^k{}_{mj} + U_j{}^m C^k{}_{im} - U_m{}^k C^m{}_{ij} = 0. \tag{12}$$

These solutions are "apparent gauge." The true gauge solutions, on the other hand, are those of the form above, but with the additional property that the $U_i{}^k$ in Eq. (11) vanish.

Under what condition are all apparent gauge solutions in fact true gauge? Note that we may, without altering the perturbed metric, add to the τ^a in Eq. (9) any Killing field, i.e., any field of the form $w^i\underset{i}{\xi}{}^a$ with the w^i constant. The condition, then, is that by adding such a field to τ^a we may achieve vanishing of the $U_i{}^k$ in Eq. (11), i.e., that

$$U_i{}^j = w^m C^j{}_{mi}, \tag{13}$$

for some w^i. We conclude the following: *Every solution that is apparent gauge is actually true gauge provided every* $U_i{}^k$ *satisfying Eq.* (12) *is of the form* (13). Note that this condition involves only the Lie-algebra structure on the Killing fields. It may, in fact, be restated thus: Every infinitesimal automorphism on the Lie algebra is inner. Note also that, by the Jacobi relation, every $U_i{}^k$ of the form (13) automatically satisfies Eq. (12).

It is generally easy to decide whether a given Lie algebra satisfies the condition above, i.e., whether apparent gauge, in the presence of passive symmetries with that Lie algebra, must be true gauge. For the zero-dimensional Lie algebra (no passive fields), the condition is of course satisfied. For all one- or two-dimensional Lie algebras, the condition is not. For three-dimensional Lie algebras, it depends on the algebra. It is satisfied, for example, for the Lie algebra of SO(3) (the rotations), and for SO(2,1), but not for the three-dimensional commutative Lie algebra. We remark that, when the condition above fails, then there normally do exist solutions that are apparent gauge but not true gauge.

The class of solutions considered here is restricted in that we allow no sources in Einstein's equation and no passive fields other than Killing fields, and yet broad enough to include curved background metrics and arbitrary Lie algebras of symmetries. It appears that apparent gauge is pervasive for Einstein's equation in the presence of passive symmetries.

IV. RELIABILITY

We now return again to the static, plane-symmetric example of Sec. I. We found in that example that there exist solutions of the first-order perturbation equations—namely, those of the form (3), with τ^a given by Eq. (7) with the constant p_i so chosen that $(p_1+p_2+p_3)^2 \neq {p_1}^2+{p_2}^2+{p_3}^2$—that do not come from any one-parameter families of exact solutions of the system. We interpreted this phenomenon as indicating that the first-order perturbation equations do not adequately reflect the full equations. In the general framework for perturbation theory in Sec. II, we introduced, for virtually any system of equations on fields, a closely related notion: The nth-order perturbation equations are reliable if every solution of those equations can be extended to a solution of the equations at next-highest order in perturbation theory. We now ask how this specific example fits into the general framework.

The first-order perturbation equations for this example are (1) and (2). The second-order perturbation equations are

$$\nabla^m\nabla_m i_{ab} - 2\nabla^m\nabla_{(a}i_{b)m} + \nabla_a\nabla_b i^m{}_m = t_{ab} - \tfrac{1}{2}g_{ab}t^m{}_m \tag{14}$$

and

$$\mathcal{L}_t i_{ab} = \mathcal{L}_x i_{ab} = \mathcal{L}_y i_{ab} = 0, \tag{15}$$

where i_{ab} denotes the second-order metric perturbation, and t_{ab} is given by

$$t_{ab} - \tfrac{1}{2} g_{ab} t^{m}{}_{m} = -4h^{mn}\nabla_{[a}\nabla_{|[b}h_{m]|n]} - \tfrac{1}{2}(\nabla_a h^{mn})(\nabla_b h_{mn}) - 2(\nabla^m h^{n}{}_{a})(\nabla_{[m}h_{n]b}) + (\nabla_m h^{mn} - \tfrac{1}{2}\nabla^n h^{m}{}_{m})(2\nabla_{(a}h_{b)n} - \nabla_n h_{ab}). \tag{16}$$

We now claim the following: *Given h_{ab} satisfying Eqs.* (1) *and* (2), *there exists i_{ab} satisfying Eqs.* (14) *and* (15) *if and only if*

$$u^{mn}u^{pq}s_{mp}s_{nq} - (u^{mn}s_{mn})^2 = 0, \tag{17}$$

where z^a is the unit translation orthogonal to the other three, $u^{ab} = g^{ab} - z^a z^b$ is the projection orthogonal to z^a, and s_{ab} is given by

$$\nabla_a h_{bc} = z_a s_{bc} \tag{18}$$

[noting that the existence of such an s_{ab} is guaranteed by (2)]. To prove this claim, let h_{ab} satisfy (1) and (2). Set, by Eq. (15), $\nabla_a\nabla_b i_{cd} = z_a z_b \kappa_{cd}$, for some κ_{cd}, and substitute into Eq. (14) to obtain the following: There exists i_{ab} satisfying Eqs. (14) and (15) if and only if $z^m t_{am} = 0$. Next, solve Eq. (16) [noting that, by (1) and (2), the first term on the right vanishes] for t_{ab}, contract with z^b, and substitute Eq. (18). The claim follows.

In particular, for the h_{ab} by given by Eq. (3) with τ^a given by (7), the condition, Eq. (17), for the existence of a second-order perturbed metric becomes $(p_1 + p_2 + p_3)^2 = p_1^2 + p_2^2 + p_3^2$. This is precisely the same as the condition, obtained in Sec. I, that our h_{ab} come from some one-parameter family of exact solutions. We conclude, then, that the first-order perturbation equations in this example are not reliable. That perturbation theory gives the "wrong answer" in this example is correctly detected by the notion of reliability. This is a local version of the linearization instability noted earlier.[2,6]

When, in general relativity, can local perturbation theory be trusted, and when can it not? To answer this question fully appears to be difficult. But the following discussion does suggest that reliability in general relativity is more the rule than the exception.

Let the dynamic field be a Lorentz-signature metric g_{ab}, the passive fields m vector fields $\underset{i}{\xi}{}^a$. Let the equations be Einstein's equation with vanishing sources, Killing's equation for each of the $\underset{i}{\xi}{}^a$, and set of commutation relations on the $\underset{i}{\xi}{}^a$. Fix a background solution. Then the nth-order perturbation equations are

$$\nabla^m\nabla_m n_{ab} - 2\nabla^m\nabla_{(a}n_{b)m} + \nabla_a\nabla_b n^{m}{}_{m} = t_{ab} - \tfrac{1}{2} g_{ab} t^{m}{}_{m} \tag{19}$$

and

$$\mathcal{L}_{\underset{i}{\xi}} n_{ab} = 0, \tag{20}$$

where n_{ab} is the nth-order perturbed metric and $t_{ab} = t_{(ab)}$ is some expression involving the perturbed metrics up to order $(n-1)$. It follows from the $(n-1)$st-order perturbation equations that this t_{ab} respects the symmetries $\mathcal{L}_{\underset{i}{\xi}} t_{ab} = 0$ and is conserved, $\nabla_m t^{am} = 0$. Hence, *perturbation theory is reliable at every order provided that, for any t_{ab} that respects the symmetries and is conserved there exists an n_{ab} satisfying Eqs.* (19) *and* (20).

We conjecture that one can always solve (19) and (20) locally, and so perturbation theory is locally reliable, whenever the number of passive Killing fields is two or less. Evidence for this conjecture comes from the important special case in which the Killing fields are linearly independent at each point and all Killing fields are spacelike. Choose, in this case, a spacelike slice S to which all Killing fields are tangent. The idea is to solve Eq. (19) for n_{ab} using an initial-value formulation on S. While this equation is not hyperbolic as it stands, it becomes such if there is imposed on n_{ab} the Lorentz-gauge condition

$$\nabla^m(n_{am} - \tfrac{1}{2} g_{am} n^{s}{}_{s}) = 0. \tag{21}$$

The initial data then consist of n_{ab}, together with its first normal derivative, evaluated on S. These must be so chosen that the gauge condition (21), together with its first normal derivative, are satisfied on S. To this end, choose as the data $\alpha t_a t_b + 2\alpha_{(a}t_{b)}$ for n_{ab} and $\beta t_a t_b + 2\beta_{(a}t_{b)}$ for its first normal derivative, where t^a is the unit normal to S and α^a and β^a are both orthogonal to t^a. Substituting, the gauge condition (21) gives expressions for β and β^a in terms of α and α^a, while its first normal derivative becomes

$$D^2\alpha = \mu, \quad D^2\alpha^a = \mu^a, \tag{22}$$

where D^2 is the Laplacian operator on S, and the source terms on the right involve α and α^a only through their values and first derivatives in S. We thus conclude the following: There exists a solution n_{ab} of Eqs. (19) and (20) provided there exists a solution α, α^a of the elliptic system (22) with α and α^a invariant under the Killing fields. With two or fewer Killing fields, linearly independent at each point, there does exist a solution of Eq. (22) with α, α^a invariant under the Killing fields, as one sees by passing to the manifold of trajectories[10] of the symmetries. It is false in general that there exists such a solution with three or more Killing fields. (The situation here is analogous to trying to solve the Newtonian gravitational equation $D^2\varphi = \rho$ such that the symmetries of ρ are also carried by φ. For ρ invariant under three translations, i.e., constant, there is no solution φ with the same symmetries.) It seems likely that one could prove the full conjecture by similar arguments.

So, the local unreliability apparently sets in only for space-times with a high degree of symmetry—three or more Killing fields. The situation may be contrasted with the global linearization instability, which sets in already with a single Killing field.

How pervasive is unreliability for space-times with high symmetry? A simple class of examples is that provided by certain spatially homogeneous space-times: those for which a symmetry group acts simply transitively on spacelike slices. Recall[11] that such a space-time is determined by a three-dimensional Lie algebra together with positive-definite metric q_{ab} and symmetric tensor p^{ab} over the vector space of the Lie algera, satisfying the constraint equations

$$(p^{m}{}_{m})^2 - p^{mn}p_{mn} - \tfrac{3}{2}v^m v_m - s^{mn}s_{mn} + \tfrac{1}{2}(s^{m}{}_{m})^2 = 0, \tag{23}$$

$$p^{m}{}_{n}s^{np}\epsilon_{pma} - \tfrac{1}{2}p^{m}{}_{m}v_a + \tfrac{3}{2}p_a{}^{m}v_m = 0. \tag{24}$$

Here, ϵ_{abc} is the alternating tensor for q_{ab}, and $s^{ab} = s^{(ab)}$ and

v_a result[12] from the decomposition of the structure-constant tensor of the Lie algebra

$$C^a{}_{bc} = s^{am}\epsilon_{mbc} - \delta^a{}_{[b}v_{c]}. \tag{25}$$

The Jacobi relation in the Lie algebra is just $s^{am}v_m = 0$.

Fix the Lie algebra, and consider the 12-dimensional manifold $\mathscr{S}$ of all pairs consisting of positive-definite q_{ab} and symmetric p^{ab}. Denote by $\mathscr{C}$ the subset of this manifold consisting of those points for which q_{ab}, p^{ab} satisfy the constraint equations (23) and (24). Thus, a point of $\mathscr{C}$ determines such a spatially homogeneous space-time. Fix a point q_{ab}, p^{ab} of $\mathscr{C}$, and suppose that $\mathscr{C}$ is a submanifold of $\mathscr{S}$ in a neighborhood of this point. Then a solution of the first-order perturbation equations with background determined by this point yields a tangent vector to $\mathscr{C}$ at this point. Since $\mathscr{C}$ is there a submanifold, this tangent vector is tangent to some curve in $\mathscr{C}$ through the point. Thus, every solution of the first-order perturbation equations with this background arises from some one-parameter family of exact solutions of the system. We conclude that the first-order perturbation equations (and, similarly, the higher-order equations) are reliable whenever the background is given by a point of $\mathscr{C}$ in a neighborhood of which $\mathscr{C}$ is a submanifold of $\mathscr{S}$. The issue of at what points $\mathscr{C}$ is a submanifold is analyzed in the Appendix, with the following result: *The perturbation equations are reliable to all orders for all such spatially homogeneous space-times, with the possible exception of those with* $v_a = 0$ *and* s^{ab} *and* p_{ab} *of the form*

$$s^{ab} = \alpha q^{ab} + \beta w^a w^b, \tag{26}$$

$$p_{ab} = \gamma q_{ab} + \delta w_a w_b, \tag{27}$$

where w^a is any q-unit vector in the Lie algebra, and α, β, γ, and δ are constants satisfying

$$4\gamma(3\gamma + 2\delta) + (\alpha + \beta)(3\alpha - \beta) = 0. \tag{28}$$

Thus, the perturbation equations are reliable to every order for the vast majority of spatially homogeneous background space-times, including in particular all those of Bianchi types III, IV, V, VI_h, and VII_h. The static, plane-symmetric example discussed earlier is analogous to the case $\alpha = \beta = \gamma = \delta = 0$ above. It is curious that all the "exceptional" backgrounds have a fourth Killing field.

Thus, the perturbation equations of general relativity have a pronounced tendency to be reliable.

V. DISCUSSION

We have been concerned here with two types of difficulties associated with perturbation theory in general relativity. The first is that there can exist solutions of the perturbation equations that appear to be gauge—in the sense that the perturbed metric is the symmetrized derivative of a vector field—but which nonetheless are not true gauge. The second is that there can exist solutions of the perturbation equations that cannot be extended to solutions at the next-higher order in perturbation theory. The first difficulty is more pervasive in general relativity than the second, but it also much easier to deal with.

One might think that, at least for situations similar to that of the static, plane-symmetric example, these difficulties are merely the result of a poor choice of variable. Recall that first-order perturbation theory for the vacuum Einstein equation with flat background is often formulated as follows. Let the field be, not the perturbed metric h_{ab}, but rather the perturbed Weyl tensor K_{abcd}, a field having all the symmetries and traces of a Weyl tensor. Let the first-order perturbation equation be, not Eq. (1), but rather the first-order Bianchi identity

$$\nabla_{[a}K_{bc]de} = 0. \tag{29}$$

There is no gauge freedom within K_{abcd} itself, only in its "potential," h_{ab}. This formulation, which does not fit within the general framework of Sec. II, has no second-order version, and is generally inapplicable with sources or curved backgrounds.

It is instructive to see what happens when this formulation is applied to the static, plane-symmetric example. Take as the first-order perturbation equations in this case Eq. (29) together with

$$\mathscr{L}_t K_{abcd} = \mathscr{L}_x K_{abcd} = \mathscr{L}_y K_{abcd} = 0. \tag{30}$$

The most general solution to this system is K_{abcd}, any constant field having the symmetries and traces of a Weyl tensor. By contrast, every one-parameter family of exact solutions of the system gives rise to a vanishing first-order perturbed Weyl tensor. Thus, on the one hand, the first-order perturbation equations in this formulation admit too many solutions; all constant K_{abcd} rather than just $K_{abcd} = 0$. But, on the other hand, these perturbation equations admit too few solutions to reflect adequately the full theory, for all families of exact solutions—even those that do not represent first-order gauge—collapse to $K_{abcd} = 0$ in the perturbation theory. Since a family of exact solutions of the system gives rise to a nonzero Weyl tensor only at second order, one might expect better agreement between the resulting second-order Weyl-tensor perturbations and the solutions K_{abcd} of Eqs. (29) and (30). While all such second-order perturbations of the Weyl tensor are in fact constant tensor fields, not all constants are allowed; there must be satisfied a condition analogous to $(p_1 + p_2 + p_3)^2 = p_1{}^2 + p_2{}^2 + p_3{}^2$ of Sec. I. Thus, there continue to be too many solutions of Eqs. (29) and (30). In short, use of the Weyl tensor as the perturbed field does not seem to alleviate these difficulties with perturbation theory.

One might think that, alternatively, these difficulties are merely the result of our introduction of passive fields. So, let the $\underset{i}{\xi}{}^a$ be active. Include, with the perturbed fields, those resulting from the $\underset{i}{\xi}{}^a$, and include, with the perturbation equations, those resulting from the commutation relations on the $\underset{i}{\xi}{}^a$. Nothing essential then changes. It turns out that there are still solutions of the first-order perturbation equations for which the perturbed metric is a symmetrized derivative but which are not gauge solutions. And there are still solutions that cannot be extended to second order in perturbation theory.

In the discussion of apparent gauge in Sec. III, we treated only the case of Einstein's equation with vanishing sources. The introduction of sources, constructed from addi-

tional dynamic fields, does not significantly change the situation. An apparent gauge solution is again determined by a vector field τ^a, where the perturbed fields are those that result from applying $\mathcal{L}_\tau$ to each of the dynamic background fields. Invariance of the active fields under the $\underset{i}{\xi}{}^a$ then requires invariance of these perturbed fields under the $\underset{i}{\xi}{}^a$, which is turn requires invariance of each of the dynamic background fields under the $\underset{i}{\mathcal{L}}\,\tau^a$. Thus, apparent gauge again results whenever each of the $\underset{i}{\mathcal{L}}\,\tau^a$ is some linear combination, with constant coefficients, of the $\underset{i}{\xi}{}^a$. For true gauge, on the other hand, we again require $\underset{i}{\mathcal{L}}\,\tau^a = 0$. So, there will again be solutions that are apparent gauge but not true gauge. A further restriction in Sec. III was that all passive fields be vector fields satisfying Killing's equation. What happens when other types of passive fields are included? Are there versions of apparent gauge at higher order in perturbation theory?

Is the conjecture of Sec. IV, that at least three Killing fields are required for unreliability, true? We showed in Sec. IV that the spatially homogeneous space-times, with certain possible exceptions, are reliable. Are all the exceptional cases actually unreliable? Does the presence of sources, or passive fields other than Killing fields, increase or decrease the chances of reliability? Is reliability less prevalent as one goes to higher orders in perturbation theory? A possible conjecture is that if the nth-order perturbation equations are not reliable, then neither are the $(n+1)$st. One might attack this by trying to show that, whenever Eqs. (19) and (20) admit no solution for some t_{ab}, then that t_{ab} can be reached through some choice of the lower-order perturbations. Even for the simple static, plane-symmetric example the situation is not immediately clear. Is it true that the nth-order perturbation equations in this example are unreliable for all n?

APPENDIX: RELIABILITY OF SPATIALLY HOMOGENEOUS SPACE-TIME

Fix a three-dimensional Lie algebra. As in Sec. IV, denote by $\mathcal{S}$ the 12-dimensional manifold of pairs (q_{ab}, p^{ab}) of tensors over the vector space of the Lie algebra, and by $\mathcal{C}$ the subset of $\mathcal{S}$ consisting of those pairs for which the constraint equations (23) and (24) are satisfied, where $s^{ab} = s^{(ab)}$ and v_a are given by the decomposition (25) of the structure-constant tensor.

Denote by $H(q,p)$ and $H_a(q,p)$ the respective left sides of the constraint equations (23) and (24). Thus, $\mathcal{C}$ consists of those points at which these functions on $\mathcal{S}$ vanish. The issue of at which of its points $\mathcal{C}$ is a submanifold turns on the issue of at which points of $\mathcal{C}$ the gradients in $\mathcal{S}$ of these functions are linearly independent. So, consider the linear combination $\mu H + \nu^a H_a$. Equating to zero the gradient in $\mathcal{S}$ of this combination, i.e., equating to zero the q_{ab}- and p^{ab}-partial derivatives, keeping the structure-constant tensor fixed, we obtain

$$\mu\left[2(p^{ab}p^m{}_m - p^a{}_m p^{bm}) + \tfrac{3}{2}v^a v^b + (s^{mn}s_{mn} - \tfrac{1}{2}(s^m{}_m)^2)q^{ab} - 2s^a{}_m s^{bm} + s^{ab}s^m{}_m\right] + U^{(a}{}_m p^{b)m} = 0, \tag{A1}$$

$$2\mu\left[q_{ab}p^m{}_m - p_{ab}\right] + U_{(ab)} = 0, \tag{A2}$$

where we have set

$$U^a{}_b = s^{am}\epsilon_{mbn}\nu^n + \tfrac{3}{2}\nu^a v_b - \tfrac{1}{2}\delta^a{}_b \nu^m v_m. \tag{A3}$$

Let, q_{ab}, p^{ab}, μ, and ν^a satisfy Eqs. (23), (24), (A1), and (A2). We first derive, as a consequence, the following three equations:

$$\mu p_{ab} = 0, \tag{A4}$$

$$\mu v_a = 0 \tag{A5}$$

$$\mu(s^a{}_m s^{bm} - \tfrac{1}{2}s^{ab}s^m{}_m) = 0. \tag{A6}$$

To derive Eq. (A4), first contract Eq. (A2) with q^{ab} and use Eq. (A3), to obtain $\mu p^m{}_m = 0$. Then contract Eq. (A2) with p^{ab}, using this and Eqs. (A3) and (24). To derive Eq. (A5), first multiply Eq. (23) by μ, noting that, by Eq. (A4), the first two terms drop out. Were μv_a nonzero, then the third term would be negative, while, by the Jacobi relation $s^{ab}v_b = 0$, s^{ab} would have rank at most two, whence the last two terms together would be nonpositive. This contradiction establishes $\mu v_a = 0$. To derive Eq. (A6), multiply Eq. (A1) by μ, using (A4), (A5), and the result of multiplying Eq. (23) by μ. We next derive, as a further consequence, the following four equations:

$$U_{(ab)} = 0, \tag{A7}$$

$$U^{(a}{}_m p^{b)m} = 0, \tag{A8}$$

$$U^{(a}{}_m s^{b)m} = 0, \tag{A9}$$

$$U_{ab}v_c = 0. \tag{A10}$$

To derive Eq. (A7), use Eq. (A2) with Eq. (A4). To derive Eq. (A8), use Eq. (A1) with Eqs. (A4)–(A6). To derive Eq. (A9), contract Eq. (A3) with s^{cb} and symmetrize over c and a. The first two terms give zero, while the last also vanishes as a consequence of contracting Eq. (A7) with $\nu^a\nu^b$. To derive Eq. (A10), note that, by antisymmetry of U_{ab}, it suffices to check this equation when contracted with q^{ac} and when the antisymmetrized over all indices. But these both follow from Eq. (A3).

To summarize, we have shown so far that Eqs. (23), (24), (A1), and (A2) together imply Eqs. (A4)–(A10).

Now let (q_{ab}, p^{ab}) be a point of $\mathcal{C}$ at which $\mathcal{C}$ is not a submanifold. Then there must be some linear combination $\mu H + \nu^a H_a$, other than the zero function, whose gradient vanishes at this point. Since $\mu H + \nu^a H_a$ is not the zero function, it follows from Eqs. (23) and (24) that either μ or $U^a{}_b$ is nonzero. But for μ nonzero it follows from Eqs. (A4)–(A6), and for $U^a{}_b$ nonzero from (A7)–(A10), that this is one of the exceptional cases of Sec. IV. We have shown, then, that any point of $\mathcal{C}$ at which $\mathcal{C}$ is not a submanifold must be of the form (26) and (27) with $v_a = 0$. Equation (28) is merely the result of substituting Eqs. (26) and (27) into Eq. (23).

[1]E. Balbinski, S. Detweiler, L. Lindblom, and B. Schutz, Mon. Not. R. Astron. Soc. **213**, 553 (1985).
[2]D. Brill, "Isolated solutions in general relativity," in *Gravitation: Problems and Prospects* (Naukova Dumka, Kiev, 1972), p. 17.

[3]D. Brill and S. Deser, Commun. Math. Phys. **32**, 291 (1973).
[4]A. Fischer and J. Marsden, Bull. Am. Math. Soc. **79**, 997 (1973); A. Fischer and J. Marsden, Proc. Symp. Pure Math. **27**, 219 (1975).
[5]V. Moncrief, J. Math. Phys. **16**, 493 (1975); **17**, 1893 (1976).
[6]A. Fischer, J. Marsden, and V. Moncrief, Ann. Inst. H. Poincaré **33**, 147 (1980); J. Arms, J. Marsden, and V. Moncrief, Ann. Phys. **144**, 81 (1982).
[7]E. Kasner, Am. J. Math. **43**, 217 (1921).
[8]A. Taub, Ann. Math. **53**, 472 (1951).
[9]In order that these diffeomorphisms be defined globally, one must require that the relevant vector fields on $\mathcal{M}$ be complete. But for $n = 1$, this global condition can be avoided by "cutting off" the diffeomorphisms in λ. A similar argument may work for other n.
[10]R. Geroch, J. Math. Phys. **13**, 394 (1972).
[11]R. Wald, *General Relativity* (Chicago U. P., Chicago, 1984), pp. 168–178.
[12]G. F. R. Ellis and M. A. H. MacCallum, Commun. Math. Phys. **12**, 108 (1969).

Generation of new solutions of the stationary axisymmetric Einstein equations by a double complex function method

Zai-Zhe Zhong
Department of Physics, Liaoning Normal University, Dalian, Liaoning, The Peoples's Republic of China

(Received 31 December 1984; accepted for publication 24 May 1985)

A new approach to the solution of certain differential equations, the double complex function method, is developed, combining ordinary complex numbers and hyperbolic complex numbers. This method is applied to the theory of stationary axisymmetric Einstein equations in general relativity. A family of exact double solutions, double transformation groups, and n-soliton double solutions are obtained.

I. INTRODUCTION

Besides the ordinary complex numbers, of which the imaginary unit is i ($i^2 = -1$), there exist other generalized complex numbers,[1] namely, the *hyperbolic complex numbers* (or double numbers), of which the imaginary unit is ϵ, where $\epsilon^2 = +1$. Corresponding to the ordinary complex number field C, all hyperbolic complex numbers $a + \epsilon b$ (a, b are real) constitute a commutative ring H. Some applications of H in physics have been found.[2] Recently, Zhong[3,4] found some principal results of H concerning geometry and group theory. In this paper, we advance a new approach to the solution of certain differential equations, which we shall call the "double complex function (with real variables) method." Furthermore, we apply this method to the stationary axisymmetric Einstein vacuum field equations, and obtain a series of new results.

Tanabe[5] had discussed the Ernst equation,[6] and suggested the introduction of a second imaginary unit j, where $j^2 = -1$, and under conjugation $j^* = j$. Then a "new" Ernst solution $\mathscr{K}' = \rho f^{-1} + ij\omega$ is generated from an "old" Ernst solution $\mathscr{K} = f + i\psi$, where (ρ,z,ϕ) are canonical cylindrical coordinates, and $\partial_\rho \psi = \rho^{-1} f^2 \partial_z \omega$, $\partial_z \psi = -\rho^{-1} f^2 \partial_\rho \omega$. However, the Tanabe transformation $\mathscr{K} \to \mathscr{K}'$, in essence, is equivalent to the Kramer–Neugebauer[7,8] substitution $(f,\psi) \to (\rho f^{-1}, i\psi)$, because j and i are the same except under complex conjugation. It is known that, generally, without analytic continuation the Kramer–Neugebauer substitution with $\psi \neq 0$ cannot generate a new gravitational real solution of the field equations.[9] Similarly, a new real solution cannot be obtained directly by the Tanabe transformation with $\psi \neq 0$ unless the parameters in this solution can be analytically continued. This restricts greatly the usefulness of the Kramer–Neugebauer and Tanabe transformations. It should be noted that $ij = \epsilon$. Therefore $\mathscr{K}' = \rho f^{-1} + \epsilon\omega$, in fact, is a *hyperbolic complex Ernst potential*. In this paper, a double complex Ernst equation is just an organic combination of an ordinary complex Ernst equation and a hyperbolic complex Ernst equation, and a *double complex solution* of such an equation provides the simultaneous derivation of a pair of real solutions automatically related by a Kramer–Neugebauer substitution and analytic continuation.

In Sec. II we discuss the mathematical basis of the double complex function method. In Sec. III we discuss the double complex Ernst equation corresponding to a stationary axisymmetric Einstein vacuum field. Sections IV–VI contain applications of this new method. In Sec. IV a family of exact double solutions of the stationary axisymmetric Einstein equations are generated directly by the double complex function method, including double versions of the Timimatsu–Sato solutions. In Sec. V we discuss the double complex transformation groups and the infinite chain of solutions. Another family of exact double solutions and double transformation groups, in fact, are generated, and the Tanabe series is naturally generalized. In Sec. VI we discuss the n-soliton solutions. The Belinsky–Zakharov[10,11] soliton technique is generalized to the double complex Ernst equation, and a double n-soliton solution is given.

II. GENERAL MATHEMATICAL METHOD

Let J denote the double imaginary unit, i.e., J may be the ordinary unit i, it also may be the hyperbolic imaginary unit ϵ. An element $\mathscr{E} = a + \epsilon b$ is called a double complex number, where a and b are real. The conjugate element of $\mathscr{E}$ is $\mathscr{E}^* = a - Jb$. According to the common way, in the set DC of all double complex numbers we can introduce addition, multiplication, etc., so DC $= C$ when $J = i$, and DC $= H$ when $J = \epsilon$. The inverse of $a + ib$ exists if $|a| + |b| \neq 0$, and the inverse of $a + \epsilon b$ exists if $|a| - |b| \neq 0$. Therefore, in the following we suppose that $|a| \neq |b|$ when an abstract double complex number $\mathscr{E} = a + Jb$ is taken as a divisor. The $\mathscr{E}$ corresponds to a pair $(\mathscr{E}_C, \mathscr{E}_H)$, where

$$\mathscr{E}_C = \mathscr{E}(J = i) = a + ib, \tag{1}$$

$$\mathscr{E}_H = \mathscr{E}(J = \epsilon) = a + \epsilon b. \tag{2}$$

This pair $(\mathscr{E}_C, \mathscr{E}_H)$ plays an important role in this paper.

Now, we discuss the definition of a double complex function. Without loss of generality, we can only consider the case in which the function has only two real independent variables (x,y).

Definition 1: $\mathscr{E} = \mathscr{E}(J) = \mathscr{E}(x,y;J) = F(x,y;J) + J\Omega(x,y;J)$ is called a double complex function defined on a region $N \subset R^2$ if the following conditions are satisfied.

(a) $\mathscr{E}_C = F_C + i\Omega_C$ and $\mathscr{E}_H = F_H + \epsilon\Omega_H$ are an ordinary complex function and a hyperbolic complex function, respectively, where

 0022-2488/85/102589-07$02.50

$$F_C = F_C(x,y) = F(J = i),$$
$$F_H = F_H(x,y) = F(J = \epsilon),$$
$$\Omega_C = \Omega_C(x,y) = \Omega(J = i), \tag{3}$$
$$\Omega_H = \Omega_H(x,y) = \Omega(J = \epsilon)$$

all are real functions defined on N.

(b) J is an "analysis link" between $\mathscr{C}_C$ and $\mathscr{C}_H$. This means that if we take x and y as constants, and take J as a formal real independent variable, then $F(J)$ and $\Omega(J)$ are real analytic functions of J, and they can be written into infinite series form as follows:

$$F(J) = \sum_{n=0}^{\infty} a_n J^{2n},$$
$$\Omega(J) = \sum_{n=0}^{\infty} b_n J^{2n}, \tag{4}$$

where $a_n(x,y)$ and $b_n(x,y)$ are real functions, and the series Σa_n, $\Sigma(-1)^n a_n$, Σb_n, and $\Sigma(-1)^n b_n$ all are convergent on N.

Notice that the above condition (b) of analyticity links $\mathscr{C}_C$ with $\mathscr{C}_H$. As a result of condition (b), $\mathscr{C}(J)$ is not a mechanical combination of $\mathscr{C}_C$ and $\mathscr{C}_H$ in which J only plays a role as a discrete index ($J = i,\epsilon$), there is some analytic relation between the ordinary complex function $\mathscr{C}(x,y;i)$ and the hyperbolic complex function $\mathscr{C}(x,y;\epsilon)$.

In this paper, we shall use the double exponential function $\exp(\theta J)$. It is defined as follows:

$$\exp(\theta J) = \sum_{n=0}^{\infty} \frac{1}{n!} \theta^n J^n = c(\theta J) + J \cdot s(\theta J), \tag{5}$$

where $\theta(x,y)$ is a real function, and

$$c(\xi) = \sum_{n=0}^{\infty} \frac{1}{(2n)!} \xi^{2n},$$
$$s(\xi) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \xi^{2n}. \tag{6}$$

Therefore, we obtain a useful function pair

$$(e^{i\theta}, e^{\epsilon\theta}) = (\cos\theta + i\sin\theta,\ \cosh\theta + \epsilon\sinh\theta). \tag{7}$$

From here we see that $\mathscr{C}_H$, in general, is not the hyperbolic complex function in which the imaginary unit i is simply substituted by the imaginary unit ϵ.

Definition 2: If $\mathscr{C}(J) = F(J) + J\Omega(J)$ is a double complex function with $\mathrm{Im}(\mathscr{C}) = \Omega = 0$, then $\mathscr{C}$ is called a double real function. Next, if F_C, F_H, Ω_C, and Ω_H are real constants, then $\mathscr{C}(J)$ and $F(J)$ are called a double complex constant and a double real constant, respectively.

Thus, a double constant, in fact, is a number pair $(\mathscr{C}_C, \mathscr{C}_H)$ connected by a formal analytic function. In particular, a double complex number $a + Jb$, of course, is a double complex constant.

A differential equation

$$D(\mathscr{C}, \mathscr{C}^*, \partial_x, \partial_y, \ldots) = 0, \tag{8}$$

of which the unknown function $\mathscr{C}(J)$ is a double complex function, is called a double complex equation. Equation (8) corresponds to a combination of an ordinary complex equation

$$D_C = D(J = i) = 0 \tag{9a}$$

and a hyperbolic complex equation

$$D_H = D(J = \epsilon) = 0. \tag{9b}$$

Similarly, since the condition (b) of Definition 1, there is the same analytic relation between Eq. (9a) and Eq. (9b). Equation (8) can split up into two parts, i.e., the real part

$$R = \mathrm{Re}(D) = 0 \tag{10a}$$

and the imaginary part

$$I = \mathrm{Im}(D) = 0. \tag{10b}$$

By Eq. (10) we obtain two ordinary real equations, i.e.,

$$R_C = R(J = i) = 0,$$
$$I_C = I(J = i) = 0, \tag{11a}$$

and

$$R_H = R(J = \epsilon) = 0,$$
$$I_H = I(J = \epsilon) = 0. \tag{11b}$$

Notice that Eq. (11a) and Eq. (11b), in fact, are two distinct equations because $i^2 = -1$ but $\epsilon^2 = +1$. However, we are just interested in such a case that a physical field F can be equivalently described by Eq. (11a) or Eq. (11b). Let us suppose that the field F can be characterized enough by two real functions $f(x,y)$ and $\omega(x,y)$, i.e., all components of F are determined by f and ω. Let the field equation of F be

$$F_k(f,\omega,\partial_x,\partial_y,\ldots) = 0 \quad (k = 1,2). \tag{12}$$

Definition 3: Equation (8) is called a double complex field equation of field F if there exist two such transformations,

$$T: f \to \hat{f} = T(f),$$
$$V: (f,\omega) \to \hat{\omega} = V(f,\omega) = V_f(\omega), \tag{13}$$

that under transformations T and V the physical properties of F are not broken down, and the following conditions are satisfied.

(a) Equation (12) is a deformation of Eq. (11a) under the transformation

$$(F_C, \Omega_C) \to (f,\omega) = (F_C, V_{F_C}(\Omega_C)).$$

(b) Equation (12) is a deformation of Eq. (11b) under the transformation

$$(F_H, \Omega_H) \to (f,\omega) = (T(F_H), \Omega_H).$$

Now Eq. 8, indeed, is "double," since both Eq. (9a) and Eq. (9b) can be taken as the field equations of F.

Notice that though we can obtain a solution $(f, V_f^{-1}(\omega))$ of Eq. (11a) and a solution $(T^{-1}(f),\omega)$ of Eq. (11b) from a solution (f,ω) of Eq. (12), there generally is not a double complex solution $\mathscr{C}(J)$ of Eq. (8), which preserves both equations $\mathscr{C}_C = f + iV_f^{-1}(\omega)$ and $\mathscr{C}_H = T^{-1}(f) + \epsilon\omega$, unless $\omega = 0$. However, we can consider this problem otherwise as follows. Sometimes a solution of the ordinary complex equation (9a) can be generalized to a double complex function by some way (see Secs. III–VI). This means that we obtain a solution $\mathscr{C}(J) = F(J) + J\Omega(J)$ of Eq. (8). Therefore, we obtain two solutions $(F_C, V_{F_C}(\Omega_C))$ and $(T(F_H), \Omega_H)$ of Eq. (12).

According to the above illustration concerning $\mathscr{E}_C$ and $\mathscr{E}_H$, the above solutions of Eq. (12), in general, are different. In addition, if the right side of Eq. (8) is symmetric for $\mathscr{E}$ and $\mathscr{E}^*$, then $\mathscr{E}^*$ is also a double complex solution. In this case we obtain four solutions of field equation (12), i.e., $(F_C, V_{F_C}(\pm\Omega_C))$ and $(T(F_H), \pm\Omega_H)$. Now, a technique of generating "new" solution from an "old" solution of a differential equation, in fact, has been given by the double complex function. In the following sections of this paper, we shall make full use of this technique in general relativity.

Definition 4: A field F, which is described by Eq. (8), is called double complex normal if the following conditions are satisfied.

(a) The right side of Eq. (8) is symmetric for $\mathscr{E}$ and $\mathscr{E}^*$.

(b) When $\mathscr{E} = F + J\Omega$ is a double complex function, then $\tilde{\mathscr{E}} = T(F) + JV_F(\Omega)$ is also.

(c) $$T^2 = 1, \quad V_{T(f)} V_f = \tau = \pm 1, \tag{14}$$

where 1 denotes the identical transformation.

Definition 5: If F is double complex normal, and $\mathscr{E}(J) = F(J) + J\Omega(J)$ is a double complex solution of the field equation (8), then the solution $(\hat{f},\hat{\omega}) = (T(F_H), \tau\Omega_H)$ of the field equation (12) is called a dual solution of the solution $(f,\omega) = (F_C, V_{F_C}(\Omega_C))$ of the field equation (12). Sometimes, we simply call $\mathscr{E}_H$ the dual solution of $\mathscr{E}_C$.

Theorem 1: If $(\hat{f},\hat{\omega})$ is a dual solution of (f,ω), then (f,ω) is also a dual solution of $(\hat{f},\hat{\omega})$, i.e., the dual relation is reflexive.

Proof: According to the assumption, there is a double complex solution $\mathscr{E} = F + J\Omega$ which makes $F_C = f$, $V_f(\Omega_C) = \omega$, $T(F_H) = \hat{f}$, and $\tau\Omega_H = \hat{\omega}$. Since $(\hat{f},\hat{\omega})$ is a solution of Eq. (12), $\mathscr{E}_1 = \hat{f} + iV_{\hat{f}}^{-1}(\hat{\omega}) = T(F_H) + iV_{T(F_H)}^{-1}(\tau\Omega_H)$ is an ordinary complex solution of Eq. (9a). Similarly, $\mathscr{E}_2 = T^{-1}(f) + \epsilon\omega = T^{-1}(F_C) + \epsilon V_{F_C}(\Omega_C)$ is a hyperbolic complex solution of Eq. (9b). If the formal operation

$$(iJ)^{2n} = i^{2n}J^{2n} = (-1)^n J^{2n}$$

is allowed temporarily, then both functions $F'(J) = F(x,y;iJ)$ and $\Omega'(J) = \Omega(x,y;iJ)$ are significative and real by Definition 1. Let

$$\begin{aligned}\tilde{\mathscr{E}}(J) &= \tilde{F}(J) + J\tilde{\Omega}(J) \\ &= T(F'(J)) + J\cdot V_{F'(J)}(\Omega'(J)).\end{aligned} \tag{15}$$

According to the conditions (b) and (c) of Definition 4, we can examine directly that $\tilde{\mathscr{E}}$ is a double complex solution of Eq. (8), $\tilde{\mathscr{E}}_C = \mathscr{E}_1$, and $\tilde{\mathscr{E}}_H = \mathscr{E}_2$. Q.E.D.

The relation among the above solutions of a double complex normal field can be figured clearly as follows:

$$\begin{array}{ccccc}
(F_H, & \Omega_H) & \overset{H}{\leftarrow}\ \mathscr{E}\ \overset{C}{\rightarrow} & (F_C, & \Omega_C) \\
\downarrow T & \downarrow \tau & & \downarrow +1 & \downarrow V \\
(T(F_H), & \tau\Omega_H) & \overset{\text{dual}}{\leftrightarrow} & (F_C, & V_{F_C}(\Omega_C)). \\
\uparrow +1 & \uparrow V & & \uparrow T & \uparrow +1 \\
(T(F_H), & V_{F_H}(\Omega_H)) & \overset{C}{\leftarrow}\ \tilde{\mathscr{E}}\ \overset{H}{\rightarrow} & (T(F_C), & V_{F_C}(\Omega_C))
\end{array} \tag{16}$$

Since the dual solutions probably do not exist for some solutions, the dual relation is some correspondence between two proper subsets of the set of total solutions. The intersection of such two subsets, in fact, is not empty because there are some self-dual solutions (e.g., $\omega = 0$).

Lastly, when the double complex solution of some field F does not exist, we can use singly Eq. (11b) only. This means that the way of solving Eq. (9a) cannot be transformed to solve Eq. (9b), and vice versa.

III. THE STATIONARY AXISYMMETRIC EINSTEIN VACUUM FIELD IS DOUBLE COMPLEX NORMAL

Let us write the metric of a stationary axisymmetric Einstein vacuum field as the Papapetrou form

$$ds^2 = f(dt - \omega\, d\phi)^2 - f^{-1}[e^{2\gamma}(dz^2 + d\rho^2) + \rho^2\, d\phi^2], \tag{17}$$

where f, ω, and γ are real functions of ρ and z only. It is known that γ is determined by f and ω. The field equation is[6]

$$\begin{aligned} f\nabla^2 f &= \nabla f\cdot\nabla f - \rho^{-2}f^4\nabla\omega\cdot\nabla\omega, \\ \nabla\cdot(\rho^{-2}f^2\nabla\omega) &= 0, \end{aligned} \tag{18}$$

where ∇ denotes the understood three-dimensional divergence operator in the cylindrical coordinate system (ρ,z,ϕ). Transformations T and V are defined by

$$\begin{aligned} &T: \quad f \rightarrow T(f) = \rho f^{-1}, \\ &V: \quad (f,\psi) \rightarrow V_f(\psi) = \omega, \\ &\qquad \rho^{-1}f^2\nabla\omega = \mathbf{n}\times\nabla\psi, \end{aligned} \tag{19}$$

where $\mathbf{n}$ denotes the azimuthal direction, and ψ is real function of ρ, z only. Thus, under the transformation $(f,\omega)\rightarrow(F_C,\Omega_C) = (f, V_f^{-1}(\omega)) = (f,\psi)$, Eq. (18) changes into

$$\begin{aligned} F_C\nabla^2 F_C &= \nabla F_C\cdot\nabla F_C - \nabla\Omega_C\cdot\nabla\Omega_C, \\ 2\nabla F_C\cdot\nabla\Omega_C &= F_C\nabla^2\Omega_C, \end{aligned} \tag{20a}$$

and under the transformation $(f,\omega)\rightarrow(F_H,\Omega_H) = (T^{-1}(f),\omega) = (\rho f^{-1},\omega)$, Eq. (18) changes into

$$\begin{aligned} F_H\nabla^2 F_H &= \nabla F_H\cdot\nabla F_H + \nabla\Omega_H\cdot\nabla\Omega_H, \\ 2\nabla F_H\cdot\nabla\Omega_H &= F_H\nabla^2\Omega_H. \end{aligned} \tag{20b}$$

Therefore, we naturally obtain a double complex field equation

$$\mathrm{Re}(\mathscr{E})\nabla^2\mathscr{E} = \nabla\mathscr{E}\cdot\nabla\mathscr{E} \tag{21}$$

by the combination of Eq. (20a) and Eq. (20b), where $\mathscr{E}(J) = F(J) + J\Omega(J)$. We call Eq. (21) and $\mathscr{E}$ the double complex Ernst equation and the double complex Ernst potential, respectively. If $J = i$, then $\mathscr{E} = \mathscr{E}_C = F_C + i\Omega_C$, and Eq. (21) changes into the ordinary Ernst equation[6]

$$\mathrm{Re}(\mathscr{E}_C)\nabla^2\mathscr{E}_C = \nabla\mathscr{E}_C\cdot\nabla\mathscr{E}_C, \tag{22a}$$

which is equivalent to Eq. (20a). If $J = \epsilon$, then $\mathscr{E} = \mathscr{E}_H = F_H + \epsilon\Omega_H$ and Eq. (21) changes into

$$\mathrm{Re}(\mathscr{E}_H)\nabla^2\mathscr{E}_H = \nabla\mathscr{E}_H\cdot\nabla\mathscr{E}_H, \tag{22b}$$

which is equivalent to Eq. (20b). We call this equation the hyperbolic complex Ernst equation. Notice that though Eq. (22b) is equivalent to the ordinary Ernst equation (22a), sometimes to use Eq. (22b) is more convenient than to use

Eq. (22a), i.e., there is no need to introduce a twist potential ψ in Eq. (22b). In fact, the potential transformation V has been substituted by the direct transformation T.

Equation (21) is symmetric for $\mathscr{E}$ and $\mathscr{E}^*$. According to Eq. (19) and condition (b) of Definition 1, we see that condition (b) of Definition 4 is satisfied for T and V defined by Eq. (14). Evidently, we have $T^2 = 1$. In addition, we have

$$\begin{aligned}&\rho^{-1}(T(f))^2\nabla(V_{T(f)}V_f(\psi))\\&\quad=\rho^{-1}(\rho f^{-1})^2\nabla(V_{\rho f^{-1}}V_f(\psi))=\mathbf{n}\times\nabla V_f(\psi)\\&\quad=\mathbf{n}\times\rho f^{-2}(\mathbf{n}\times\nabla\psi)=\rho f^{-2}(-\nabla\psi),\end{aligned}$$

this means $\nabla(V_{T(f)}V_f(\psi)) = -\nabla\psi$. Since we, in fact, are interested in the function ω, without loss of generality we can select such V that $V_{T(f)}V_f = -1$. Therefore, we have the following theorem.

Theorem 2: The stationary axisymmetric Einstein vacuum field is double complex normal.

According to Sec. II we, in fact, have obtained a technique of generating new solutions. Namely, if $\mathscr{E}(J)$ is a double complex solution of Eq. (21), then we obtain four solutions of Eq. (18):

$$(f,\omega)=(F_C,V_{F_C}(\Omega_C)),\quad(\hat{f},\hat{\omega})=(\rho F_H^{-1},\Omega_H),$$
$$(f,-\omega),\quad(\hat{f},-\hat{\omega}).$$

Evidently, the sign before ω or $\hat{\omega}$ is immaterial for the gravitational field, so that in the following we discuss solution $\mathscr{E}_C$ and its dual solution $\mathscr{E}_H$ only.

Now, we discuss the relation between the above generation technique and the Kramer–Neugebauer theorem.[7,8] If (f,ω) is a solution of Eq. (18), then $(\rho f^{-1},i\psi)$ is also a solution of Eq. (18). However, this solution is not real. This solution will be real if it is possible to continue the parameters in the solution analytically so that the imaginary unit i can be compensated. From here we see that the Kramer–Neugebauer thorem only points out a possibility to find real solutions of Eq. (18), it cannot concretely give a real solution. However, the above double complex Ernst method is a process of solving the typical equation (21), in which the solutions of Eq. (18) are surely obtained in pairs. As for how to solve Eq. (21), there is much in the ordinary Ernst method that we can make use of (see Secs. IV–VI). In short, though both the Kramer–Neugebauer theorem and the double complex Ernst method make use of the transformation T, the use of the latter, evidently, is more advantageous than the former. Now we prove that the result of the double complex Ernst method also can be introduced directly by the Kramer–Neugebauer theorem.

Theorem 3: If $\mathscr{E}(J) = F(J) + J\Omega(J)$ is a double complex solution of Eq. (21), then $(\rho F_C^{-1},i\Omega_C)$ can generate $(\rho F_H^{-1},\Omega_H)$ by analytic continuation.

Proof: According to Definition 1, both $F(J)$ and $\Omega(J)$ are analytic functions of J if we take J as a formal real independent variable. Therefore, if we take J as a complex parameter and $J(\theta) = e^{i\theta}$, then both $F(x,y,\theta)$ and $\Delta(x,y,\theta) = J(\theta)\Omega(x,y,\theta)$ are analytic functions, and $\mathscr{E} = F_C + i\Omega_C = F(\theta=\pi/2) + i\hat{\Omega}(\theta=\pi/2)$ and $(f',\omega') = (\rho F_C^{-1},i\Omega_C)$ are solutions of Eq. (20a) and Eq. (18), respectively. When θ continuously changes from $\pi/2$ into 0, $(\rho[F(\theta)]^{-1},\Delta(\theta))$ changes from $(\rho F_C^{-1},i\Omega_C)$ into $(\rho F_H^{-1},\Omega_H)$. This means that the theorem is true.

As for the Tanabe[5] transformation $\mathscr{K} = f + i\psi \to \mathscr{K}' = \rho f^{-1} + ij\psi$, it, in fact, still is similar to the Kramer–Neugebauer substitution $(f,\omega)\to(\rho f^{-1},i\psi)$. Since $\mathrm{Re}(\mathscr{E}_C)\cdot\mathrm{Re}(\mathscr{E}_H)\neq\rho$ unless $\omega = 0$, the Tanabe transformation is essentially different from the transformation $\mathscr{E}_C\to\mathscr{E}_H$.

IV. APPLICATION A: THE FAMILY OF DUAL NEW EXACT SOLUTIONS GENERATED DIRECTLY

In this section, we discuss a family of the new exact solutions, which are generated directly by double complex Ernst method from some known solutions. Under the transformation

$$\mathscr{E}(J) = [\xi(J) - 1]/[\xi(J) + 1], \tag{23}$$

the double complex Ernst equation (21) changes into

$$(\xi\xi^* - 1)\nabla^2\xi = 2\xi^*\nabla\xi\cdot\nabla\xi. \tag{24}$$

When $J = i$, Eq. (24) changes into the ordinary complex form

$$(\xi_C\xi_C^* - 1)\nabla^2\xi_C = 2\xi_C^*\nabla\xi_C\cdot\nabla\xi_C. \tag{25a}$$

When $J = \epsilon$, Eq. (24) changes into the hyperbolic complex form

$$(\xi_H\xi_H^* - 1)\nabla^2\xi_H = 2\xi_H^*\nabla\xi_H\cdot\nabla\xi_H. \tag{25b}$$

(a) We discuss the dual family DSTδ of the Sato–Timimatsu family STδ. The Sato–Tomimatsu[12,13] family of exact solutions of Eq. (25a) is

$$\xi_C = \alpha(x,y;p,q,\delta)/\beta(x,y;p,q,\delta), \tag{26}$$
$$p^2 + q^2 = 1, \tag{27}$$

where α and β are ordinary complex polynomials of the prolate spheroidal coordinates x and y, and p and q are real numbers. The concrete expressions of α and β for $\delta = 1,2,3,4$ have been given by Refs. 13 and 14. It is known[13] that in α and β the coefficients of the real terms include q with an even power index, and those of the pure-imaginary terms include q with an odd power index. Therefore, under the substitution

$$(p,q^{2n},iq^{2n-1})\to(\zeta,(-1)^n\eta^{2n},(-1)^{n-1}\eta^{2n-1}) \tag{28}$$

(where ζ and η are formal variables, and $n\geqslant 1$), α and β change into real polynomials $A(x,y,\zeta,\eta;\delta)$ and $B(x,y,\zeta,\eta;\delta)$, respectively. Let

$$\xi = A(x,y,\zeta,\eta;\delta)/B(x,y,\zeta,\eta;\delta), \tag{29}$$
$$\zeta^2 - \eta^2 = 1. \tag{30}$$

It is clear that when $\zeta = p$ and $\eta = iq$, Eqs. (29) and (30) change exactly into Eqs. (26) and (27), respectively. Further, if $\zeta = \hat{p}$ and $\eta = \epsilon\hat{q}$, where $\hat{p}$ and $\hat{q}$ are real, then we obtain a hyperbolic complex function

$$\zeta_H = \frac{\hat{\alpha}(x,y;\hat{p},\hat{q},\delta)}{\hat{\beta}(x,y;\hat{p},\hat{q},\delta)} = \frac{A(x,y,\hat{p},\epsilon\hat{q},\delta)}{B(x,y,\hat{p},\epsilon\hat{q},\delta)}, \tag{31}$$
$$\hat{p}^2 - \hat{q}^2 = 1, \tag{32}$$

and ξ_H is exactly a hyperbolic complex solution of Eq. (25b).

Theorem 4: If λ is an arbitrary real number, then

$$\xi_{H\lambda} = \frac{\hat{\alpha}(x,y,\cosh\lambda,\sinh\lambda,\delta)}{\hat{\beta}(x,y,\cosh\lambda,\sinh\lambda,\delta)} \tag{33}$$

is the DSTδ of the STδ

$$\xi_{C\lambda} = \frac{\alpha(x,y;\cos\lambda,\sin\lambda,\delta)}{\beta(x,y;\cos\lambda,\sin\lambda,\delta)}. \tag{34}$$

Proof: According to Eq. (6), the expressions $c(J)$ and $s(J)$ may be regarded as an analytic functions of J, and

$$[c(J)]^2 - J^2[s(J)]^2 = 1. \tag{35}$$

Let

$$\xi(J) = \frac{A(x,y,c(J),J\cdot s(J);\delta)}{B(x,y,c(J),J\cdot s(J);\delta)}. \tag{36}$$

Evidently, $\xi(J)$ is a double complex solution of Eq. (24), and $\xi(J=i) = \xi_{C\lambda}, \xi(J=\epsilon) = \xi_{H\lambda}$. Q.E.D.

Example: Let $p = \cos\lambda$, $q = \sin\lambda$, $\hat{p} = \cosh\lambda$, $\hat{q} = \sinh\lambda$. When $\delta = 2$,

$$\text{ST2: } \xi_C = \frac{p^2x^4 + q^2y^4 - 1 - 2ipqxy(x^2 - y^2)}{2px(x^2 - 1) - 2iqy(1 - y^2)}, \tag{37}$$

$$\text{DST2: } \xi_H = \frac{\hat{p}^2x^4 - \hat{q}^2y^4 - 1 - 2\epsilon\hat{p}\hat{q}(x^2 - y^2)}{2\hat{p}x(x^2 - 1) - 2\epsilon\hat{q}y(1 - y^2)}. \tag{38}$$

When $\delta = 1$ we obtain the Kerr solution[6] and its dual solution

$$\xi_C = px + iqy, \tag{39}$$

$$\xi_H = \hat{p}x + \epsilon\hat{q}y. \tag{40}$$

When $\delta = 1, \lambda = 0$, we obtain the Schwarzschild solution[6]

$$\xi'_C = \xi'_H = x. \tag{41}$$

Thus the Schwarzschild solution is "self-dual." It should be noted that though $\xi'_C = \xi'_H$, ξ'_C and ξ'_H correspond to the solutions $((x-1)/(x+1),0)$ and $((x+1)/(x-1)\rho,0)$, respectively.

(b) If $\xi'(J)$ is a double complex solution of Eq. (24), then

$$\xi(J) = \exp(\kappa J)\xi'(J), \tag{42}$$

evidently, is also a double complex solution of Eq. (24), where κ is an arbitrary real number. Thus, if ξ'_H is the dual solution of ξ'_C, then

$$(\xi_C,\xi_H) = (e^{i\kappa}\xi'_C, e^{\epsilon\kappa}\xi'_H) \tag{43}$$

is also a dual solution pair.

Example: By ST1 we obtain the Demianski–Newman solution[14] and its dual solution, i.e.,

$$\xi_C = e^{i\kappa}(x\cos\lambda + iy\sin\lambda),$$
$$\xi_H = e^{\epsilon\kappa}(x\cosh\lambda + \epsilon y\sinh\lambda). \tag{44}$$

In particular, let $\kappa = 0$, then we obtain the NUT–Taub solution $\xi_C = e^{i\kappa}x$ and the dual solution $\xi_H = e^{\epsilon\kappa}x$.

(c) According to (a), (b), and Sec. III, a family of new exact solutions of the stationary axisymmetric Einstein equation (18), which corresponds to the above ξ_H, can be written directly as follows:

$$(f,\omega) = \left(\rho\left(\mathrm{Re}\,\frac{\xi_H - 1}{\xi_H + 1}\right)^{-1}, \quad \mathrm{Im}\,\frac{\xi_H - 1}{\xi_H + 1}\right). \tag{45}$$

It should be noted that, generally, the solution (45) does not belong to the family of ST.

Example: For DST1, we obtain

$$f = \frac{(\hat{p}x+1)^2 - \hat{q}^2y^2}{\hat{p}^2x^2 - \hat{q}^2y^2 - 1}\sqrt{(x^2-1)(1-y^2)},$$
$$\omega = \frac{2\hat{q}y}{(\hat{p}x+1)^2 - \hat{q}^2y^2}, \quad \hat{p}^2 - \hat{q}^2 = 1. \tag{46}$$

V. APPLICATION B: THE DOUBLE COMPLEX TRANSFORMATION GROUPS AND THE INFINITE CHAIN OF SOLUTIONS

In this section we discuss the transformation groups G acting on the double complex solutions, i.e., the double complex linear transformation groups which link up the known ordinary complex transformation groups and the new transformation groups—hyperbolic complex transformation groups.[3,4]

Let

$$\mathscr{E} = (u - w)/(u + w), \tag{47}$$

where both u and w are double complex functions. Therefore, Eq. (21) changes into[15]

$$z^\alpha z^*_\alpha \nabla^2 z^\beta = 2z^*_\alpha \nabla z^\alpha \cdot \nabla z^\beta, \tag{48}$$

where $\alpha,\beta = 1,2$, $(z^1z^2) = (u,w)$, signature $\eta_{\alpha\beta} = \mathrm{diag}(1,-1)$ and $z_\alpha = \eta_{\alpha\beta}z^\beta$. A double complex linear tranformation K: $z \to z' = K(z)$ is defined by

$$z'^\beta = K^\beta_\alpha z^\alpha, \tag{49}$$

where $K(J) = (K^\beta_\alpha(J))$ is a 2×2 double complex matrix, i.e., every $K^\beta_\alpha(J)$ is a double complex constant. A transformation K is called unitary if the form of K is

$$K(J) = \begin{bmatrix} a(J) & b^*(J) \\ b(J) & a^*(J) \end{bmatrix}, \quad \det(K) = 1. \tag{50}$$

Evidently, the set of all unitary K forms a group SU(1,1;J), and we call it the double complex unitary transformation group. Under the action of an element of SU(1,1;J), Eq. (48) is covariant. Therefore, if z is a known solution of Eq. (48), then $z' = K(z)$ [$K \in$ SU(1,1;J)] is a new solution. Notice that

$$\mathrm{SU}(1,1;J=i) = \mathrm{SU}(1,1,C),$$
$$\mathrm{SU}(1,1;J=\epsilon) = \mathrm{SU}(1,1,H), \tag{51}$$

where SU(1,1;C) is just the ordinary complex group SU(1,1) and SU(1,1;H) is the hyperbolic complex unitary group.[3,4] It is very interesting that though Eq. (22a) is very similar to Eq. (22b), the internal transformation group of the set $\{\mathscr{E}_C\}$ is not isomorphic to the internal transformation group of the set $\{\mathscr{E}_H\}$. In fact, Zhong[3,4] has proved the Lie group SU(1,1;H) to be isomorphic to the group SL(2,R).

Equation (21) can be obtained from the Lagrangian

$$\mathscr{L}(J) = [\nabla\mathscr{E}(J)\cdot\nabla\mathscr{E}(J)]/\mathrm{Re}(\mathscr{E}(J)) \tag{52}$$

by means of the variational principle. The $\mathscr{L}(J)$ is invariant under the transformation

$$K(J):\mathscr{E}(J) \to \mathscr{E}'(J),$$
$$\mathscr{E}'(J) = (a\mathscr{E} + Jb)/(J_C\mathscr{E} + d), \quad ad - J^2bc = 1, \tag{53}$$

where a, b, c, and d are real numbers. We can examine directly that the set of all transformations K forms a group, and this group is isomorphic to the matrix group $G_K(J)$ consisting of those 2×2 matrices having the form

$$K(J)=\begin{bmatrix} a & Jb \\ Jc & d \end{bmatrix}, \quad \det(K)=1. \tag{54}$$

Under the action of an element $K\in G_k$, Eq. (21) is covariant. Therefore, if $\mathscr{E}$ is a known double complex solution, then a new double complex solution $\mathscr{E}'=K(\mathscr{E})$ is generated by $K\in G_k$. This generation technique, evidently, is a generalization of the known ordinary complex method. In fact, if $J=i$, then

$$K_C:\mathscr{E}_C\to\mathscr{E}'_C,$$
$$\mathscr{E}'_C=(a\mathscr{E}_C+ib)/(ic\mathscr{E}+d), \quad ad+bc=1. \tag{55}$$

This is a known transformation[16] acting on the ordinary complex Ernst solution $\mathscr{E}_C$.

When $J=\epsilon$ we obtain

$$K_H:\mathscr{E}_H\to\mathscr{E}'_H,$$
$$\mathscr{E}'_H=(a\mathscr{E}_H+\epsilon b)/(\epsilon c\mathscr{E}+d), \quad ad-bc=1. \tag{56}$$

According to Sec. III, a solution (f,ω) of Eq. (18) uniquely corresponds to a hyperbolic complex Ernst solution $\mathscr{D}_H=\rho f^{-1}+\epsilon\omega$. Evidently, the group $G_k(J=\epsilon)$ is isomorphic to the group SL(2,R). Therefore, we prove once again a known result[16] in a new form, i.e., the group SL(2,R) is the group of transformations acting on the stationary axisymmetric gravitational vacuum field solutions. However, the action of an element $K'\in$ SL(2,R) should be explained as the action of K in Eq. (56) concerning a hyperbolic complex Ernst solution $\mathscr{E}_H$ now, where K corresponds to K' under the isomorphic mapping from $G_k(J=\epsilon)$ to SL(2,R).

To sum up, we have the following.

Theorem 6: If $\mathscr{E}(J)$ is a double complex solution of Eq. (21), then $K(\mathscr{E})$ is also a double complex solution of Eq. (21), and $K_H(\mathscr{E}_H)$ is the dual solution of $K_C(\mathscr{E}_C)$, where K is a transformation defined by Eq. (49), by Eq. (50), or by Eq. (54).

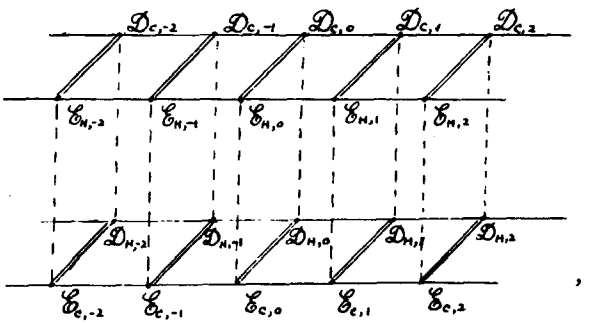
According to Sec. III, corresponding to a solution (f,ω) and dual solution $(\hat{f},\hat{\omega})$ of Eq. (18) we can write out four Ernst solutions as follows:

$$\mathscr{E}_C=f+i\psi, \quad \mathscr{E}_H=\rho\hat{f}^{-1}+\epsilon\hat{\omega},$$
$$\mathscr{D}_H=\rho f^{-1}+\epsilon\omega, \quad \mathscr{D}_C=\hat{f}+i\hat{\psi}. \tag{57}$$

Notice that $\mathscr{D}_H$ and $\mathscr{D}_C$ are equivalent to $\mathscr{E}_C$ and $\mathscr{E}_H$, respectively, in view of solving Eq. (18). The relation among these solutions can be figured as follows:

(58)

where the twin line and the dotted line denote the equivalent relation and the dual relation, respectively. Now, applying the transformation K, we obtain the following "infinite prismatic chain of new solutions":



(59)

where $\mathscr{E}_{C,0}$ is a known ordinary complex "seed solution," and a solid line denotes a generation by means of the transformation K defined by Eq. (49), by Eq. (50), or by Eq. (53). According to Sec. I and Sec. III, the Tanabe[5] series of new solutions, in fact, corresponds to a toothlike part of the chain (59):

$$\cdots\cdots\mathscr{D}_{H,-2}-\mathscr{D}_{H,-1}=\mathscr{E}_{C,-1}-\mathscr{E}_{C,0}=\mathscr{D}_{H,0}-\mathscr{D}_{H,1}=\mathscr{E}_{C,1}-\mathscr{E}_{C,2}\cdots\cdots. \tag{60}$$

VI. APPLICATION C: THE DOUBLE-SOLITON SOLUTIONS

Let

$$g_C=\frac{1}{F_C}\begin{bmatrix} 1 & \Omega_C \\ \Omega_C & \Omega_C^2+F_C^2 \end{bmatrix}, \tag{61}$$

then it can be proved that Eq. (20a) is equivalent to the equation

$$\partial_\rho(\rho\,\partial_\rho g_C\cdot g_C^{-1})+\partial_z(\rho\,\partial_z g_C\cdot g_C^{-1})=0. \tag{62}$$

Similarly, let

$$g_H=\frac{1}{F_H}\begin{bmatrix} 1 & \Omega_H \\ \Omega_H & \Omega_H^2-F_H^2 \end{bmatrix}, \tag{63}$$

then Eq. (20b) is equivalent to the equation

$$\partial_\rho(\rho\,\partial_\rho g_H\cdot g_H^{-1})+\partial_z(\rho\,\partial_z g_H\cdot g_H^{-1})=0. \tag{64}$$

Therefore, we obtain a double Belinsky–Zakharov equation

$$\partial_\rho U+\partial_z W=0, \tag{65}$$

where

$$U(\rho,z;J)=\partial_\rho g(\rho,z;J)\cdot[g(\rho,z;J)]^{-1},$$
$$W(\rho,z;J)=\partial_z g(\rho,z;J)\cdot[g(\rho,z;J)]^{-1}, \tag{66}$$

$$g(J)=\frac{1}{F(J)}\begin{bmatrix} 1 & \Omega(J) \\ \Omega(J) & [\Omega(J)]^2-J^2[F(J)]^2 \end{bmatrix}. \tag{67}$$

Equation (65) is equivalent to the double complex Ernst equation (21). Notice that $g(J)$ is real, therefore the Belinsky–Zakharov[10,11] soliton technique can be generalized to solve Eq. (21) as follows. Let $\mathscr{E}_0(J)=F_0(J)+J\Omega_0(J)$ be a known double complex solution of Eq. (21), and $g_0(J)$ be defined by $\mathscr{E}_0$ as in Eq. (67). Let us consider the equations

$$\left(\partial_\rho + \frac{2\lambda\rho}{\lambda^2+\rho^2}\partial_\lambda\right)\Psi = \frac{\rho U + \lambda W}{\lambda^2+\rho^2}\Psi,$$
$$\left(\partial_z - \frac{2\lambda^2}{\lambda^2+\rho^2}\partial_\lambda\right)\Psi = \frac{\rho W - \lambda U}{\lambda^2+\rho^2}\Psi, \tag{68}$$

where the spectral parameter λ is an ordinary complex number, and

$$\Psi(\lambda,\rho,z;J) = A(\lambda,\rho,z;J) + iB(\lambda,\rho,z;J) \tag{69}$$

is a "double ordinary complex function matrix," i.e., both A and B are 2×2 matrices of which the elements are double real functions (Definition 2). We choose such a solution Ψ_0 of Eq. (68) that

$$\Psi_0(\lambda = 0;J) = g_0(J), \tag{70}$$

thus a new double n-soliton solution $g(J)$ of Eq. (62) can be generated from the double soliton solution $g_0(J)$ as follows:

$$g = (-J^2 \det g')^{-1/2} g',$$
$$g' = \left(1 - \sum_{K=1}^{n} \frac{R_k}{\mu_k}\right) g_0, \tag{71}$$
$$\mu_k = \nu_k - z \pm [(\nu_k - z)^2 + \rho]^{1/2}.$$

Here $\nu_k(J)$ is an arbitrary double real constant, $R_k(J)$ is a 2×2 double real matrix

$$(R_k)^A_B = n^{(k)A} m^{(k)}_B \quad (A,B = 1,2), \tag{72}$$

and the double real constants $m^{(k)}(J)$, $n^{(k)}(J)$ are defined by the following equations:

$$m^{(k)}_A = m^{0(k)}_C ([\Psi_0(\lambda = \mu_k)]^{-1})^C_A,$$
$$m^{0(k)}_C = \text{arbitrary double real constant},$$
$$\sum_{l=1}^{n} \Gamma_{kl} n^{(l)A} = \frac{1}{\mu_k} m^{(k)}_C (g_0)^C_B \delta^{AB},$$
$$\Gamma_{kl} = m^{(k)}_C (g_0)^C_A m^{(k)}_A / (\rho^2 + \mu_k \mu_l). \tag{73}$$

This new n-soliton can be written out in the form of the double complex Ernst solution $\mathscr{E}(J)$,

$$\mathscr{E} = [(g)^1_1]^{-1} + J(g)^1_2 [(g)^1_1]^{-1}. \tag{74}$$

Evidently, the n-soliton solution $\mathscr{E}_H$ is the dual solution of n-soliton solution $\mathscr{E}_C$ in the sense of Definition 5.

Example: When $n = 2$, and $\mathscr{E}_0$ is the Minkowski solution, i.e., $\mathscr{E}_0 = 1$, we choose $\Psi_0 = g_0$,

$$g_0(J) = \begin{bmatrix} 1 & 0 \\ 0 & -J^2 \end{bmatrix}. \tag{75}$$

We introduce the parameters[17]

$$(m^{0(1)}_1)^2 + (m^{0(1)}_2)^2 = \sigma_1^2,$$
$$(m^{0(2)}_1)^2 + (m^{0(2)}_2)^2 = \sigma_2^2, \tag{76}$$
$$m^{0(1)}_1 m^{0(2)}_1 + m^{0(1)}_2 m^{0(2)}_2 = \sigma_1\sigma_2 \cdot c[(\varphi_1 - \varphi_2)J],$$

where $c[\lambda J]$ is defined by Eq. (6). By a direct calculation we obtain a new double two-soliton $\mathscr{E} = (\xi - 1)/(\xi + 1)$,

$$\xi(J) = e^{(\varphi_1 + \varphi_2)J}(x \cdot c[(\varphi_1 - \varphi_2)J] + Jy \cdot s[(\varphi_1 - \varphi_2)J]), \tag{77}$$

where (x,y) are prolate spheroidal coordinates. This is just the same result as can be obtained in Eq. (44) with $\kappa = \varphi_1 + \varphi_2$ and $\lambda = \varphi_1 - \varphi_2$.

VII. CONCLUSION AND DISCUSSION

It may be useful to attempt to apply the double complex function method to certain equations of mathematical physics. A solution of the double complex equation (8) always corresponds to two solutions of Eq. (12). Thus we have obtained a new technique of generating solutions of a differential equation. This method has been used to solve the stationary axisymmetric Einstein vacuum field equation, and the relation between the double complex Ernst method and the Kramer–Neugebauer theorem has been discussed. By the above method we have obtained a family of double exact solutions, the double complex transformation groups, the infinite chain of solutions, and the double n-soliton solutions of the stationary axisymmetric Einstein vacuum field equation.

As for the Einstein–Maxwell field, its ordinary complex Ernst equation is[18]

$$(\mathrm{Re}(\mathscr{E}) + |\Phi|^2)\nabla^2\mathscr{E} = (\nabla\mathscr{E} + 2\Phi^* \nabla\Phi)\cdot\nabla\mathscr{E},$$
$$(\mathrm{Re}(\mathscr{E}) + |\Phi|^2)\nabla^2\Phi = (\nabla\mathscr{E} + 2\Phi^* \nabla\Phi)\cdot\nabla\Phi, \tag{78}$$

where Φ is determined by the potential of the electromagnetic field. It is a pity that we have not yet found how Eq. (78) can be extended to a double complex form. However, according to a result of Ernst,[18] under very special conditions Eq. (78) is equivalent to an ordinary complex Ernst equation. Therefore, in this limited sense, we can also apply the double complex function method to consider the Einstein–Maxwell field.

[1] M. Yaglom, *Complex Numbers in Geometry* (Academic, London, 1968).
[2] G. Kunstatter, J. W. Moffat, and R. Malzan, J. Math. Phys. **24**, 866 (1983).
[3] Z-Z. Zhong, J. Math. Phys. **25**, 3538 (1984).
[4] Z-Z. Zhong, J. Math. Phys. **26**, 404 (1985).
[5] Y. Tanabe, J. Math. Phys. **20**, 1486 (1979).
[6] F. J. Ernst, Phys. Rev. **167**, 1175 (1968).
[7] G. Neugebauer and D. Kramer, Ann. Phys. **24**, 62 (1969).
[8] D. Kramer, H. Stephani, E. Herlt, M. MacCallum, and E. Schmutzer, *Exact Solution of Einstein's Field Equations* (Cambridge U.P., Cambridge, 1980).
[9] If $\psi = 0$, a vacuum solution generates a solution of Einstein–Maxwell equations by the Bonnor transformation [W. B. Bonnor, Z. Phys. **161**, 439 (1961)]. In addition, the "seed solution" of a Tanabe series, in fact, must be a Weyl solution (e.g., Schwarzschild solution).
[10] V. A. Belinsky and V. E. Zakharov, Zh. Eksp. Teor. Fiz. **75**, 1953 (1978).
[11] V. A. Belinsky and V. E. Zakharov, Zh. Eksp. Teor. Fiz. **77**, 3 (1979).
[12] A. Tomimatsu and H. Sato, Phys. Rev. Lett. **29**, 1344 (1972).
[13] A. Tomimatsu and H. Sato, Prog. Theor. Phys. **50**, 95 (1973).
[14] M. Demianski and E. T. Newman, Bull. Acad. Polon. Sci. Ser. Math. Astro. Phys. **14**, 653 (1966).
[15] M. Carmeli, *Classical Field: General Relativity and Gauge Theory* (New York, 1982).
[16] Y. Nakamura, J. Math. Phys. **24**, 606 (1983).
[17] J. Gruszczak, J. Phys. A: Math. Gen. **14**, 3247 (1983).
[18] F. J. Ernst, Phys. Rev. **168**, 1415 (1968).

On unifying electromagnetism and gravitation without curvature

Dieter Schütze
Institut für Mathematik, Ruhr-Universität Bochum, D-4630 Bochum 1, Federal Republic of Germany

(Received 4 October 1984; accepted for publication 12 April 1985)

This paper is devoted to a five-dimensional unification of the gravitational theory of Hayashi and Shirafuji with electromagnetism. Interference effects are found between gravitational contributions of matter spin and electromagnetism. This unification becomes the classical Kaluza–Klein theory if contributions of the torsion tensor related with spin are neglected.

I. INTRODUCTION

There are some attempts for extending Einstein's gravitational theory to contain matter spin. This forces a change in the space-time geometry. While Einstein used the torsion-free Riemann space-time V_4, later a Riemann–Cartan space-time geometry[1] was developed with nonvanishing torsion and curvature. This Riemann–Cartan space-time U_4 is a paracompact, Hausdorff, connected C^∞ four-dimensional manifold endowed with a locally Lorentzian metric and a linear affine connection, which is metric.

From the general Riemann–Cartan space-time two more special space-time models can be derived. One is the Riemann space-time V_4, which is obtained from U_4 by setting the torsion tensor to be identically vanishing. Another interesting model is the Weitzenböck space-time M_4 (see Refs. 2), which is obtained from U_4 by setting the curvature equal to zero.

Or, to put it equivalently, the Weitzenböck space-time is obtained by requiring the U_4 to admit absolute parallelism, i.e., to have a quadruplet (specified by $k = 1,2,3,4$) of linearly independent parallel vector fields $\{b_k{}^\mu\}$ with a vanishing covariant derivative:

$$D_\nu b_k{}^\lambda = \partial_\nu b_k{}^\lambda + \Gamma_\mu{}^\lambda{}_\nu b_k{}^\mu = 0.$$

Here, $\partial_\nu \equiv \partial/\partial x^\nu$ denotes the derivative in an open neighborhood $U \subset M_4$, $x \in U$, with components x^ν.

Solving this equation one finds the nonsymmetric affine connection

$$\Gamma_\mu{}^\lambda{}_\nu = b_k{}^\lambda \partial_\nu b^k{}_\mu \equiv b_k{}^\lambda b^k{}_{\mu,\nu}.$$

For the torsion tensor we get

$$T_{\mu\cdot\nu}^{\ \ \lambda} = b_k{}^\lambda (\partial_\nu b^k{}_\mu - \partial_\mu b^k{}_\nu).$$

Hayashi and Shirafuji[3] constructed a theory of gravitation on this space-time. The main consequences of their theory are analogous to those of Einstein's general relativity so far as macroscopic phenomena are concerned. This theory attributes gravitation to the torsion tensor formed from the parallel vector fields. This formation is a great advantage of the theory because those vector fields are connected by means of a local Lorentz transformation with "four-beins," which are used in field theoretical descriptions of general relativity.

Moffat[4] made use of a space-time with a nonsymmetric Hermitian metric tensor to describe a gravitational theory with torsion. He took fermion currents as the second gravitational charge (the first is the mass). Furthermore, he constructed the five-dimensional unification of his theory with electromagnetism.[5,6] Later Kalinowski[7] achieved this unification, too. One gets interference effects between gravitational and electromagnetic fields.

In the following we will consider the Kaluza–Klein approach[8–10] to the theory of Hayashi and Shirafuji. In Sec. II we shall introduce the framework of principal fiber bundles reduced to the U(1)-structure group. We shall identify the bundle connection projected into the base manifold M_4 with the electromagnetic four-potential A_μ, which one knows from relativistic electromagnetism. This connection transforms under changing local sections as gauge potentials A_μ under gauge transformations.

In Sec. III we shall consider the bundle metric. One can divide every tangent vector in the bundle in view of a given connection in only one way into two parts—horizontal and vertical. We project the horizontal part onto space-time and the vertical one into the Lie algebra of the gauge group. One can measure now independently the length of both parts and after this add these two results. This construction was first emphasized by Trautman.[11]

Having the bundle metrized in this way we introduce liner connections. They are compatible in some sense with the metric, i.e., their covariant derivatives vanish. The coefficients of the torsion tensor can now be calculated.

In Sec. IV A we will summarize the used gravitational theory.[3] In Sec. IV B it will be extended in a Kaluza–Klein approach. Variations of the extended Lagrangian will be taken with respect to the gravitational and electromagnetic components $b^k{}_\mu$ (resp. A_μ) of the used parallel vector fields. Setting a torsion tensor contribution equal to zero, one gets the well-known Maxwell and Einstein equations.

II. SPECIALIZED PRINCIPAL FIBER BUNDLE

One may introduce a principal fiber bundle[12] as follows. The base space M is the four-dimensional Weitzenböck manifold used in Ref. 3. The bundle space P is a locally trivial five-dimensional Weitzenböck manifold. This means, for any $x \in M$, there exists an open neighborhood $U \subset M$ such that the set $\pi^{-1}(U) = \{ p \in P \mid \pi(p) \in U \}$ is isomorphic to $U \times \mathrm{U}(1)$ under the smooth projection $\pi: P \to M$. Here, U(1) is the Abelian group $\mathrm{U}(1) \equiv \{ g \mid g = e^{-it}, t \in \mathbf{R} \}$.

There exists a diffeomorphism

$$\Phi_U: \pi^{-1}(U) \to U \times \mathrm{U}(1), \quad \Phi_U(p) = (\pi(p), \overset{\triangle}{\Phi}(p)),$$

such that

$$\overset{\triangle}{\Phi}(pg) = \overset{\triangle}{\Phi}(p)g, \quad \forall g \in \mathrm{U}(1), \quad p \in P.$$

 0022-2488/85/102596-05$02.50

Consider a basis of the tangent space

$$\{E_\mu\} = \{(\partial/\partial x^\mu)_x\}, \quad \text{at } \pi(p) = x \in M,$$

in the base manifold, $\mu = 1,...,4$.

For a basis[13] of U(1) choose the vector field ζ_5 which satisfies

$$L_g \zeta_5 = \zeta_5, \quad \forall g \in \mathrm{U}(1).$$

The left multiplication L_g is induced by the maps $\mathrm{U}(1) \to \mathrm{U}(1)$, $g' \to gg'$.

One can view ζ_5 as a one-dimensional Lie algebra basis of U(1).

Let e denote the group unit element. Each ζ_5 generates a one-parameter group of transformations $\varphi(t)$ on U(1) with

$$\varphi(t) \cdot g = \varphi(t) \cdot L_g \cdot e = L_g\, \varphi(t) \cdot e$$
$$= g \cdot (\varphi(t) \cdot e) = R_{\varphi(t)\cdot e} g,$$

for every $g \in \mathrm{U}(1)$.

Since the group U(1) acts smoothly to the right on P without fixed point [$R_g : (p, g) \in P \times \mathrm{U}(1) \to pg \in P$, $p(gg') = (pg)g'$, $\forall g, g' \in \mathrm{U}(1)$, $p \in P$, and $pg = p$ for some p implying $g = e$] the basis ζ_5 of U(1) can be mapped into the bundle, this means $R_{\varphi(t)}$ can be viewed as a one-parameter group of transformations acting on the bundle.

Call the vector field induced by $R_{\varphi(t)\cdot e}$ the fundamental vector field E_5 and write

$$E_5 = \partial_5 \equiv \frac{\partial}{\partial p^5}, \quad p \in P.$$

Because the bundle space P is locally trivial, hence diffeomorphic to the direct product space $U \times \mathrm{U}(1)$, one can choose the basis in $\pi^{-1}(U)$ as the local direct product basis. Write

$$\{\partial_\mu\} = \{(\partial_\mu, \partial_5)\}, \quad \mu = 1,...,4.$$

Here, $\mathcal{U}(1)$ denotes the Lie algebra of U(1). We introduce a connection[12] on P as a $\mathcal{U}(1)$-valued one-form $\overline{\omega}(v_p)$. Here, v_p belongs to the tangent space T_pP of P at each $p \in P$. One has

(i) $\overline{\omega}_p(\partial_s) = \zeta_5$;

(ii) $\overline{\omega}_p$ is differentiable at p;

(iii) $\overline{\omega}_{R_g p}(R_{g*} v_p) = \mathrm{Ad}(g^{-1})\overline{\omega}_p(v_p)$,

$$v_p \in T_pP, \quad \forall g \in U(1),$$

with the adjoint representation Ad(g).

In our direct product basis one needs the connection on $U \subset M$. Let $f: U \subset M \to f(U) \subset P$, $\pi \circ f = \mathrm{id}_M$ a local section. Define a $\mathcal{U}(1)$-valued one-form $\omega = f^* \overline{\omega}$ on U via pullback:

$$\omega \equiv (f^* \overline{\omega})_x(u) = \overline{\omega}_{f(x)}(f' u), \quad \forall u \in T_x M, \quad x \in M.$$

We are able to write ω using the basis ζ_5 of U(1):

$$\omega \equiv \omega^5 \zeta_5,$$

with $\overline{\omega}_p{}^5(\partial_5) = 1$ in the bundle space P. This immediately tells us that $\overline{\omega}^5$ is nothing more than the vector field in P dual to ∂_s. Note that the one-forms ω_i, ω_j in the trivializations Φ_{U_i}, Φ_{U_j} at a point $x \in U_i \cap U_j$ related to the same connection ω transform like

$$\omega_{i,x} = \omega_{j,x} - i g^{-1} \partial_\mu g \tag{1}$$

with respect to the transition function[12]

$$U_i \cap U_j \to U(1), \quad g = \overset{\Delta}{\Phi}_j(p)(\overset{\Delta}{\Phi}_i(p))^{-1}.$$

III. GEOMETRY OF THE BUNDLE MANIFOLD

Now we construct a metric $\mathfrak{g}$ on the bundle space P. The metric condition (7) will yield a flat connection with torsion on P. The components of the torsion tensor will be obtained. The components of the metric tensor on P satisfy

$$\mathfrak{g}_{\mu\nu} \equiv g(\pi' \partial_\mu, \pi' \partial_\nu) - \omega^5(\partial_\mu)\omega^5(\partial_\nu),$$
$$\mathfrak{g}_{\mu 5} \equiv g(\pi' \partial_\mu, \pi' \partial_5) - \omega^5(\partial_\mu)\overline{\omega}_5(\partial_5) = -\omega^5(\partial_\mu),$$
$$\mathfrak{g}_{5\nu} \equiv g(\pi' \partial_5, \pi' \partial_\nu) - \overline{\omega}^5(\partial_5)\omega^5(\partial_\nu) = -\omega^5(\partial_\nu),$$
$$\mathfrak{g}_{55} \equiv g(\pi' \partial_5, \pi' \partial_5) - \overline{\omega}^5(\partial_5)\overline{\omega}^5(\partial_5) = -1.$$

Note $(\pi' \partial_5) = 0$, $(\pi' \partial_\nu) = \partial_\nu$. By denoting $\omega^5(\partial_\mu) \equiv A_\mu$ this can be written as

$$(\mathfrak{g}_{\underline{\mu}\underline{\nu}}) = \left(\begin{array}{c|c} g_{\mu\nu} - A_\mu A_\nu & -A_\mu \\ \hline -A_\nu & -1 \end{array}\right) \tag{2}$$

or alternatively

$$(\mathfrak{g}^{\underline{\mu}\underline{\nu}}) = \left(\begin{array}{c|c} g^{\mu\nu} & -g^{\mu\nu}A_\nu \\ \hline -A_\mu g^{\mu\nu} & -1 + g^{\mu\nu}A_\mu A_\nu \end{array}\right), \tag{3}$$

where

$$\mathfrak{g}_{\mu\nu} g^{\nu\rho} = \delta^\rho_\mu, \quad \mu,\nu,... = 1,...,4,$$
$$\mathfrak{g}_{\underline{\mu}\underline{\nu}} \mathfrak{g}^{\underline{\nu}\underline{\rho}} = \delta^{\underline{\rho}}_{\underline{\mu}}, \quad \underline{\mu},\underline{\nu},... = 1,...,5, \tag{4}$$

and

$$\delta^\rho_\mu = \begin{cases} 1, & \text{if } \mu = \rho, \\ 0, & \text{if } \mu \neq \rho \end{cases}$$

as defined in Ref. 3.

Remarks: (i) The metric is independent of the fifth dimension, i.e., it is compatible with the bundle structure

$$\partial_5 \mathfrak{g}_{\underline{\mu}\underline{\nu}} = 0.$$

(ii) For elements defined on M raise and lower Greek indices with $g_{\mu\nu}$, for those on P do it with $\mathfrak{g}_{\underline{\mu}\underline{\nu}}$.

(iii) Greek and Latin indices $(\alpha,\beta,...),(a,b,...)$ run from 1 to 4, and underlined Greek and Latin indices $(\underline{\alpha},\underline{\beta},...),(\underline{a},\underline{b},...)$ run from 1 to 5.

(iv) Call $\omega^5(\partial_\mu) \equiv A_\mu$ the electromagnetic gauge potential, and call Eq. (1) the gauge transformation of this potential.

In extension of Ref. 3 we consider parallel vector fields on the bundle space P. We shall make the usual convention of summing over doubled indices:

$$\mathcal{A}^i = \mathcal{A}^i{}_{\underline{\mu}} \partial^{\underline{\mu}}, \quad \partial_{\underline{\mu}} = \mathcal{A}^i{}_{\underline{\mu}} \mathcal{A}_i, \quad \partial^{\underline{\mu}} = \mathcal{A}_i{}^{\underline{\mu}} \mathcal{A}^i. \tag{5}$$

The coefficients $\mathcal{A}_i{}^{\underline{\mu}}$ satisfy

$$\mathcal{A}_i{}^{\underline{\mu}} \mathcal{A}^i{}_{\underline{\nu}} = \delta^{\underline{\mu}}_{\underline{\nu}}, \quad \mathcal{A}_i{}^{\underline{\mu}} \mathcal{A}^i{}_{\underline{\mu}} = \delta^i_i,$$
$$\mathcal{A}^i{}_{\underline{\mu}} \eta_{ij} \mathcal{A}^j{}_{\underline{\nu}} = \mathfrak{g}_{\underline{\mu}\underline{\nu}}, \quad \mathcal{A}_i{}^{\underline{\mu}} \mathfrak{g}_{\underline{\mu}\underline{\nu}} \mathcal{A}_j{}^{\underline{\nu}} = \eta_{ij}. \tag{6}$$

Choose the Lorentz metric of the form

$$(\eta_{ij}) = \mathrm{diag}(-1, +1, +1, +1, -1).$$

Define the covariant derivative

$$D_{\underline{\nu}} \mathcal{A}_i{}^{\underline{\lambda}} \equiv \partial_{\underline{\nu}} \mathcal{A}_i{}^{\underline{\lambda}} + \gamma^{\underline{\lambda}}_{\underline{\mu}\underline{\nu}} \mathcal{A}_i{}^{\underline{\mu}}$$

with respect to the affine connection of absolute parallelism

$$\gamma_{\underline{\mu}\cdot\underline{\nu}}^{\underline{\lambda}} = \mathcal{A}_i{}^{\underline{\lambda}} \partial_{\underline{\nu}} \mathcal{A}^i{}_{\underline{\mu}},$$

followed by the metric condition

$$D_\rho g_{\sigma\tau} \equiv \partial_\rho g_{\sigma\tau} - \gamma_\sigma{}^\alpha{}_\rho g_{\alpha\tau} - \gamma_\tau{}^\alpha{}_\rho g_{\sigma\alpha} = 0. \tag{7}$$

This condition implies the length conservation of a vector under parallel transport. The torsion tensor is given by

$$\tau_{\underline{\mu}}{}^{\underline{\lambda}}{}_{\underline{\nu}} = \gamma_{\underline{\mu}}{}^{\underline{\lambda}}{}_{\underline{\nu}} - \gamma_{\underline{\nu}}{}^{\underline{\lambda}}{}_{\underline{\mu}} = b_{\underline{k}}{}^{\underline{\lambda}}(\partial_{\underline{\nu}} b^{\underline{k}}{}_{\underline{\mu}} - \partial_{\underline{\mu}} b^{\underline{k}}{}_{\underline{\nu}}). \tag{8}$$

In the local direct product basis we have

$$\begin{aligned} b_i &= b_i{}^{\underline{\mu}}\,\partial_{\underline{\mu}} = b_i{}^{\mu}\,\partial_\mu + b_i{}^5\,\partial_5, \\ b_5 &= b_5{}^{\underline{\mu}}\,\partial_{\underline{\mu}} = b_5{}^{\mu}\,\partial_\mu + b_5{}^5\,\partial_5, \end{aligned} \tag{9}$$

with $b_i{}^\mu = b_i{}^\mu$. Here, $b_i{}^\mu$ describes the gravitational field and is introduced in Ref. 3.

In view of

$$g_{\underline{\mu\nu}} = b^i_{\underline{\mu}}\,\eta_{ij}\,b^j_{\underline{\nu}},$$

with the components of the g metric given in (2), one obtains

$$\begin{aligned} &b^5{}_5 = 1, \quad b^5{}_\nu = A_\nu, \quad b^i{}_5 = 0, \quad b^i{}_\nu = b^i{}_\nu, \\ &b_5{}^5 = 1, \quad b_5{}^\nu = 0, \quad b_i{}^5 = -b_i{}^\nu A_\nu. \end{aligned} \tag{10}$$

This yields

$$\tau_{\underline{\mu}}{}^{\underline{\lambda}}{}_{\underline{\nu}} = T_\mu{}^\lambda{}_\nu = b_k{}^\lambda(\partial_\nu b^k{}_\mu - \partial_\mu b^k{}_\nu). \tag{11}$$

The coefficients of the linear connection

$$\gamma_{\underline{\mu}}{}^{\underline{\lambda}}{}_{\underline{\nu}} = b_i{}^{\underline{\lambda}}\,\partial_{\underline{\nu}} b^i_{\underline{\mu}}$$

are given by

$$\begin{aligned} &\gamma_5{}^5{}_5 = 0, \quad \gamma_\mu{}^\lambda{}_5 = 0, \quad \gamma_\mu{}^\lambda{}_\nu = \Gamma_\mu{}^\lambda{}_\nu, \\ &\gamma_5{}^5{}_\nu = 0, \quad \gamma_\mu{}^5{}_5 = 0, \quad \gamma_\mu{}^5{}_\nu = \partial_\nu A_\mu - A_\rho \Gamma_\mu{}^\rho{}_\nu. \\ &\gamma_5{}^\lambda{}_\nu = 0, \quad \gamma_5{}^\lambda{}_5 = 0, \end{aligned} \tag{12}$$

For those of the torsion tensor (8) we get

$$\begin{aligned} &\tau_5{}^5{}_5 = 0, \quad \tau_\lambda{}^\rho{}_5 = 0, \quad \tau_5{}^5{}_\nu = 0, \\ &\tau_5{}^\lambda{}_\nu = 0, \quad \tau_5{}^\lambda{}_5 = 0, \quad \tau_\mu{}^\lambda{}_\nu = T_\mu{}^\lambda{}_\nu, \\ &\tau_\lambda{}^5{}_5 = 0, \quad \tau_\mu{}^5{}_\nu = F_{\nu\mu} + A_\lambda T_\nu{}^\lambda{}_\mu. \end{aligned} \tag{13}$$

The connection coefficients $\Gamma_\mu{}^\lambda{}_\nu = b_i{}^\lambda\,\partial_\nu b^i{}_\mu$ and the components of the torsion tensor $T_\mu{}^\lambda{}_\nu = \Gamma_\mu{}^\lambda{}_\nu - \Gamma_\nu{}^\lambda{}_\mu$, are defined by Hayashi and Shirafuji[3] for the gravitational theory.

Remark: (v) We identify the term $F_{\nu\mu}$ in (13) with the electromagnetic field tensor

$$F_{\nu\mu} = \partial_\nu A_\mu - \partial_\mu A_\nu. \tag{14}$$

IV. FIELD EQUATIONS

A. Gravitational theory

So far we have considered the unifying geometry. Now let us make physics. As indicated above the quadruplet of linearly independent parallel vector fields $\{b_k{}^\mu\}$ is able to play the role of a g-metric tensor in a gravitational theory.

It is shown[3] that a theory described by these vector fields contains spin. If we make an applicaton of the Kaluza–Klein formalism to such a theory one may expect a description of gravitation, electromagnetism, and spin.

Hayashi and Shirafuji[3] constructed a gravitational Lagrangian density in vacuum,

$$L_G = \int d^4x\sqrt{-g}\,\mathcal{L}_G,$$

in view of the following postulates.

(1) Invariance under the group of general coordinate transformations; for an arbitrary change of coordinates the parallel vector fields transform like

$$b'_k{}^\mu(x') = (\partial_\nu x'^{\,\mu}) b_k{}^\nu(x).$$

(2) Invariance under the group of global, proper, orthochronous Lorentz transformations $L^\uparrow_+$; for its element $A = (A^j{}_k)$ with $A^t \eta A = \eta$, $\det A = 1$, $A^0{}_0 \geqslant 1$, and $\partial_\mu A = 0$, the parallel vector fields transform like

$$b_k{}^\mu(x) = A^j{}_k b'_j{}^\mu(x).$$

(3) $\mathcal{L}_G$ should be invariant under parity operation. Parity operation is a Lorentz transformation with $\underline{b}_{(0)} \to \underline{b}_{(0)}$, $\underline{b}_{(a)} \to -\underline{b}_{(a)}$, where the Lorentz indices are enclosed by parentheses.

(4) $\mathcal{L}_G$ should be quadratic in the torsion tensor, besides a cosmological term. They gave an irreducible decomposition of the torsion tensor $(T^\lambda{}_{\mu\nu})$ with respect to the group of global Lorentz transformations. By defining $T^\lambda{}_{\mu\nu} = b_k{}^\lambda(\partial_\nu b^k{}_\mu - {}_\mu b^k{}_\nu)$ they have

$$\begin{aligned} t_{\lambda\mu\nu} &= \tfrac{1}{2}(T_{\lambda\mu\nu} + T_{\mu\lambda\nu}) + \tfrac{1}{6}(g_{\nu\lambda} v_\lambda + g_{\nu\mu} v_\lambda) - \tfrac{1}{3} g_{\lambda\mu} v_\nu, \\ v_\mu &= T^\lambda{}_{\lambda\mu}, \\ a_\mu &= \tfrac{1}{6}\epsilon_{\mu\nu\rho\sigma} T^{\nu\rho\sigma}. \end{aligned} \tag{15}$$

Here, $\delta_{\mu\nu\rho\sigma}$ is a completely antisymmetric tensor density of weight $+1$ with normalization $\delta_{1234} = -1$. Define $\epsilon_{\mu\nu\rho\sigma} = \sqrt{-g}\,\delta_{\mu\nu\rho\sigma}$. The tensor $\{t_{\lambda\mu\nu}\}$ has the following properties:

$$\begin{aligned} &t_{\lambda\mu\nu} = t_{\mu\lambda\nu}, \\ &g^{\mu\nu} t_{\lambda\mu\nu} = 0 = g^{\lambda\mu} t_{\lambda\mu\nu}, \\ &t_{\lambda\mu\nu} + t_{\mu\nu\lambda} + t_{\nu\lambda\mu} = 0. \end{aligned} \tag{16}$$

The general Lagrangian density is required to be of the form

$$\mathcal{L}_G = a_1(t_{\lambda\mu\nu} t^{\lambda\mu\nu}) + a_2(v_\mu v^\mu) + a_3(a_\mu a^\mu) + a_0, \tag{17}$$

where a_1, a_2, a_3 are free parameters and a_0 is a cosmological term which we set equal to zero.

There is a "relationship" between the Riemann curvature scalar R in Riemann's space-time and a combination of the above-defined parts of the torsion tensor in Weizenböck's geometry:

$$\begin{aligned} \int d^4x\sqrt{-g}R = \int d^4x\sqrt{-g}\,[&-\tfrac{2}{3}(t_{\lambda\mu\nu} t^{\lambda\mu\nu}) \\ &+ \tfrac{2}{3}(v_\mu v^\mu) - \tfrac{2}{3}(a_\mu a^\mu)]. \end{aligned} \tag{18}$$

Define $\kappa = 8\pi G/c^4 := 8\pi G$, where G denotes Newton's gravitational constant. Hence the gravitational action has the following form:

$$\begin{aligned} L_G = \int d^4x\sqrt{-g}\,\Big(&\frac{1}{2\kappa}R + c_1(t_{\lambda\mu\nu} t^{\lambda\mu\nu}) \\ &+ c_2(v_\mu v^\mu) + c_3(a_\mu a^\mu)\Big). \end{aligned} \tag{19}$$

The constants are given by

$$c_1 = a_1 + 1/3\kappa, \quad c_2 = a_2 - 1/3\kappa, \quad c_3 = a_3 + 3/4\kappa.$$

It is known from considerations[14] of macroscopic systems that

$$\kappa c_1 = 0.001 \pm 0.001, \quad \kappa c_2 = -0.005 \pm 0.005. \tag{20}$$

Note that, if R describes the gravitational field, then $(\mathcal{L}_G - R/2\kappa)$ is the contribution of the spin interaction.

B. Kaluza–Klein approach

Follow Ref. 15, p. 260, and set $F_{\mu\nu}A_\lambda = 0$. Now we get, in a slightly different notation to Ref. 3,

$$\tau_\mu{}^\lambda{}_\nu = T_\mu{}^\lambda{}_\nu, \quad \tau_{\mu\lambda\nu} = T_{\mu\lambda\nu}, \quad \tau^{\mu\lambda\nu} = T^{\mu\lambda\nu}.$$

We define

$$f_{\underline{\mu}\underline{\lambda}\underline{\nu}} = \tfrac{1}{2}(\tau_{\underline{\mu}\underline{\lambda}\underline{\nu}} + \tau_{\underline{\lambda}\underline{\mu}\underline{\nu}}) + \tfrac{1}{6}(g_{\underline{\nu}\underline{\mu}}v_{\underline{\lambda}} + g_{\underline{\nu}\underline{\lambda}}v_{\underline{\mu}}) - \tfrac{1}{3}g_{\underline{\mu}\underline{\lambda}}v_{\underline{\nu}}, \tag{21}$$

$$v_{\underline{\mu}} = \tau_{\underline{\lambda}}{}^{\underline{\lambda}}{}_{\underline{\mu}}, \quad a_\mu = \tfrac{1}{6}\epsilon_{\mu\nu\rho\sigma}T^{\nu\rho\sigma}, \quad a_5 = 0.$$

One obtains

$$v_\mu = \tau_\beta{}^\beta{}_\mu, \quad v_5 = 0, \tag{22}$$

and further

$$\begin{aligned}
&f_{555} = 0, \quad f_{55\nu} = \tfrac{1}{3}v_\nu, \quad f_{5\lambda 5} = -\tfrac{1}{6}v_\lambda, \quad f_{\mu 55} = -\tfrac{1}{6}v_\mu,\\
&f_{\mu\lambda\nu} = \tfrac{1}{2}(T_{\mu\lambda\nu} + T_{\lambda\mu\nu}) + \tfrac{1}{6}(g_{\nu\mu}v_\lambda + g_{\nu\lambda}v_\mu) - \tfrac{1}{3}g_{\mu\lambda}v_\nu,\\
&f_{5\lambda\nu} = \tfrac{1}{2}F_{\lambda\nu} - \tfrac{1}{6}A_\nu v_\lambda + \tfrac{1}{3}A_\lambda v_\nu,\\
&f_{\mu 5\nu} = \tfrac{1}{2}F_{\mu\nu} - \tfrac{1}{6}A_\nu v_\mu + \tfrac{1}{3}A_\mu v_\nu,\\
&f_{\mu\lambda 5} = -\tfrac{1}{6}(A_\mu v_\lambda + A_\lambda v_\mu).
\end{aligned} \tag{23}$$

In extension of (17) we choose the Lagrangian in our Kaluza–Klein approach to be of the form

$$\mathcal{L}_{KK} = a_1(f_{\underline{\mu}\underline{\lambda}\underline{\nu}}f^{\underline{\mu}\underline{\lambda}\underline{\nu}}) + a_2(v_{\underline{\mu}}v^{\underline{\mu}}) + a_3(a_{\underline{\mu}}a^{\underline{\mu}}). \tag{24}$$

Note that

$$\begin{aligned}
&(f_{55\nu}f^{55\nu} + f_{5\lambda 5}f^{5\lambda 5} + f_{\mu 55}f^{\mu 55}) = \tfrac{1}{6}v_\mu v^\mu,\\
&(f_{5\lambda\nu}f^{5\lambda\nu} + f_{\mu 5\nu}f^{\mu 5\nu}) = 2f_{5\lambda\nu}f^{5\lambda\nu}.
\end{aligned}$$

Define

$$h_{\mu\lambda\nu} = \tfrac{1}{3}A_\mu A_\lambda v_\nu - \tfrac{1}{6}(A_\mu A_\nu v_\lambda + A_\nu A_\lambda v_\mu), \tag{25}$$

such that with respect to (15)

$$f_{\mu\lambda\nu}f^{\mu\lambda\nu} = t_{\mu\lambda\nu}t^{\mu\lambda\nu} + h_{\mu\lambda\nu}h^{\mu\lambda\nu} + 2t_{\mu\lambda\nu}h^{\mu\lambda\nu}.$$

Since we have related in (19) the contributions $T_\mu{}^\lambda{}_\nu, v_\mu, a_\mu$ outside of R to spin interactions, we can decompose the Lagrangian (24) into a sum of a gravitational part $\mathcal{L}'_G$ and an interacting part $\mathcal{L}'_{EG}$:

$$\mathcal{L}'_G = \mathcal{L}_G + (a_1/6)v_\mu v^\mu, \tag{26}$$

$$\begin{aligned}
\mathcal{L}'_{EG} = {}& a_1 h_{\mu\lambda\nu}h^{\mu\lambda\nu} + 2a_1 f_{5\lambda\nu}f^{5\lambda\nu}\\
&+ a_1 f_{\mu\lambda 5}f^{\mu\lambda 5} + 2a_1 t_{\mu\lambda\nu}h^{\mu\lambda\nu}.
\end{aligned}$$

In view of (18) we take

$$R = -\tfrac{2}{3}t_{\mu\lambda\nu}t^{\mu\lambda\nu} + \tfrac{2}{3}v_\mu v^\mu - \tfrac{2}{3}a_\mu a^\mu. \tag{27}$$

and write

$$\mathcal{L}'_G = (1/2\kappa)R + c_1 t_{\mu\lambda\nu}t^{\mu\lambda\nu} + (c_2 + a_1/6)v_\mu v^\mu + c_3 a_\mu a^\mu, \tag{28}$$

with the gravitational constant κ.

The considered Lagrangian in our Kaluza–Klein approach can now be written explicitly as

$$\begin{aligned}
\mathcal{L}_{KK} = {}& (1/2\kappa)R + c_1 t_{\mu\lambda\nu}t^{\mu\lambda\nu} + (c_2 + a_1/6)v_\mu v^\mu\\
&+ c_3 a_\mu a^\mu + a_1 h_{\mu\lambda\nu}h^{\mu\lambda\nu} + 2a_1 f_{5\lambda\nu}f^{5\lambda\nu}\\
&+ a_1 f_{\mu\lambda 5}f^{\mu\lambda 5} + 2a_1 t_{\mu\lambda\nu}h^{\mu\lambda\nu}.
\end{aligned} \tag{29}$$

To obtain the field equations apply a variational principle: We take the vector fields $b^k{}_\nu(x)$ as independent variables which vanish for $|x| \to \infty$:

$$\begin{aligned}
\delta L_{KK} &= \delta \int_{-\infty}^{\infty} \sqrt{-g}\,\mathcal{L}_{KK}\, d^4x\\
&= \int_{-\infty}^{\infty} \left\{ \frac{\partial\sqrt{-g}\,\mathcal{L}_{KK}}{\partial b^k{}_\nu}\delta b^k{}_\nu \right.\\
&\quad \left. + \frac{\partial\sqrt{-g}\,\mathcal{L}_{KK}}{\partial b^k{}_{\nu,\mu}}\delta b^k{}_{\nu,\mu} \right\} d^4x.
\end{aligned}$$

After an integration by parts of the divergence term and setting $\delta L_{KK} = 0$, one obtains the following Euler–Lagrange equation:

$$\frac{\delta\mathcal{L}_{KK}}{\delta b^k{}_\nu} = \frac{\partial\sqrt{-g}\,\mathcal{L}_{KK}}{\partial b^k{}_\nu} - \partial_\mu \frac{\partial\sqrt{-g}\,\mathcal{L}_{KK}}{\partial b^k{}_{\nu,\mu}} = 0. \tag{30}$$

Consider in the following

$$c_1 = 0 = c_2, \tag{31}$$

which is compatible with (20).

This implies

$$\begin{aligned}
\mathcal{L}_{KK} = {}& \{(1/2\kappa)R + c_3 a_\mu a^\mu\}\\
&+ a_1\{\tfrac{1}{6}v_\mu v^\mu + h_{\mu\lambda\nu}h^{\mu\lambda\nu} + 2f_{5\lambda\nu}f^{5\lambda\nu}\\
&+ f_{\mu\lambda 5}f^{\mu\lambda 5} + 2t_{\mu\lambda\nu}h^{\mu\lambda\nu}\}.
\end{aligned} \tag{32}$$

The variation of (32) yields two equations: One for the variation with respect to A_ν, one for the corresponding variation with respect to $b^k{}_\nu$.

When we take variations with respect to the electromagnetic components A_ν we achieve, by denoting $A^2 = g^{\mu\nu}A_\mu A_\nu$, etc.

$$\begin{aligned}
&\sqrt{-g}\{\tfrac{2}{3}A^\nu A^2 v^2 - \tfrac{1}{3}A^\nu(A_\mu v^\mu)^2 - \tfrac{1}{3}v^\nu A^2(a_\mu v^\mu)\\
&\quad + \tfrac{1}{3}v^\nu(A_\mu v^\mu) + A_\rho v_\lambda(T^{\nu\rho\lambda} + T^{\rho\nu\lambda})\}\\
&\quad + \tfrac{1}{2}[\sqrt{-g}\,F^{\nu\mu}]_{,\mu} = 0.
\end{aligned} \tag{33}$$

If the bracket $\{\cdots\}$ is set equal to zero we get the second Maxwell equation

$$[\sqrt{-g}F^{\mu\nu}]_{,\mu} = 0. \tag{34a}$$

The first Maxwell equation is the existence condition of A_μ:

$$F_{\alpha\beta,\gamma} + F_{\beta\gamma,\alpha} + F_{\gamma\alpha,\beta} = 0. \tag{34b}$$

Equation (33) shows the expected interaction between electromagnetism and spin. When one takes care of $b^k{}_\nu$ the variation provides

$$\begin{aligned}
&\frac{\delta\mathcal{L}_G}{\delta b^k{}_\nu} + \frac{1}{3\kappa}g^{\mu\nu}\\
&\quad \times [2F_{\lambda\rho}F^{\lambda\rho} + \tfrac{2}{3}v^2 + \tfrac{4}{3}A^2(A^2v^2 - (A_\rho v^\rho)^2)\\
&\quad - \tfrac{1}{3}(A_\rho v^\rho)^2 + 6A_\mu A_\lambda v_\alpha T^{\mu\lambda\alpha}]
\end{aligned}$$

$$+\frac{1}{3\kappa}\left\{\left[-\frac{1}{3}+\frac{1}{3}A^2A^2\right]_{,\mu}\right.$$
$$\times\sqrt{-g}\,[v^\mu g^{\mu\nu}-g^{\mu\mu}v^\nu]$$
$$+\frac{\sqrt{-g}}{3}\,[(A_\rho v^\rho)(1-A^2)]_{,\mu}[A^\mu g^{\mu\nu}-A^\nu g^{\mu\mu}]$$
$$+\left[-\frac{1}{3}+\frac{1}{3}A^2A^2\right]D_\mu\,[\sqrt{-g}(g^{\mu\nu}v^\mu-g^{\mu\mu}v^\nu)]$$
$$+\left[\frac{1}{3}(A_\rho v^\rho)(1-A^2)\right]D_\mu[\sqrt{-g}(g^{\mu\nu}A^\mu-g^{\mu\mu}A^\nu)]$$
$$+g^{\mu\beta}D_\mu\,[\sqrt{-g}(A_\beta A^\nu v^\mu-A_\beta A^\mu v^\nu)]$$
$$\left.+D_\mu\,[\sqrt{-g}A_\alpha A_\beta(g^{\mu\nu}T^{\alpha\beta\mu}-g^{\mu\mu}T^{\alpha\beta\nu})]\right\}. \tag{35}$$

The variation of the gravitational part $\delta\mathcal{L}_G$ is given in Ref. 3. In the considered matter-free case it provides

$$G^{\mu\nu}+L^{\mu\nu}=0, \tag{36a}$$

with Einstein's field tensor $G^{\mu\nu}=R^{\mu\nu}-\frac{1}{2}g^{\mu\nu}R$. By denoting $a_i=b_i{}^\mu a_\mu$, we can choose

$$L^{\mu\nu}=(2\kappa c_3/9)\{a_\lambda\,[\epsilon^{\mu\rho\sigma\lambda}(T^\nu_{\cdot\rho\sigma}-T_{\rho\sigma}{}^\nu)$$
$$+\epsilon^{\nu\rho\sigma\lambda}(T^\mu_{\cdot\rho\sigma}-T_{\rho\sigma}{}^\mu)]-3a^\mu a^\nu-\tfrac{3}{2}g^{\mu\nu}a^\rho a_\rho$$
$$+3\epsilon^{\mu\nu\rho\sigma}(b^i_{\ \rho}\,\partial_\sigma a^i+a_\rho v_\sigma)\}. \tag{36b}$$

If we neglect spin in (35), we obtain Einstein's field equation coupled to electromagnetism:

$$G^{\mu\nu}+(2/3\kappa)\,g^{\mu\nu}F_{\lambda\sigma}F^{\lambda\sigma}=0. \tag{37}$$

V. INTERFACE EFFECTS

The theory considered above predicts three types of interference effects: interference between (i) gravity and internal spin interactions, (ii) gravity and electromagnetism, and (iii) electromagnetism and internal spin interactions.

Note, that (33) tells us that the interference between electromagnetism and internal spin can produce an electromagnetic field.

The interference in (i) was discussed by several authors.[1,3] Because of the smallness of these effects they did not show a way to measure it.

To discuss the next stated effects let us remember the following. Because of the smallness of the gravitational interaction constant (the electromagnetic interaction constant is by a factor of 10^{36} larger than that of gravity) gravitation can be neglected in elementary particle physics. The importance of gravity is related to astronomical systems. But in these systems the electromagnetic interaction now can be neglected since these systems are almost neutral in large distances. This implies that there is no hope to measure the effects in (ii).

The intrinsic spin interaction gives a contribution to gravity which can be neglected in physical mass densities.[1] Thus there is no hope to detect the interference effects between electromagnetism and the spin sources.

VI. CONCLUSION

We considered a unified field theory of gravitation and electromagnetism containing intrinsic spin without the use of curvature. As stated above, the theory corresponds with ordinary Kaluza–Klein theory if we neglect some contributions of the torsion tensor related to spin.[15] This implies we get the Einstein equation $G_{\mu\nu}=0$ and the Maxwell equations

$$(\sqrt{-g}F^{\mu\nu})_{,\mu}=0,\quad F_{[\alpha\beta,\gamma]}=0.$$

There is no hope to measure the concerned interference effects.

[1] F. W. Hehl, P. von der Heyde, and G. D. Kerlick, Ref. Mod. Phys. **48**, 393 (1976).
[2] R. Weitzenböck, *Invariantentheorie* (Noordhoff, Groningen, 1923).
[3] K. Hayashi and T. Shirafuji, Phys. Rev. D **19**, 3524 (1979).
[4] J. W. Moffat, Phys. Rev. D **19**, 3554 (1979).
[5] J. W. Moffat, Ann. Inst. H. Poincaré **1981**, 85.
[6] J. W. Moffat, Acta Phys. Austriaca, Suppl. **23**, 665 (1981).
[7] M. W. Kalinowski, J. Math. Phys. **24**, 1835 (1983).
[8] Th. Kaluza, Sitzungsber. Preuss. Akad. Wiss. Berlin Math. Phys K **1**, 966 (1921).
[9] O. Klein, Z. Phys. **37**, 895 (1926).
[10] Y. M. Cho, J.Math. Phys. **16**, 2029 (1975).
[11] A. Trautmann, Rep. Math. Phys. **1**, 29 (1970).
[12] S. Kobayashi and K.Nomizu, *Foundations of Differential Geometry* (Interscience, New York, 1963), Vol. 1.
[13] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds and Physics* (North-Holland, Amsterdam, 1977).
[14] Reference 3, p. 3538.
[15] P. G.Bergmann, *Introduction of the Theory of Relativity* (Prentice-Hall, Englewood Cliffs, NJ, 1964).

Finite hard rod systems and their thermodynamic limit. II. Ergodic properties, pair distribution function, and equation of state

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

(Received 7 November 1984; accepted for publication 21 March 1985)

Generalizing a trick used previously for one- and two-particle systems we obtain closed expressions for the positions and velocities of the hard rods as functions of time and the initial conditions. Using the corresponding solution of Liouville's equation we show which ensembles can be used to calculate time averages and discuss the conditions under which an expectation value tends toward a stationary value. We also calculate the pair distribution function for almost all stationary distributions and the pressure, defined as the time average of the force exerted onto a specified particle by its left neighbor, for three kinds of stationary ensembles (smallest stationary, microcanonical, and canonical).

I. INTRODUCTION

This is the second of a series of papers on hard rod systems. In the preceding paper,[1] henceforth referred to as I, we presented our results and discussed their physical meaning. As stated in the Introduction of I the goal of this investigation was a twofold one: First, we wanted to discuss the equivalence of time and ensemble averages and to study the approach to equilibrium; second, we tried to illustrate the nature of the thermodynamic limit in that we calculated equilibrium and transport properties of finite systems and looked at how these quantities changed with increasing size of the system. Proofs of our assertions and derivations of the general formulas, which were evaluated numerically for various parameters, were outlined in I only to such an extent that the reader could recognize the underlying ideas.

This and the following paper contain a self-contained formal derivation of our results where emphasis is put onto mathematical rigor and completeness of the proofs. An exception is only Sec. VI of this paper where we repeat the arguments of I in more detailed and general form without trying to give a formal classification of all those cases where an expectation value tends toward a constant as time procedes.

The present paper is organized as follows: In Sec. II we establish our notation and define those concepts onto which our method of calculation is based. Section III deals with the evolution and with symmetry transformations. Conserved quantities, integrability, and the equivalence of time and ensemble averages are discussed in Sec. IV, followed by the definition of those stationary distributions which are used later on (Sec. V). As already mentioned the approach to equilibrium is discussed in Sec. VI. The pair distribution function is calculated in Sec. VII and the equation of state is derived in Sec. VIII.

The calculation of the other expectation values mentioned in I (collision frequency, no-collision probability, and velocity autocorrelation function) is the content of the following paper.

II. NOTATION AND BASIC DEFINITIONS

A. Parameters of the model

We will use the following parameters in this paper:

m = mass of the particle,
d = diameter of the particle,
N = number of particles,
L = length of the movable box ("volume"),
$L_d = L - Nd$ = effective volume,
$\rho = N/L$ = particle density,
$\rho_d = N/L_d$ = effective particle density,
E = total energy of the system,
$\epsilon = E/N$ = energy per particle,
T = temperature.

In the thermodynamic limit ($N \uparrow \infty, L \uparrow \infty, E \uparrow \infty, \rho < \infty, \epsilon < \infty$), ρ_d is defined by

$$\rho_d = \rho/(1 - \rho d). \tag{1}$$

B. Linear spaces

Elements of the real vector space R^N are denoted by capitals. We write

$$Y = (y_1,\ldots, y_N) = (y_i), \quad \alpha Y = \alpha(y_i) = (\alpha y_i), \tag{2}$$

and for the scalar product

$$Y \cdot W = \sum_i y_i w_i. \tag{3}$$

We frequently use a decomposition of R^N into two orthogonal subspaces $R'(\cong R^{N-1})$ and $R''(\cong R^1)$ defined by

$$Y = Y' + Y'', \quad Y'' = \left(\frac{1}{\sqrt{N}} \sum_j y_j\right). \tag{4}$$

We have $Y = Y'$ iff $\Sigma_i y_i = 0$. Note that the vectors Y' and Y'' both have N components although they belong to subspaces of dimensions $N-1$ and 1, respectively. It is sometimes desirable to use a parametrization of the vectors $Y \in R^N$ that is more adapted to the decomposition (4). This is achieved by passing from Y to $\overline{Y} = \mathscr{W} Y$, where $\mathscr{W}$ is an orthogonal transformation and

$$\bar{y}_N = \frac{1}{\sqrt{N}} \sum_j y_j. \tag{5}$$

Then $Y' \leftrightarrow (\bar{y}_1,\ldots,\bar{y}_{N-1})$ and $Y'' \leftrightarrow \bar{y}_N$.

The most important vectors are $X = (x_i)$, representing a configuration of the particles, and $V = (v_i)$, fixing their velocities. An element $(X,V) \in R^{2N}$ is called a state of the system and the set of all allowed states its phase space. The evolution is a mapping of the phase space onto itself. Under this map-

 0022-2488/85/102601-16$02.50

ping, a state (X,V) is transformed into (X_t,V_t). The dependence of $X_t=(x_{it})$ and $V_t=(v_{it})$ on X, V, and t is described in detail in Sec. III A.

C. The lattices $\{A\}$ and $\{B\}$

Lattices are sets of vectors which are closed under addition and subtraction. In most of the calculations there appears at least one of the two lattices $\{A\}$ or $\{B\}$. The lattice $\{A\}$ can be generated by the vectors

$$A_j = L_d(\delta_{ij} - 1/N), \quad j=1,...,N, \tag{6}$$

and $\{B\}$ is the lattice generated by the vectors

$$B_{jk} = A_j - A_k, \quad j,k=1,...,N. \tag{7}$$

We have $\{B\}\subset\{A\}\subset R'$ because of (7) and (4), where $\{A\}$ and $\{B\}$ are reciprocal lattices, i.e.,

$$A\cdot B = \text{integer multiple of } L_d^2. \tag{8}$$

Moreover, if $Y'\in R'$ and $A\cdot Y'=0$ modulo L_d^2 for all $A\in\{A\}$, then $Y'\in\{B\}$; likewise, if $Y'\in R'$ and $Y'\cdot B=0$ modulo L_d^2 for all $B\in\{B\}$, then $Y'\in\{A\}$. These facts may be proved by introducing suitable (reciprocal) bases of $\{A\}$ and $\{B\}$, e.g., $\{A_j\}$ and $\{B_{jN}\}, j=1,...,N-1$.

D. The orthogonal transformations $\mathscr{T}$ and $\mathscr{P}$

Orthogonal transformations of R^N are denoted by script capitals. The transformation $\mathscr{T}$ is the inversion defined by

$$\mathscr{T}Y = -Y. \tag{9}$$

We also use a set of orthogonal transformations $\mathscr{P}$, each of which is uniquely related to a permutation P: $i\to Pi$ of the numbers $1,...,N$. The permutations are labeled by r (or s) $=1,...,N!$ and we write $i(r)$ for $P_r\, i$. The indices $r=1,2,3$ are reserved for the permutations

$$\begin{aligned} P_1 &= (1)(2)\cdots(N), \quad i(1)=i,\\ P_2 &= (1,N)(2,N-1)\cdots, \quad i(2)=N+1-i,\\ P_3 &= (1,2,...,N), \quad i(3)=i+1-N\delta_{iN}. \end{aligned} \tag{10}$$

The transformation $\mathscr{P}$ related to the permutation P is defined by

$$\mathscr{P}Y = (Y_{P^{-1}i}), \tag{11}$$

where P^{-1}: $Pi\to i$ is the inverse of P. Definition (11) ensures that the transformation related to the permutation P_rP_s is $\mathscr{P}_r\,\mathscr{P}_s$. It follows from (11) and (9) that

$$\mathscr{T}\mathscr{P} = \mathscr{P}\mathscr{T}. \tag{12}$$

The lattice $\{A\}$ (and hence $\{B\}$) is invariant under the transformations $\mathscr{T}$ and $\mathscr{P}$:

$$\{\mathscr{T}A\} = \{\mathscr{P}A\} = \{A\}. \tag{13}$$

The definitions (4) and (11) entail

$$\mathscr{P}Y'' = Y''. \tag{14}$$

E. Simplices, cells, and other subsets of R^N

Subsets of R^N will also be denoted by capitals. It will be clear from the context whether a capital represents a single vector or a set of vectors. If $G\subset R^N$ and $Y\to\mathscr{W}Y+U$ is an orthogonal transformation of R^N, the image of G under this mapping is denoted by $\mathscr{W}G+U$,

$$\mathscr{W}G+U = \{Y\,|\,Y=\mathscr{W}Y_1+U,\ Y_1\in G\}. \tag{15}$$

The closure of G is denoted by $\overline{G}$ and

$$\text{volume of } G = |G| = |\overline{G}| = |\mathscr{W}G+U|, \quad \mathscr{W} \text{ orthogonal}, \quad U\in R^N. \tag{16}$$

The first subset of R^N which is of interest here is denoted by E and defined as follows:

$$X\in E \text{ iff } X\cdot B_{jk} = n_{jk}L_d^2, \quad n_{jk} \text{ integer, for } j,k=1,...,N. \tag{17}$$

We have that $R''\subset E$ and $R'\cap E = E'$ consists of $N(N-1)/2$ sets of parallel planes, the distance between two adjacent planes being $L_d/\sqrt{2}$ (see, e.g., Fig. 7 of I for $N=3$). Note that E is invariant under $\mathscr{T}$ and all $\mathscr{P}$'s.

The smallest subsets of R' enclosed by planes belonging to E' are called simplices. Each simplex is bounded by N planes, N edges,..., and N corners. Every simplex can be obtained from a fixed one by iterative reflections at the planes of E'. We choose as a basic simplex the set

$$S_1 + D_1 = \{X\,|\,X''=0;\ x_1<x_2<\cdots<x_N<x_1+L_d\}. \tag{18}$$

The reason for considering the set (18) as a simplex S_1 shifted by a vector $D_1\in R'$ will become clear in Sec. III A; for the moment S_1+D_1 may be treated as a single entity. The closure of (18) is

$$\overline{S}_1 + D_1 = \{X\,|\,X''=0;\ x_1\leqslant x_2\leqslant\cdots\leqslant x_N\leqslant x_1+L_d\}. \tag{19}$$

Each simplex possesses a certain regularity which manifests itself in a symmetry group isomorphic to the dihedral group of order $2N$. For S_1+D_1 one finds

$$S_1 + D_1 = \mathscr{P}_3(S_1+D_1) - A_1 = \mathscr{T}\mathscr{P}_2(S_1+D_1). \tag{20}$$

The transformation $X'\to\mathscr{P}_3X' - A_1$ generates the cyclic subgroup of order N while $X'\to\mathscr{T}\mathscr{P}_2X'$ is a transformation of order 2 needed to extend the cyclic group to a dihedral one. The simplices obtained from S_1+D_1 by application of the transformations $\mathscr{P}$ are just those that are obtained from S_1+D_1 by successive reflections at the planes $X\cdot B_{jk}=0$. They are denoted by

$$S_r + D_r = \mathscr{P}_r(S_1+D_1). \tag{21}$$

Let $\mathscr{P}^{(j)}$ correspond to a permutation $P^{(j)}$,

$$\mathscr{P}^{(j)}\leftrightarrow P^{(j)}, \quad P^{(j)}1 = j. \tag{22}$$

There are then $(N-1)!$ such transformations and they may be used to define new domains

$$\overline{Q}_j = \bigcup_r{}^{(j)}\mathscr{P}_r^{(j)}(\overline{S}_1+D_1), \tag{23}$$

which turn out to be parallelepipeds

$$\overline{Q}_j = \{X\,|\,X''=0;\ 0\leqslant X\cdot B_{kj}\leqslant L_d^2, \text{ for } k(\neq j)=1,...,N\}. \tag{24}$$

The equivalence of both definitions of $\overline{Q}_j$ may be verified by showing that each element of (23) belongs to (24) and vice versa.

In the same way that $(N-1)!$ simplices $\overline{S}_r+D_r$ may be combined into one parallelepiped $\overline{Q}_j$, the N domains $\overline{Q}_j$ may be combined into one single domain $\overline{C}$:

$$\bar{C} = \bigcup_r \mathscr{P}_r(\bar{S}_1 + D_1) = \bigcup_j \bar{Q}_j$$

$$= \{X \mid X'' = 0;\ X \cdot B_{jk} \leqslant L_d^2,\ \text{for } j,k = 1,\ldots,N\}. \tag{25}$$

The domain $\bar{C}$ is the Wigner–Seitz cell of the lattice B and reflects its symmetries, i.e.,

$$\mathscr{T}\bar{C} = \mathscr{P}\bar{C} = \bar{C}. \tag{26}$$

The volume of the symmetric cell is equal to the volume of the cell spanned by vectors $B_{jN}, j = 1,\ldots,N-1$, and is thus easily computed. Observing (16) and the definitions (23) and (25) one therefore finds

$$|C| = N|Q| = N!|S| = \sqrt{N} L_d^{N-1},$$
$$|Q| = |Q_j|, \quad |S| = |S_r|. \tag{27}$$

F. The functions $B[X']$ and $\mathscr{P}[X']$

Every simplex obtained from the partition of R' by planes of E' may be transformed into any other one by a series of reflections. This allows us to define an equivalence relation for the interior points of the simplices. To do this in a concise way we exploit the fact that two successive reflections at parallel planes are the same as a translation normal to the planes by a distance twice as large as the distance of the planes. The translation group originating this way is just given by the lattice $\{B\}$. Hence the first step in establishing the equivalence relation is to transfer the interior point of a simplex, say $X' \in R' \setminus E'$, into an interior point of $\bar{C}(\setminus E')$ by adding a suitable lattice vector B. The second step consists in the application of a transformation $\mathscr{P}$ that maps $X' + B$ into the reference simplex, say, $S_1 + D_1$. Note that both B and $\mathscr{P}$ are uniquely determined by X' and that $\mathscr{P}B \in \{B\}$ if $B \in \{B\}$. This means that for each $X' \in R' \setminus E'$ there exists a vector $B[X'] \in \{B\}$ and an orthogonal transformation $\mathscr{P}[X']$ such that

$$\mathscr{P}[X'](X' + B[X']) \in S_1 + D_1. \tag{28}$$

Both $\mathscr{P}[X']$ and $B[X']$ are uniquely determined by X' because, as may be shown by an indirect proof, $S_1 + D_1$ contains exactly one element of the set $\{Y' \mid Y' = \mathscr{P}_r X' + B;\ r = 1,\ldots,N!;\ B \in \{B\}\}$ if $X' \in R' \setminus E'$. If two vectors belong to the same simplex, then their related transformations $\mathscr{P}$ and B obviously coincide. That is, if $S[X']$ is the simplex containing X', i.e.,

$$X' \notin E':\quad S[X'] = \mathscr{P}^{-1}[X'](S_1 + D_1) - B[X'], \tag{29}$$

then $S[X_1'] = S[X_2']$ implies $\mathscr{P}[X_1'] = \mathscr{P}[X_2']$ and $B[X_1'] = B[X_2']$. The uniqueness of the functions $\mathscr{P}[\cdot]$ and $B[\cdot]$ implicitly defined by (28) and the definitions of Secs. II C and II D imply the following relations:

$$\mathscr{P}[\mathscr{P}_r X' + B] = \mathscr{P}[X']\mathscr{P}_r^{-1},$$
$$B[\mathscr{P}_r X' + B] = \mathscr{P}_r B[X'] - B; \tag{30}$$

$$\mathscr{P}[-X'] = \mathscr{P}_2\mathscr{P}[X'], \quad B[-X'] = -B[X']; \tag{31}$$

$$\mathscr{P}[X' - A_1] = \mathscr{P}_3\mathscr{P}[X'],$$
$$B[X' - A_1] = B[X'] + A_1 - \mathscr{P}^{-1}[X' - A_1]A_1. \tag{32}$$

In deriving (32) one has to take into account that $A_j - \mathscr{P}_r A_j \in \{B\}$.

G. Extended phase space functions

It turns out that the calculation of expectation values may be simplified by passing from phase space functions $f,g,h,\ldots$, defined on a subset of R^{2N} ("true phase space"), to functions $\bar{f},\bar{g},\bar{h},\ldots$, defined on the whole space R^{2N} ("extended phase space"). A function $\bar{h}$ is obtained from h according to the following scheme: Let χ^G be the characteristic function of the domain $G(\subset R')$, i.e.,

$$\chi^G(X') = \begin{cases} 1, & \text{for } X' \in G \\ 0, & \text{for } X' \notin G \end{cases}, \tag{33}$$

then a new function h^r is obtained from h by the definition

$$h^r(X + D_r, V) = \chi^{S_r}(X')h(X,V). \tag{34}$$

That is, h^r is the function h put equal to zero outside S_r and then shifted by D_r such that its support is now $S_r + D_r$. The vector D_r appearing in (34) is defined by

$$D_r = \mathscr{P}_r D_1, \quad D_1 = D_1' = d((N+1)/2 - i). \tag{35}$$

In the second step the "cutoff" function h^r is used to define the "extended" function $\bar{h}^r$:

$$\bar{h}^r(X,V) = \sum_{s,B} h^r(\mathscr{P}_s X + B, \mathscr{P}_s V). \tag{36}$$

These definitions imply

$$\bar{h}^r(X + D_r, V) = h(X,V), \quad \text{for } X' \in S_r \tag{37}$$

and

$$\bar{h}^r(\mathscr{P}X + B, \mathscr{P}V) = \bar{h}^r(X,V). \tag{38}$$

Equations (33)–(38) show that the correspondence between $\chi^{S_r}h$ and $\bar{h}^r$ is one-to-one. The expectation values we are interested in are of the form

$$\langle f,g\rangle_G = \int_G dX' \int dX'' \int dV f(X,V)g(X,V), \tag{39}$$

where $G = S_r$ and all unspecified integrations extend from $-\infty$ to $+\infty$. Now let G be a subset of R' generated by successive reflections from $S_1 + D_1$,

$$G = \bigcup_{s,B}^{(G)} [\mathscr{P}_s(S_1 + D_1) + B]; \tag{40}$$

then

$$\langle f,g\rangle_{S_r} = |G|^{-1}|S|\langle \bar{f}^r, \bar{g}^r\rangle_G. \tag{41}$$

Examples of domains of the form (40) are the parallelepipeds $\bar{Q}_j$, Eq. (23), and the cell $\bar{C}$, Eq. (25). Which choice of G is most advantageous depends on r and the definition of the functions f and g.

III. EVOLUTION IN TIME AND SYMMETRY TRANSFORMATIONS

A. The evolution in time

The set of configurations where the particles are ordered from the left to the right according to their numbers $i = 1,\ldots,N$ is $\{X \mid X' \in S_1\}$ where

$$S_1 = \Big\{ X \,|\, x_1 - d < x_2 - 2d < \cdots < x_N - Nd < x_1 - d + L_d;$$
$$\sum_j x_j = 0 \Big\}. \tag{42}$$

Note that $S_1 = (S_1 + D_1) - D_1$, in accordance with Eqs. (15), (18), and (35). Here, S_1 is of the same form and size as $S_1 + D_1$ but shifted against the origin $X = 0$ (all particles at $x = 0$) by the vector $-D_1$ (closest packing of the particles with center of mass at $x = 0$). Whether the configuration space, i.e., the set of allowed configurations, is $S_1 \times R''$ or $M \times R''$ with

$$M = \cup_r S_r, \quad S_r = \mathscr{P}_r S_1, \tag{43}$$

is a matter of convention. In the first case the particles are ordered in a certain way once and forever; in the second all orders of the particles are admitted. We choose the second possibility so that the phase space of the system is $M \times R'' \times R^N$. It decomposes into $N!$ disjoint parts since the $N!$ constituents S_r of M do not overlap (cf. I, Figs. 3 and 8).

Since the order of the particles is not changed by their interaction the state (X_t, V_t) at some instant t belongs always to the same part of the phase space as the initial state (X_0, V_0) from which it evolved. Using the functions $\mathscr{P}[\cdot]$ and $B[\cdot]$ defined in Sec. II F and assuming that $X'_0 \in S_1$ we have

$$X'_0 \in S_1: \quad Y_{t,1} = Y_{t,1}[X_0, V_0] = X_0 + V_0 t + D_1; \tag{44}$$

$$X_t = X_t[X_0, V_0] = \mathscr{P}[Y'_{t,1}](Y_{t,1} + B[Y'_{t,1}]) - D_1, \tag{45}$$

$$V_t = V_t[X_0 V_0] = \mathscr{P}[Y'_{t,1}] V_0.$$

Equations (44) and (45) are the mathematical formulation of the situation described in Sec. 4 of I (see especially Fig. 8). Here, $Y_{t,1}$ describes the uniform motion of a point in the extended configuration space R^N which is transformed into a billiardlike trajectory within the true configuration space $S_1 \times R'$. (In Fig. 8 of I, R'' points normal to the plane of drawing.) Note that the definition (44) is merely a convention introduced to allow for a more concise formulation of the dynamics, Eqs. (45).

We now prove the validity of Eq. (45) by induction: Assuming (45) to be valid for t we show that it is also for $t + \tau$, $\tau > 0$. Choosing τ sufficiently small we can assume without loss of generality that at most one single collision takes place within the time interval $(t, t + \tau)$. We therefore have to distinguish three types of events: (i) no collision, (ii) a binary collision (two particles involved), and (iii) a multiple collision (three or more particles involved).

(i) If no particles collide within the time interval $(t, t + \tau)$ the evolution is given by

$$X_{t+\tau} = X_t + V_t \tau, \tag{46}$$

$$V_{t+\tau} = V_t. \tag{47}$$

Expression (45) for X_t may be transformed into

$$\mathscr{P}^{-1}[Y'_{t,1}](X_t + D_1) - B[Y'_{t,1}] = Y_{t,1}. \tag{48}$$

From this and Eqs. (45) and (44), we get

$$\mathscr{P}^{-1}[Y_{t,1}](X_t + V_t \tau + D_1) - B[Y'_{t,1}]$$
$$= Y_{t,1} + V_0 \tau = Y_{t+\tau,1}. \tag{49}$$

Now $X_t \in S_1$ and $X_{t+\tau} \in S_1$, so that

$$\mathscr{P}^{-1}[Y'_{t,1}](S_1 + D_1) - B[Y'_{t,1}]$$
$$= S[Y'_{t,1}] = S[Y'_{t+\tau,1}]. \tag{50}$$

The last equation implies $\mathscr{P}[Y'_{t+\tau,1}] = \mathscr{P}[Y'_{t,1}]$ and $B[Y'_{t+\tau,1}] = B[Y'_{y,1}]$, which may be used to write (47) in the form $V_{t+\tau} = \mathscr{P}[Y_{t+\tau,1}] V_0$ and (49) in the form

$$\mathscr{P}^{-1}[Y'_{t+\tau,1}](X_{t+\tau} + D_1) - B[Y'_{t+\tau,1}] = Y_{t+\tau,1}. \tag{51}$$

(ii) For the case of binary collisions, we assume the collision to occur at time t. Then $X_t \in \partial S_1$, i.e.,

$$x_{j+1,t} - d = x_{j,t}, \quad \text{for some } j \in \{1, \ldots, N-1\} \tag{52}$$

or

$$x_{N,t} - Nd = x_{1,t} - d + L_d. \tag{53}$$

These equations may be cast into the form

$$\mathscr{P}_{(jk)}(X_t + D_1) + n_{jk} B_{jk} = X_t + D_1. \tag{54}$$

Here $\mathscr{P}_{(jk)}$ is the transformation corresponding to the transposition (jk); accordingly

$$\mathscr{P}_{(jk)} B_{jk} = -B_{jk} \tag{55}$$

and

$$\mathscr{P}_{(jk)} Y = Y \text{ iff } Y \cdot B_{jk} = 0 \ (y_j = y_k). \tag{56}$$

For (54) to be equivalent to (52) and (53) the integers n_{jk} have to be chosen in the following way:

$$j = 1, \ldots, N-1, \quad k = j+1: \quad n_{j,j+1} = 0; \tag{57}$$

$$j = N, \quad k = 1: \quad n_{N,1} = 1. \tag{58}$$

The transformation $(X + D_1) \to \mathscr{P}_{(jk)}(X + D_1) + n_{jk} B_{jk}$ is a reflection at the plane $(X + D_1) \cdot B_{jk} = n_{jk} L_d^2$. Equation (54) therefore expresses the fact that X_t lies on such a plane which forms part of the boundary of S_1.

The positions immediately before the collision are given by $X_{t-0} (= X_t)$, the velocities by V_{t-0}. For a collision between particles j and k to occur $v_{j,t-0} > v_{k,t-0}$ or, equivalently,

$$V_{t-0} \cdot B_{jk} > 0 \tag{59}$$

must hold true. During the collision, particles j and k exchange their velocities so that for sufficiently small (but finite) times after the collision,

$$X_{t+\tau} = X_{t-0} + \mathscr{P}_{(jk)} V_{t-0} \tau, \tag{60}$$

$$V_{t+\tau} = \mathscr{P}_{(jk)} V_{t-0}. \tag{61}$$

Because of $X'_{t-0} + D_1 \in S_1 + D_1$, (54), and (59),

$$X'_{t-0} + D_1 + V_{t-0} \tau \in \mathscr{P}_{(jk)}(S_1 + D_1) + n_{jk} B_{jk}, \tag{62}$$

where the domain on the right-hand side (rhs) of (62) is just the simplex obtained from $S_1 + D_1$ by reflection at the boundary plane containing X_t. Applying $\mathscr{P}^{-1}[Y'_{t-0,1}]$ to (62) and using (44) and (45) one obtains

$$Y'_{t+\tau,1} + B[Y'_{t-0,1}] - \mathscr{P}^{-1}[Y'_{t-0,1}] n_{jk} B_{jk}$$
$$\in \mathscr{P}^{-1}[Y'_{t-0,1}] \mathscr{P}_{(jk)}(S_1 + D_1). \tag{63}$$

From this and the uniqueness of the functions $\mathscr{P}[\cdot]$ and $B[\cdot]$, we get

$$\mathscr{P}[Y'_{t+\tau,1}] = \mathscr{P}_{(jk)} \mathscr{P}[Y'_{t-0,1}], \tag{64}$$

$$B[Y'_{t+\tau,1}] = B[Y'_{t-0,1}] - \mathscr{P}^{-1}[Y'_{t-0,1}]n_{jk}B_{jk}. \tag{65}$$

Equations (64) and (45) imply $V_{t+\tau} = \mathscr{P}[Y'_{t+\tau,1}]V_0$, and (65) combined with (55), (44), (45), and (60) yields

$$\begin{aligned}
&\mathscr{P}[Y'_{t+\tau,1}](Y_{t+\tau,1} + B[Y'_{t+\tau,1}]) \\
&\quad = \mathscr{P}_{(jk)}\mathscr{P}[Y'_{t-0,1}](Y_{t-0,1} + V_0\tau + B[Y'_{t-0,1}] \\
&\qquad - \mathscr{P}^{-1}[Y'_{t-0,1}]n_{jk}B_{jk}) \\
&\quad = \mathscr{P}_{(jk)}(X_{t-0} + D_1 - n_{jk}B_{jk} + V_{t-0}\tau) \\
&\quad = \mathscr{P}_{(jk)}(X_{t-0} + D_1) + n_{jk}B_{jk} + \mathscr{P}_{(jk)}V_{t-0}\tau \\
&\quad = X_{t-0} + \mathscr{P}_{(jk)}V_{t-0}\tau = X_{t+\tau}.
\end{aligned} \tag{66}$$

(iii) For the case of multiple collisions, note that if only two particles hit each other they exchange their velocities since the energy and the total momentum of the two colliding particles are assumed to be conserved during the interaction. These conservation laws do not suffice to fix what happens after the collision if three or more particles collide at the same instant. To define the dynamics also in these cases we consider multiple collisions as limiting cases of sequences of binary collisions. Since formula (45) is correct for a binary collision it also holds true for sequences of such collisions and therefore, by definition, also for multiple collisions. It should be noted that this definition is a very natural one because the limit X_t of a sequence $X_t^{(n)}$ depends only on X_0, the limit of $X_0^{(n)}$, and not on the particular sequence of orbits that has been chosen. This stability of the evolution against small perturbations of the initial conditions, easily visualized for $N = 3$ (cf. I, Figs. 7 and 8), is an essential feature of the present model. In its formulation as a billiard problem in N dimensions it may be traced back to the fact the angles between the confining planes are either $\pi/3$ or $\pi/2$.

If the particles are initially not ordered according to their numbers, $X'_0 \notin S_1$ and Eqs. (44) and (45) are not applicable. The evolution is instead given by the following formulas, specified for a certain r:

$$X'_0 \in S_r: \quad Y_{t,r} = X_0 + V_0 t + D_r; \tag{67}$$

$$\begin{aligned}
X_t &= X_t[X_0, V_0] \\
&= \mathscr{P}_r\mathscr{P}[Y'_{t,r}](Y_{t,r} + B[Y'_{t,r}]) - D_r, \\
V_t &= V_t[X_0, V_0] = \mathscr{P}_r\mathscr{P}[Y'_{t,r}]V_0.
\end{aligned} \tag{68}$$

Equations (68) may be proved in a similar way as Eqs. (45).

The evolution in point mechanics determines uniquely the evolution in statistical mechanics. For if f_t is the distribution function representing our knowledge of the state of the system at time t, this function is related to its initial value f_0 by

$$f_t(X,V) = f_0(X_{-t}[X,V], V_{-t}[X,V]). \tag{69}$$

Since the explicit form of the functions X_{-t} and V_{-t} depends on the simplex to which the argument X belongs it is convenient to decompose f_t as

$$f_t(X,V) = \sum_r \chi^{S_r}(X')f_t(X,V). \tag{70}$$

Each function appearing on the rhs may be extended to a function $\bar{f}_t^r$ defined on R^{2N} according to the scheme described in Sec. II G:

$$\bar{f}_t^r(X,V) = \bar{f}_t^r(\mathscr{P}X + B, \mathscr{P}V) = f_t(X - D_r, V) \quad \text{for } X' \in S_r. \tag{71}$$

The advantage of the functions $\bar{f}_t^r$ consists in their simple time dependence. For instance,

$$\begin{aligned}
X' \in S_1: \quad \bar{f}_t^1(X + D_1, V) &= f_t(X,V) = f_0(X_{-t}[X,V], V_{-t}[X,V]) \\
&= f_0(\mathscr{P}[Y'_{-t,1}](Y_{-t,1} \\
&\quad + B[Y'_{-t,1}]) - D_1, \mathscr{P}[Y'_{-t,1}]V) \\
&= \bar{f}_0^1(\mathscr{P}[Y'_{-t,1}](Y_{-t,1} \\
&\quad + B[Y'_{-t,1}]), \mathscr{P}[Y'_{-t,1}]V) \\
&= \bar{f}_0^1(Y_{-t,1}, V) = \bar{f}_0^1(X - Vt + D_1, V),
\end{aligned} \tag{72}$$

and the combination of (72) and (71) yields

$$\bar{f}_t^1(X,V) = \bar{f}_0^1(X - Vt, V), \tag{73}$$

for all $X \in R^N$. More generally,

$$\bar{f}_t^r(X,V) = \bar{f}_0^r(X - Vt, X), \tag{74}$$

for all X and arbitrary r. The price to be paid for the simple time dependence (74) is that the equivalent of one single function f_t is now a set of $N!$ functions $\bar{f}_t^r$. However, if the order of particles is definitely known, all but one of these functions vanish identically.

B. Symmetry transformations

A symmetry transformation $X \to \mathscr{W}X + U$, $V \to \mathscr{W}V$ is, by definition, a (linear canonical) transformation commuting with the evolution. Some of these symmetries, namely those of a general closed N-particle system, are expected to exist while others are less obvious and characteristic for the present model.

(i) Rigid translations: A rigid translation is given by $x_i \to x_i + u$, $v_i \to v_i$ or $X \to X + U''$, $V \to V$. Because of (14),

$$\begin{aligned}
X_t[X_0 + U'', V_0] &= X_t[X_0, V_0] + U'', \\
V_t[X_0 + U'', V_0] &= V_t[X_0, V_0].
\end{aligned} \tag{75}$$

(ii) Rigid inversion: This is the transformation $x_i \to -x_i$, $v_i \to -v_i$ or $X \to \mathscr{T}X = -X$, $V \to \mathscr{T}V = -V$. Taking into account $-X' \in \mathscr{P}_2 S_1 = S_2$ if $X' \in S_1$, $-D_1 = \mathscr{P}_2 D_1 = D_2$, and Eqs. (31), one finds

$$\begin{aligned}
X_t[-X_0, -V_0] &= -X_t[X_0, V_0], \\
V_t[-X_0, -V_0] &= -V_{-t}[X_0, V_0].
\end{aligned} \tag{76}$$

(iii) Permutation of particles: A typical transformation is $x_i \to x_{i(r)}$, $v_i \to v_{i(r)}$ or $X \to \mathscr{P}_r^{-1}X$, $V \to \mathscr{P}_r^{-1}V$. Equations (30) may be used to show that

$$\begin{aligned}
X_t[\mathscr{P}_r X_0, \mathscr{P}_r V_0] &= \mathscr{P}_r X_t[X_0, V_0], \\
V_t[\mathscr{P}_r X_0, \mathscr{P}_r V_0] &= \mathscr{P}_r V_t[X_0, V_0].
\end{aligned} \tag{77}$$

(iv) Cyclic transformations: A typical transformation is

$$X \to \begin{cases} \mathscr{P}_3 X - (L/L_d)A_1, & \text{if } X' \in S_1, \\ X, & \text{if } X' \notin S_1, \end{cases} \qquad V \to \mathscr{P}_3 V. \tag{78}$$

That is, if $X \in S_1$, x_i is substituted by $x_{i-1+N\delta_{i1}}$

$-L(\delta_{i1}-N^{-1})$. It follows from definitions (6), (7), (10), and (35) that

$$\mathscr{P}_3X_0-(L/L_d)A_1+D_1+\mathscr{P}_3V_0t$$
$$=\mathscr{P}_3(X_0+V_0t+D_1-A_1)+B_{21}. \tag{79}$$

Note that $\mathscr{P}_3X_0-(L/L_d)A_1\in S_1$ if, and only if, $X_0\in S_1$ [cf. Eqs. (20), (6), (10), (11), and (35)]. Therefore, if this configuration and the velocity $\mathscr{P}_3V_0$ are inserted as initial values in (45) one obtains, using once more Eqs. (30),

$$X_t[\mathscr{P}_3X_0-(L/L_d)A_1,\mathscr{P}_3V_0]+D_1$$
$$=\mathscr{P}[Y'_{t,1}-A_1](Y_{t,1}-A_1+B[Y'_{t,1}-A_1]), \tag{80}$$

$$V_t[\mathscr{P}_3X_0-(L/L_d)A_1,\mathscr{P}_3V_0]=\mathscr{P}[Y'_{t,1}-A_1]V_0.$$

By means of (32) we finally arrive at the following result:

$$\begin{aligned} X'_0\in S_1:\quad & X_t[\mathscr{P}_3X_0-(L/L_d)A_1,\mathscr{P}_3V_0] \\ & =\mathscr{P}_3X_t[X_0,V_0]-(L/L_d)A_1, \\ & V_t[\mathscr{P}_3X_0-(L/L_d)A_1,\mathscr{P}_3V_0] \\ & =\mathscr{P}_3V_t[X_0,V_0]. \end{aligned} \tag{81}$$

The symmetry transformation (78) maps $S_1\times R''\times R^N$ onto itself leaving the rest of the phase space unaffected. It generates a cyclic group of order N and, combining (81) with (77), isomorphic cyclic symmetry groups may be found for each of the $N!$ parts of the phase space. The cyclic symmetry group of a constituent may be extended to a dihedral group by adding a suitable combination of (76) and (77). For example, the transformation

$$X\to-\mathscr{P}_2X,\quad V\to-\mathscr{P}_2V \tag{82}$$

transforms the parts $S_r\times R''\times R^N$ of the phase space in such a way that $S_1\times R''\times R^N$ is mapped onto itself.

The existence of cyclic symmetry groups not contained in the general symmetry transformations of a closed N-particle system (rigid motions, permutations) is a special feature but not reserved to the present model. As approximate symmetries such groups also appear in the lattice dynamics of solids where they are referred to as "translation symmetry of the crystal."[2] The common origin of these groups is always an interaction which preserves the initial order of the particles (at least classically and within a certain range of energy).

IV. INTEGRABILITY AND TIME AVERAGES

A. Conserved quantities

There exist three types of conserved quantities that are easily found by simple inspection of the system's evolution but may also be derived from formulas (67) and (68).

(i) Total momentum: Equations (67), (68), (4), (6), (7), (14), and (35) imply that the center of mass undergoes a free motion:

$$X''_t=X''_0+V''_0t,\quad V''_t=V''_0. \tag{83}$$

By Noether's theorem, (83) may also be derived from (75). Since the motion of the center of mass is trivial and does not interfere with the complicated relative motion of the particles we exclude it in the following by imposing the constraints

$$X''_0=0,\quad V''_0=0 \tag{84}$$

on all initial conditions. For statistical mechanics this means that we consider only distribution functions of the form

$$f(X,V)=f'(X',V')\delta(X'')\delta(V''). \tag{85}$$

(ii) Order of the particles: As mentioned several times before, the initial order of the particles is conserved in time. This also follows formally from the first of Eqs. (68): $X'_t\in S_r$, if $X'_0\in S_r$.

(iii) Set of initial velocities—invariants: The second of Eqs. (68) shows that the set of particle velocities v_i is the same for almost all times, i.e., $\{v_{it}\}=\{v_{i0}\}$ whenever $V_t[X_0,V_0]$ is defined. This set can be uniquely characterized by the values of N symmetric functions, e.g.,

$$I_n(V)=\sum_i v_i^n,\quad n=1,...,N. \tag{86}$$

Considered as phase space functions,

$$F_n(X_0,V_0)=I_n(V_0)=I_n(V_t\,X_0,V_0)=F_n(X_t,V_t), \tag{87}$$

if $X'_t\notin\partial\overline{S}_r$ and hence V_t is defined. Without this restriction we were faced with N analytic invariant functions suggesting that the system is integrable.[3,4] It should, however, be kept in mind that $F_n(X_t,V_t)$ is only defined for those time intervals where all the particles move freely; ignoring this fact would change the dynamics of the present model into that of a system of noninteracting particles. The singular nature of the hard core interaction excludes, in fact, the possibility to identify $(m/2)F_2$ with the Hamiltonian and to describe the dynamics in terms of Hamiltonian dynamics.

B. Action and angle variables

For the above-mentioned reasons it is not possible to use the scheme offered in literature[3,4] to derive action and angle variables from the invariants. But this does not imply that such variables do not exist for hard rod systems. Assume, for instance, that $X'_t\in S_1$; the action variables are then

$$V_0=\mathscr{P}[\tilde{Y}_{t,1}]V_t \tag{88}$$

and the angle variables can be defined as

$$Y_{t,1}=\mathscr{P}[\tilde{Y}_{t,1}](X_t+D_1+B[\tilde{Y}_{t,1}]), \tag{89}$$

where

$$\tilde{Y}_{t,1}=X_t-V_t\,t+D_1. \tag{90}$$

Equations (88)–(90) express that (X_0,V_0) is obtained from (X_t,V_t) by going back from time t to time 0. If the definition (90) is inserted into Eq. (89) and the relation (88) is used, one obtains $Y_{t,1}$ in the form (44) or, equivalently, as

$$Y_{t,1}=Y_{0,1}+V_0\,t, \tag{91}$$

which is the solution of the equation of motion in terms of action and angle variables. Analogous solutions exist if $X_t\in S_r$, $r\neq1$.

At first sight the mappings $(X_t,V_t)\to(Y_{t,1},V_0)$, defined by Eqs. (88)–(90), and $(Y_{t,1},V_0)\to(X_t,V_t)$, given by Eqs. (45), look like ordinary linear canonical transformations. There are, however, two objections against such a simple view. The first is that the objects mapped onto each other belong to different spaces, namely $(X_t,V_t)\in S_1\times R''\times R^N$ (phase space) and $(Y_{t,1},V_0)\in R^{2N}$ (extended phase space). Equations (30) imply that under the mapping (45) both $(Y_{t,1},V_0)$ and $(Y_{t,1}+B,V_0)$

are mapped onto the same (X_t,V_t) irrespectively of what $B\in\{B\}$ has been chosen. To obtain a one-to-one correspondence Eqs. (44) and (89)–(91) containing $Y_{t,1}$ or $\tilde{Y}_{t,1}$ have to be read as relations modulo $\{B\}$ and the element $(Y_{t,1},V_0)$ has to be considered as representative of the lattice $\cup_B(Y_{t,1}+B,V_0)$ (cf. I, Fig. 4). The second point is that the mappings (45) and (88)–(90) are undefined for $Y_t\in E$ and $\tilde{Y}_t \in E$, respectively, where E is the set defined in (17). The mappings of the two phase spaces onto each other, i.e., the transition from position and momentum variables to action and angle variables or vice versa, are therefore defined only up to a set of hyperplanes, the definition of which depends on the considered instant. If this set of measure zero is removed the remaining part of the phase space decomposes into disjoint subsets that are mapped onto similar subsets of the other phase space, each pair of subsets being related by its own linear transformation.

That the angle variables are multivalued functions of the position and momentum variables is well known from Hamiltonian mechanics. There also exist some familiar examples (e.g., the pendulum) where different action and angle variables have to be constructed for various parts of the phase space. What makes the hard rod systems considered here so peculiar are only the following two points: (i) the partition of the phase space varies with the time; and (ii) the canonical transformations relating the parts of the phase space to the action and angle variables are always linear transformations.

C. Time averages

The theory of integrable systems[3,4] tells us that a time average of a phase space function coincides with the average value of this function for that subset of the phase space which is the closure of all states evolving from the given initial state. Each of these subsets is invariant under the evolution and topologically equivalent to a torus. These invariant tori, which densely fill the whole phase space, fall into two categories: irrational (or nonresonant) tori corresponding to general (or likely) initial conditions and rational (or resonant) tori corresponding to special (or unlikely) initial conditions. In the first case the torus is uniquely determined by the values of the action variables which are linearly independent over the integers (and therefore also over the rationals). In the second case there exist one or more linear relations with integer coefficients between the values of the action variables so that they do not suffice to characterize the torus uniquely. Rational tori are dense in phase space but of measure zero.

Since we already solved the equations of motion in terms of action and angle variables it would be sufficient to list all the rational and irrational tori in order to see which ensemble averages can be used as substitutes for time averages. However, the concepts introduced in Sec. II G are easily combined with the evolution (Sec. III A) to derive these results straightforwardly without any recourse to the general theory. This is done in the following since there are slight modifications of the general scheme originating from the constraint (84). Moreover, since the general theory is rather abstract and not well known outside the field of nonlinear dynamics it may be of interest to illustrate it with the example at hand. In this discussion and the following we stick to the terminology of paper I ("smallest stationary ensemble" instead of "invariant torus"), which was introduced before we became aware that our results on time averages are best explained in terms of modern mechanics.

The time average of a phase space function g for a sequence of states beginning with $X_0=X_0'$, $V_0=V_0'$ is defined by

$$\langle g\rangle_{X_0',V_0',t}=\lim_{\tau\uparrow\infty}\frac{1}{\tau}\int_0^\tau dt\, g(X_t[X_0',V_0'],V_t[X_0',V_0']). \tag{92}$$

The integrand in (92) is a periodic function if, and only if, $V_0't\in\{B\}$ for some $t>0$. Though this condition will not be satisfied in general $V_0't$ can be approximated to any desired degree of accuracy if t is chosen to be sufficiently large. The integrand in (92) is therefore in general an almost periodic function (this even in the sense of the precise mathematical definition[5]). We now assume $X_0=X_0'\in S_1$, which implies $X_t=X_t'\in S_1$. The integrand in (92) may therefore be replaced by

$$\begin{aligned} &g^1(X_t[X_0',V_0']+D_1,V_t[X_0',V_0'])\\ &\quad=\bar{g}^1(X_t[X_0',V_0']+D_1,V_t[X_0',V_0'])\\ &\quad=\bar{g}^1(X_0'+V_0't+D_1,V_0') \end{aligned} \tag{93}$$

[cf. Eqs. (33)–(38), (44), (45), (72), and (73)]. To evaluate

$$\langle g\rangle_{X_0',V_0',t}=\lim_{\tau\uparrow\infty}\frac{1}{\tau}\int_0^\tau dt\,\bar{g}^1(X_0'+V_0't+D_1,V_0'), \tag{94}$$

we note that $\bar{g}^1$ is periodic in X', the periods being given by the vectors $B\in\{B\}$. It is therefore natural to expand $\bar{g}^1$ as a Fourier series

$$\bar{g}^1(X',V')=\sum_A G_A^1(V')\exp[i\alpha A\cdot X'], \tag{95}$$

$$\alpha=2\pi/L_d^2. \tag{96}$$

For the Fourier coefficients

$$G_A^1(V')=|C|^{-1}\int_C dX'\,\bar{g}^1(X',V')\exp[-i\alpha A\cdot X'] \tag{97}$$

to exist, the function $\bar{g}^1$ (and hence g) must belong to a certain class of (generalized) functions. A sufficiently general condition is, for instance,

$$\sum_\tau\int_{S_\tau}dX'\,|g(X',V')|<\infty. \tag{98}$$

Inserting (95) into (94) one arrives at

$$\begin{aligned} \langle g\rangle_{X_0',V_0',t}&=\sum_A G_A^1(V_0')\exp[i\alpha A\cdot(X_0'+D_1)]\\ &\quad\times\lim_{\tau\uparrow\infty}\tau^{-1}I(A\cdot V_0',\tau), \end{aligned} \tag{99}$$

$$I(A\cdot V_0',\tau)=\int_0^\tau dt\exp[i\alpha A\cdot V_0't].$$

Now

$$\begin{aligned} &I(A\cdot V_0',\tau)=\tau, \quad \text{for } A\cdot V_0'=0,\\ &|I(A\cdot V_0',\tau)|\leq 2|\alpha A\cdot V_0'|^{-1}, \quad \text{for } A\cdot V_0'\neq 0, \end{aligned} \tag{100}$$

so that

$$\langle g\rangle_{X_0',V_0',t}=\sum_A^{(V_0')}G_A^1(V_0')\exp[i\alpha A\cdot(X_0'+D_1)], \tag{101}$$

$$\sum_A^{(V_0')}=\sum \text{ over all } A\in\{A\} \text{ for which } A\cdot V_0'=0. \tag{102}$$

(102)

Using the definition (97) of G_A^1, the partition of the cell C into simplices S_r+D_r, and the symmetry properties (38) of $\bar{g}^1$, Eq. (101) may be transcribed as follows:

$$\begin{aligned}\langle g\rangle_{X_0',V_0',t}&=|C|^{-1}\sum_r\int_{S_r+D_r}dX'\,\bar{g}^1(X',V_0')\\&\quad\times\sum_A^{(V_0')}\exp[i\alpha A\cdot(X_0'+D_1-X')]\\&=|C|^{-1}\sum_r\int_{S_1+D_1}dY'\,\bar{g}^1(\mathscr{P}_rY',V_0')\\&\quad\times\sum_A^{(V_0')}\exp[i\alpha A\cdot(X_0'+D_1-\mathscr{P}_rY')]\\&=|C|^{-1}\sum_r\int_{S_1+D_1}dY'\,\bar{g}^1(Y',\mathscr{P}_r^{-1}V_0')\\&\quad\times\sum_A^{(\mathscr{P}_r^{-1}V_0')}\exp[i\alpha A\\&\quad\cdot(\mathscr{P}_r^{-1}X_0'+\mathscr{P}_r^{-1}D_1-Y')]\\&=\int_{S_1}dX'\int dV'\,|C|^{-1}\sum_r\sum_A^{(\mathscr{P}_rV_0')}\exp[i\alpha A\\&\quad\cdot(\mathscr{P}_rX_0'+D_r-X'-D_1)]\\&\quad\times\delta(V'-\mathscr{P}_rV_0')g(X',V').\end{aligned} \tag{103}$$

In the second step we used $U\cdot\mathscr{P}V=\mathscr{P}^{-1}U\cdot V$ and the equivalence of $A\cdot V_0'=0$ and $\mathscr{P}A\cdot\mathscr{P}V_0'=0$. It follows from (103) that the time average is equal to the average over a stationary ensemble, namely

$$\langle g\rangle_{X_0',V_0',t}=\int dX'\int dV'\,f^I_{X_0',V_0'}(X',V')g(X',V'), \tag{104}$$

with

$$\begin{aligned}f^I_{X_0',V_0'}(X',V')&=|S|^{-1}\chi^{\bar{S}_1}(X')\frac{1}{N!}\sum_r\sum_A^{(\mathscr{P}_rV_0')}\exp[i\alpha A\\&\quad\cdot(\mathscr{P}_rX_0'+D_r-X'-D_1)]\\&\quad\times\delta(V'-\mathscr{P}_rV_0').\end{aligned} \tag{105}$$

The function (105) is real since $(-A)\cdot V_0'=0$ if $A\cdot V_0'=0$, and reversing the transformations of (103) one easily verifies that

$$\int dX'\int dV'\,f^I_{X_0',V_0'}(X',V')=1. \tag{106}$$

Moreover if $X'_{-t}[X',V'],V'_{-t}[X',V']$ is substituted for X', V' in (105) one finds

$$f^I_{X_0',V_0'}(X'_{-t}[X',V'],V'_{-t}[X',V'])=f^I_{X_0',V_0'}(X',V'). \tag{107}$$

On account of Eqs. (106) and (107) and the reality of $f^I_{X_0',V_0'}$, one might be tempted to consider the function (105) [supplemented with $\delta(X'')\delta(V'')$] as a stationary distribution function but before doing so one has to prove that this distribution is positive.

In proving the positivity of the distribution (105) we have to distinguish whether the sum over A contains one or more terms. This leads us to introduce a set $Q\subset R'$ defined as follows:

$$\begin{aligned}&V'\in Q \text{ iff } V'\cdot A=0, \quad \text{for some } A\neq 0,\\&\text{or, equivalently, iff } \sum_i v_i=0 \text{ and } \sum_i m_iv_i=0,\\&\text{for some set of integers } m_i \text{ with } m_i\neq m_j,\\&\text{for some pair } i\neq j.\end{aligned} \tag{108}$$

For $N>2$ this set has two remarkable properties:

(i) Q is of measure zero (with respect to R');

(ii) Q is dense in R'. (109)

The proof of the first assertion follows from the fact that for every $V'\in Q$ there exists a sublattice of $\{A\}$ consisting of all vectors A for which $A\cdot V'=0$. If D is the dimension of this lattice, then $1\leqslant D\leqslant N-2$. Thus, V' lies in the orthogonal complement of the subspace spanned by the sublattice, i.e., in a subspace of R' of dimension $D^c=N-1-D$. Accordingly the vectors $V'\in Q$ can be classified according to their corresponding sublattices. The set of all sublattices is countable, since each of them is uniquely determined by a basis $A_1,\ldots,A_D$. The set Q is therefore a countable union of subspaces of dimension D^c, $1\leqslant D^c\leqslant N-2$, which implies that Q is of measure zero with respect to R'. The proof of proposition (ii) of (109) rests on the properties of the rationals and the lattice $\{B\}$. Every $V'\in R'$ may be represented as real linear combination of at most $N-1$ vectors $B_j\in\{B\}$. Since real numbers may be approximated to any desired degree of accuracy by rationals, every $V'\in R'$ may be approximated by a linear combination of B's with rational coefficients. This means that for every $V'\in R'$ and $\epsilon>0$ there exists a rational number β and a vector $B\in\{B\}$ such that $|V'-\beta B|<\epsilon$. But $\beta B\in Q$ because $B\in Q$. This follows from the fact that every $B\in\{B\}$ may be considered as an element of a basis of $\{B\}$. Since there exists a basis of $\{A\}$ that is reciprocal to the basis of $\{B\}$, there exists always an $(N-2)$-dimensional sublattice of $\{A\}$ orthogonal to a given $B\in\{B\}$.

Definition (108) and proposition (i) of (109) tell us that for a "general" set of initial velocities v_{iQ} the vector 0 is the only element of $\{A\}$ which is orthogonal to V_0':

$V_0'\notin Q$:

$$f^I_{X_0',V_0'}(X',V')=|S|^{-1}\chi^{\bar{S}_1}(X')\frac{1}{N!}\sum_r\delta(V'-\mathscr{P}_rV_0'). \tag{110}$$

This distribution is obviously positive. That (105) is also positive for "special" initial velocities, $V_0'\in Q$, follows from

$$\begin{aligned}\Delta^{(V_0')}(Y')&=\sum_A^{(V_0')}\exp(i\alpha A\cdot Y')\\&=|Z|^{-1}\sum_C^{(V_0')}\delta(\Pi^{(V_0')}Y'-C),\end{aligned} \tag{111}$$

where

$$Z = \text{unit cell of } \{C\},$$

$$\{C\} = \text{reciprocal lattice of } \{A\}^{(V_0')},$$

$$\{A\}^{(V_0')} = \text{sublattice of } \{A\} \text{ orthogonal to } V_0', \quad (112)$$

$$\Pi^{(V_0')} = \text{projection operator into the subspace spanned by the vectors of } \{A\}^{(V_0')}.$$

Note that $\{B\} \subset \{C\}$ on account of $\{A\}^{(V_0')} \subset \{A\}$. Since $\Delta^{(V_0')}$ is positive, the distribution (105) is also positive:

$$V_0' \in Q: \quad f^I_{X_0', V_0'}(X', V') = |S|^{-1} \chi^{\bar{S}_1}(X') \frac{1}{N!} \times \sum_r \Delta^{(\mathscr{P}_r, V_0')}(\mathscr{P}_r X_0' + D_r - X' - D_1) \times \delta(V' - \mathscr{P}_r V_0'). \quad (113)$$

Both (110) and (113) represent uniform distributions on sets which are the closure of the orbit $O(X_0', V_0') = \{X', V' | X' = X_t'[X_0', V_0'], V' = V_t'[X_0', V_0'], 0 \leqslant t < \infty\}$ but the structure of this set depends on the properties of V_0'. For $V_0' \notin Q$, $\bar{O}(X_0', V_0') = \bar{S}_1 \times \cup_r \{\mathscr{P}_r V_0'\}$, where all the $N!$ velocities $\mathscr{P}_r V_0'$ are different because $\mathscr{P}_r V_0' = V_0'$ implies $V' \in Q$ [cf. definitions (11) and (108)]. For $V_0' \in Q$, $\bar{O}(X', V')$ decomposes into a finite union of subsets of dimension D', $1 \leqslant D' \leqslant N-2$. Both D' and the number of subsets depend on V_0', i.e., on the initial velocities of the particles. In the most extreme case $D' = 1$ and the sets $\{X' | X' = X_t'[X_0', \mathscr{P}_r V_0']\}$ form a closed orbit within $\bar{S}_1$. Since $D = N - 1 - D^c = N - 2$ in this case, this happens if, and only if, V_0' is proportional to some $B \in \{B\}$, that is if the evolution is periodic in all variables. Note that V_0' is proportional to some $B \in \{B\}$ if it is proportional to some $A \in \{A\}$ because the definitions (6) and (7) imply

$$NA_j = \sum_{k \neq j} B_{jk}. \quad (114)$$

States with $V_0' \propto A$ are exceptional states in the sense of (109), proposition (i); but it is just these states which are needed if one wants to compare the classical version of the present model with the quantum mechanical one.

There is another class of special initial conditions that is of interest even if the system is treated only classically. Suppose that $N = 2N^*$, $L = 2L^*$, and

$$x_{i,0} = x_{N+1-i,0} \text{ or } \mathscr{P}_{(i,n+1-i)} X_0' = X_0', \quad (115)$$

$$v_{i,0} = v_{N+1-i,0} \text{ or } \mathscr{P}_{(i,N+1-i)} V_0' = V_0'.$$

Then it is easily visualized and also formally proved that (115) is not only valid for $t = 0$ but also for all $t > 0$. The first N^* particles therefore move as if they were trapped between fixed walls at $x = -L^*$ and $x = 0$ while the second half of the system behaves like the mirror image of the first one. Geometrically the configuration space of the fixed wall system N^*, L^* is obtained from the configuration space of the movable wall system N, L by intersecting it with the planes $\mathscr{P}_{(i,N+1-i)} X = X$. This intersection contains the hyperplane $X'' = 0$ and the configuration space of N^*, L^* consists of simplices of dimension N^* each of which may be extended by iterative reflections to a simple covering of R^{N^*}. The methods of calculation used in the following therefore can be transferred to the fixed wall system but the details become more complicated, since there is no cyclic symmetry. It is to be expected that the differences between the two kinds of systems decrease as the systems increase; for the velocity autocorrelation function agreement has been found already for surprisingly small systems ($N^* = 3$).

We conclude this section with a comment on the symmetry properties of the distribution function (105). For general initial conditions this function is invariant under a large number of transformations, which, in general, do not commute with the evolution nor even are canonical transformations:

$$V_0' \notin Q: f^I_{X_0', V_0'}(X', V') = f^I_{X_0', V_0'}(-\mathscr{P}_2 X', V') = f^I_{X_0', V_0'}(\mathscr{P}_3 X' + D_3 - (L/L_d)A_1 - D_1, V') = f^I_{X_0', V_0'}(X', \mathscr{P} V'). \quad (116)$$

Whether a distribution belonging to special initial conditions ($V_0' \in Q$) has some symmetries depends on X_0' and V_0'; there are, however, always many fewer symmetries for special initial conditions than for general ones.

Up to now we have assumed $X_0' \in S_1$; it is obvious how to modify the above results if $X_0' \in S_r$, $r \neq 1$.

V. STATIONARY DISTRIBUTIONS

In the following we shall calculate the expectation values of several observables g for three kinds of stationary ensembles, namely general smallest stationary, microcanonical, and canonical ensembles. The corresponding distribution functions are all of the form

$$f^s_\sigma(X, V) = (1/|S|) \chi^{\bar{S}_s}(X') \varphi_\sigma(V') \delta(X'') \delta(V'') \quad (117)$$

or

$$f_\sigma(X, V) = \frac{1}{N!} \sum_s f^s_\sigma(X, V). \quad (118)$$

It is obvious from Eqs. (117) and (118) that in the first case the order of the particles is known with certainty while it is completely unknown in the second case.

As is clear from the discussion in the preceding section only those smallest stationary ensembles which contain general relative velocities ($V' \notin Q$) are of the form (117):

$$\sigma = \{v_{i0}\}: \quad \varphi_{\{v_{i0}\}}(V') = \frac{1}{N!} \sum_r \delta(V' - \mathscr{P}_r V_0'). \quad (119)$$

We write $\{v_{i0}\}$ to indicate that the corresponding expectation values can be considered as time averages, however, not for one single initial condition X_0', V_0' only but for a whole class of them ($X' \in S_s$, $V' = \mathscr{P} V_0'$).

The microcanonical ensemble (with fixed order and resting center of mass) has a distribution of relative velocities given by

$$\sigma = E: \quad \varphi_E(V') = (v_E^{N-2}|K|)^{-1}\delta(|V'| - v_E), \tag{120}$$

where

$$v_E = \sqrt{2E/m}, \tag{121}$$

$$K = \{V' | V'^2 = 1\}$$

$$= \text{surface of the } (N-1)\text{-dimensional unit sphere},$$

$$|K| = (N-1)\pi^M/\Gamma(M+1), \tag{122}$$

$$M = (N-1)/2. \tag{123}$$

The distribution (120) may be obtained from a superposition of all distributions (119) for which $m\Sigma_i v_{i0}^2 = 2E$. Let $\omega = (\omega_1,...,\omega_{N-2}) =$ set of angles labeling a point of K, $d\omega\,\mu(\omega) =$ surface element of K,

$$E'(\omega) = (e_i(\omega)) \in K, \tag{124}$$

then

$$\varphi_E(V') = |K|^{-1} \int_K d\omega\,\mu(\omega)\varphi_{\{v_E e_i(\omega)\}}(V'). \tag{125}$$

That the special smallest ensembles contained in a microcanonical one may be neglected is not a direct consequence of (109), proposition (i), but emerges from the proof of this assertion.

The velocity distribution of the canonical ensemble is given by

$$\sigma = \mathrm{T}: \quad \varphi_T(V') = (\gamma/\pi)^M \exp[-\gamma V'^2], \tag{126}$$

$$\gamma = m/2kT. \tag{127}$$

The distribution (126) may be written as

$$\varphi_T(V') = \int_0^\infty dE\, h_T(E)\varphi_E(V'), \tag{128}$$

the weight function h_T being given by

$$h_T E = [(kT)^M \Gamma(M)]^{-1} E^{M-1} \exp[-E/(kT)]. \tag{129}$$

All stationary distributions considered here satisfy

$$f_\sigma^s(X,V) = f_\sigma^s(X, \mathscr{P} V). \tag{130}$$

Moreover

$$f_\sigma^1(X,V) = f_\sigma^1(\mathscr{P}_3 X - [L/L_d]A_1, V) = f_\sigma^1(-\mathscr{P}_2 X, V) \tag{131}$$

because of (20), (6), (10), (11), and (35); similar relations hold for $s \neq 1$. Finally

$$f_\sigma(X,V) = f_\sigma(\mathscr{P} X, V) = f_\sigma(X, \mathscr{P} V). \tag{132}$$

These symmetry properties may be used to relate expectation values of different observables and to simplify their calculation.

VI. APPROACH TO EQUILIBRIUM

In this section we discuss under which conditions a nonstationary distribution f_t tends towards a stationary distribution $f_{(\infty)}$. That is, we ask whether the expectation value

$$\langle f_t, g\rangle = \int dX \int dV f_t(X,V) g(X,V) \tag{133}$$

has a limit for $t \uparrow \infty$,

$$\lim_{t\uparrow\infty} \langle f_t, g\rangle = \langle f_{(\infty)}, g\rangle, \tag{134}$$

and how the "equilibrium" distribution $f_{(\infty)}$ depends on the initial distribution f_0.

For the integral (133) to make sense, f and g have to be elements of dual spaces, i.e., one has to be a test function if the other one is a distribution. From a purely mathematical point of view it would be convenient to fix these dual spaces once and forever; this would allow us to arrive at precise conditions for the existence of weak limits $f_t \to f_{(\infty)}$. But every choice of well-known classes of distributions and test functions[6] would be too restrictive to cover all distribution functions f and observables g of interest.

Except for the distribution f_t evolving from $f_0(X,V) = \delta(X - X_0)\delta(V - V_0)$ the f's and g's considered here can be characterized as follows: (i) both f and g factorize in X and V; (ii) f is a test function in X and a distribution in V, and (iii) g is a distribution in X and a test function in V. In the following these conditions will be replaced by three assumptions on f and g which guarantee the existence of a weak limit (134).

The first assumption is that f is of the form

$$f_0^1(X,V) = \begin{cases} f_0^{1I}(X',V')\delta(X'')\delta(V''), \\ 0 \quad \text{for } X' \notin S_1. \end{cases} \tag{135}$$

This assumption serves only to simplify the arguments. It is easy to generalize the following calculations for different orders of the particles and for a moving center of mass. If we define the function g^{1I} by

$$g^{1I}(X',V') = \chi^{\bar{S}_1}(X') g(X',V'), \tag{136}$$

then

$$\langle f_t, g\rangle = \int dX'\, dV' f^{1I}(X',V') g^{1I}(X',V'). \tag{137}$$

Next we assume that the functions $f^{1I}(\cdot,V')$ and $g^{1I}(\cdot,V')$ belong to dual spaces of test functions and distributions, respectively, such that the integrals $I_0^1(V')$, defined by

$$I_0^1(V') = \int dX' f_t^{1I}(X',V') g^{1I}(X',V'), \tag{138}$$

may be computed by means of a Fourier series. This means that the following relations are assumed to hold true for $t = 0$:

$$I_t^1(V') = \sum_A I_{tA}^1(V'),$$

$$I_{tA}^1(V') = |S| F_{tA}^1(V') G_A^{1*}(V'). \tag{139}$$

The Fourier coefficients $H_A^1 (= F_{tA}^1, G_A^1)$ are related to the functions $h^{1I} (= f_t^{1I}, g^{1I})$ by

$$H_A^1(V') = |C|^{-1} \sum_r \int_{S_1} dX'\, h^{1I}(X', \mathscr{P}_r V') \times \exp[-i\alpha A \cdot \mathscr{P}_r(X' + D_1)], \tag{140}$$

with $\alpha = 2\pi/L_d^2$. It is sufficient to assume that (139) holds true for $t = 0$ since, as will be shown later on,

$$F_{tA}^{1I}(V') = F_{0A}^{1I}(V') \exp(-i\alpha A \cdot V' t). \tag{141}$$

It follows from (137), (139), and (141) that the limit (134) exists if, and only if,

$$\lim_{t\uparrow\infty}\int dV'\, I^1_{0A}(V')\exp(-i\alpha A\cdot V' t)=0,\quad \text{for all } A\neq 0. \tag{142}$$

The validity of (142) constitutes our third assumption. To see that (142) is indeed satisfied in many cases it is useful to consider a few examples. To this end let us assume that both f^{1I} and g^{1I} factorize in X' and V',

$$h^{1I}(X',V')=h^{1IX}(X')h^{1IV}(V'). \tag{143}$$

Then

$$H^1_A(V')=\sum_r H^1_{Ar}h^{1IV}(\mathscr{P}_r V'), \tag{144}$$

$$H^1_{Ar}(V')=|C|^{-1}\int_{S_1} dX'\; h^{1IX}(X')\times\exp[-i\alpha A\cdot\mathscr{P}_r(X'+D_1)],$$

and

$$I^1_{0A}(V')=\sum_{r,s}F^1_{Ar}G^{1*}_{As}f_0^{1IV}(\mathscr{P}_r V')g^{1IV}(\mathscr{P}_s V'). \tag{145}$$

Furthermore, let g^{1IV} be a polynomial in the velocities v_i and consider the following velocity distribution functions.

(i) f_0^{1IV} is a Maxwellian distribution (subject to $\Sigma_i v_i=0$). Then $f_0^{1IV}(\mathscr{P}_r V')g^{1IV}(\mathscr{P}_s V')$ is a rapidly decreasing function. Its Fourier transform is therefore also rapidly decreasing so that (142) is satisfied.

(ii) Let f_0^{1IV} be a Gaussian centered around V_0' or the characteristic functions of a parallelepiped containing V_0'. The Fourier transforms of these functions are well-known analytical functions which satisfy (142). Note that the smaller the width of f_0^{1IV} is, the slower is the decay of its Fourier transform and hence the approach to equilibrium.

(iii) Let $f_0^{1IV}(V')$ be proportional to $\delta(|V'|-v_E)$. If one introduces polar coordinates for V', i.e., a radial variable v and angular variables $\omega=(\omega_1,\ldots,\omega_{N-2})$, then the integration over v may be carried out. The remaining integrals are of the form $\int d\omega\, A(\omega)\exp[i\phi(\omega)t]$, where A is a real polynomial in trigonometric functions of $\omega_1,\ldots,\omega_{N-2}$ and $\phi(\omega)$ proportional to $\cos(\omega_1)$, if the coordinate system is properly chosen. The method of the stationary phase or the Riemann–Lebesgue lemma can then be used to show that (142) is satisfied.

(iv) Let $f_0^{1IV}(V')=\int d\eta\, a(\eta)\delta(V'-\eta V_0')$, where $V_0'\notin Q$ and $a(\eta)$ is a smooth rapidly decreasing function. Then $V_0'\cdot A\neq 0$ for all A and the integral

$$\int dV'\, I^1_{0A}(V')\exp(-i\alpha A\cdot V' t)$$

$$=\sum_{r,s}F^1_{Ar}G^{1*}_{As}\int d\eta\, a(\eta)g^{1IV}(\eta\mathscr{P}_s\mathscr{P}_r^{-1}V_0')\times\exp(-i\alpha A\cdot V_0' t) \tag{146}$$

vanishes for $t\uparrow\infty$ on account of the Riemann–Lebesgue lemma.

Note the difference between the first and the second two examples: In (i) and (ii) $\langle f_t, g\rangle$ is proportional to the product of $(N-1)$ integrals each of which vanishes for $t\uparrow\infty$, while in (iii) and (iv) the expectation value contains only one such integral so that much more time is needed to approach the equilibrium value.

It is instructive to look not only for examples where (142) is valid but also for some counterexamples.

(i') Assume that the initial velocities are known. Then $f_0^{1IV}(V')=\delta(V'-V_0')$ and $\langle f_t, g\rangle$ is an almost periodic function. This is also true for any finite linear combination of delta functions but the larger the number of terms, the larger will, in general, become the "period" of the expectation value. The present example may be considered as the "limit" of examples (ii) or (iv) above: The more the velocity distribution becomes concentrated around V_0' the slower tends the expectation value toward a constant value.

(ii') Let $f_0^{1IV}(V')=\int d\eta\, a(\eta)\delta(V'-V_0'-\eta B_{12})$; this means that the initial (relative) velocities are completely known except for a certain indeterminacy of the relative velocity v_1-v_2. In this case the integral $\int d\eta\, a(\eta)\times\exp(-i\alpha A\cdot B_{12}t\eta)$ will not tend to zero for all A's. A weak limit will exist only if for all A's with $A\cdot B_{12}=0$, i.e., $A=(a_i)$ with $a_1=a_2$, at least one of the prefactors F^1_{Ar} or G^1_{As} vanishes. This, however, happens only if f_0^{1IX} or g^{1IX} is constant with respect to the variable x_1-x_2.

The conclusion to be drawn from all these examples is that a certain spread in the initial velocities is necessary (and in many cases even sufficient) for a time-dependent expectation value to tend toward a constant value.

If Eqs. (135), (139), (141), and (142) hold true, then

$$\lim_{t\uparrow\infty}\langle f_t, g\rangle$$

$$=\int dV'\, I^1_{t0}(V')=\int dV'\,|S|F^1_{t0}(V')G^{1*}_0(V')$$

$$=\frac{1}{N!}\sum_r\int dX'\int dV'\,F^1_{00}(V')g^{1I}(X',\mathscr{P}_r V')$$

$$=\int dX'\int dV'\,\frac{1}{N!}\sum_r F^1_{00}(\mathscr{P}_r^{-1}V')g^{1I}(X',V')$$

$$=\langle f^1_{(\infty)}, g\rangle, \tag{147}$$

where the equilibrium distribution is given by

$$f^1_{(\infty)}(X,V)=|S|^{-1}\chi^{\bar S_1}(X')(1/N!)\times\sum_r\int_{S_1}dY'\,f_0^{1I}(Y',\mathscr{P}_r V')\delta(X'')\delta(V''). \tag{148}$$

The various factors of this stationary distribution are easily interpreted: The function of X' shows that all configurations with the same order as the initial one become equally probable for large times; this is a consequence of the confinement of the particles. The function of V' indicates that in equilibrium every distribution of the initial velocities onto the N particles is equally probable. Since this does not occur for noninteracting particles this fact has to be attributed to the collisions between the particles (cf. I, Sec. 3).

It remains to prove that (147) follows indeed from the previous definitions and assumptions. The first equality follows from (142). Because of (41),

$$\langle f^1_t, g\rangle=\langle f_t^{1I}, g^{1I}\rangle_{S_1}=|C|^{-1}|S|\langle\bar f_t^{1I},\bar g^{1I}\rangle_C, \tag{149}$$

where the barred functions $\bar{h}^{1I}(=\bar{f}_t^{1I}, \bar{g}^{1I})$ are, by definition (36), periodic in X' with respect to translations B. According to our second assumption the function $\bar{h}^{1I}$ may be represented as Fourier series with coefficients

$$\begin{aligned} H_A^1(V') &= |C|^{-1}\int_C dX'\,\bar{h}^{1I}(X',V')\exp(-i\alpha A\cdot X') \\ &= |C|^{-1}\sum_r \int_{S_r+D_r} dX'\,\bar{h}^{1I}(X',V') \\ &\quad\times\exp(-i\alpha A\cdot X') \\ &= |C|^{-1}\sum_r \int_{S_1+D_1} dX'\,\bar{h}^{1I}(X',\mathscr{P}_r^{-1}V') \\ &\quad\times\exp(-i\alpha A\cdot \mathscr{P}_r X') \\ &= |C|^{-1}\sum_r \int_{S_1} dX'\,h^{1I}(X',\mathscr{P}_r^{-1}V') \\ &\quad\times\exp[-i\alpha A\cdot \mathscr{P}_r(X'+D_1)] \end{aligned} \tag{150}$$

[cf. (140)]; this explains the second equality. The third one follows from (141), which in turn is a consequence of (74).

We conclude this section by pointing out that the discussion is easily extended to time-dependent correlation functions of the form

$$\langle g\,h_t\rangle_0^1 = \int dX \int dV f_0^1(X,V)g(X,V)h_t(X,V), \tag{151}$$

with

$$h_t(X,V) = h_0(X_t[X,V],V_t[X,V]). \tag{152}$$

All the previous arguments may be repeated if $f(X,V)g(X,V)$ is substituted for $g(X,V)$ and $h_t(X,V)$ for $f_t(X,V)$. For a complete coincidence one would have to change also $+t$ into $-t$ [cf. Eqs. (152) and (69)], but this does not affect the result because if (142) is valid for $+t$ it is also valid for $-t$. Thus if the limit of (151) exists,

$$\lim_{t\uparrow\infty}\langle g\,h_t\rangle_0^1 = \langle g\,h_{(\infty)}^1\rangle_0^1, \tag{153}$$

then $h_{(\infty)}^1$ is given by

$$h_{(\infty)}^1(X,V) = |S|^{-1}\chi^{S_1}(X')\frac{1}{N!}\sum_r\int_{S_1} dX'\,h_0(X,\mathscr{P}_r V). \tag{154}$$

The velocity autocorrelation function, calculated in the following paper in full detail, is an example that fits into this scheme. For a smallest stationary ensemble this function is almost periodic [cf. example (i') above]. For the microcanonical and the canonical ensemble the limit (153) exists and it is zero because of the symmetrization with respect to the velocities, Eq. (154) and the factor $\delta(V'')$ in f_0^1. From examples (i) and (iii) it may conjectured that the function decays faster for the canonical ensemble than for the microcanonical one; this is indeed verified by explicit calculation (see, e.g., I, Figs. 6 and 13).

VII. THE PAIR DISTRIBUTION FUNCTION

The pair distribution function is defined by

$$g(r) = \frac{1}{N}\sum_j\sum_{k(\neq j)}\langle f_\sigma^s,\delta(j,k;r)\rangle, \tag{155}$$

where f_σ^s is a stationary distribution function of the form (117) and

$$\delta(j,k;r,X) = \delta(x_j - x_k - r). \tag{156}$$

Since the observable (115) depends only on X the rhs of (155) is the same for $\sigma=\{v_{i0}\}$, E, T, i.e., for almost all stationary distributions (cf. Secs. IV and V). The quantity $\langle f_\sigma^s,\delta(j,k;r)\rangle dr$ is the probability for $x_j - x_k\in(r+dr)$; therefore $\Sigma_{k\neq j}\cdots dr$ is the number of particles expected to be in the interval $(r,r+dr)$, where r is counted from the position of the particle j. The normalized sum $(1/N)\Sigma_j\cdots$ expresses the fact that the tagged particle is chosen at random. The normalization

$$\int dr\,g(r) = N-1 \tag{157}$$

is a direct consequence of the definition (155) and consistent with the interpretation given above. In this interpretation it has been tacitly assumed that the particles are arranged in a definite order s. But the relation

$$f_\sigma^s(X,V) = f_\sigma^1(\mathscr{P}_s^{-1}X,V) \tag{158}$$

can be used to show that the rhs of (155) is the same for all orders s as has been already anticipated in the notation of the lhs.

Choosing $s=1$ the mathematical task reduces to the computation of the quantities

$$q_{j,k}(r) = \langle f_\sigma^1,\delta(j,k;r)\rangle. \tag{159}$$

In doing so we may exploit the symmetry properties of the simplex S_1 and its characteristic function χ^{S_1} contained in f_σ^1. From (131) and the orthogonality of the transformations $\mathscr{P}$, and $\mathscr{P}_3^{-1}A_1 = A_N$, we get

$$\begin{aligned} &\int dX\int dV f_\sigma^1(X,V)\delta(j,k;r,X) \\ &\quad=\int dX\int dV f_\sigma^1\left(\mathscr{P}_3 X - \left[\frac{L}{L_d}\right]A_1,V\right)\delta(j,k;r,X) \\ &\quad=\int dX\int dV f_\sigma^1(X,V)\delta\left(j,k;r,\mathscr{P}_3^{-1}X + \left[\frac{L}{L_d}\right]A_N\right). \end{aligned} \tag{160}$$

This implies that many of the q's coincide, namely

$$\begin{aligned} &1\leqslant j<k\leqslant N: \quad q_{j,k}(r) = q_{1,1+k-j}(r), \\ &N\geqslant j>k\geqslant 1: \quad q_{j,k}(r) = q_{N,N+k-j}(r), \end{aligned} \tag{161}$$

so that only $q_{1,k}$, $2\leqslant k\leqslant N$, and $q_{N,j}$, $1\leqslant j\leqslant N-1$, have to be calculated. The second equality of (131) entails

$$\begin{aligned} &\int dX\int dV f_\sigma^1(X,V)\delta(j,k;r,X) \\ &\quad=\int dX\int dV f_\sigma^1(-\mathscr{P}_2 X,V)\delta(j,k;r,X) \\ &\quad=\int dX\int dV f_\sigma^1(X,V)\delta(j,k;r,-\mathscr{P}_2 X), \end{aligned} \tag{162}$$

whence

$$q_{j,k}(r) = q_{N+1-j,N+1-k}(r), \tag{163}$$

and it is sufficient to calculate $q_{N,k}(r)$ for $k=1,\ldots,N-1$.

Note that these functions have a simple geometrical meaning from which their support is easily deduced:

$$q_{N,k}(r) = |S|^{-1}|S_1 \cap \text{plane } x_N - x_k = r| = 0, \quad \text{for } r \notin [[N-k]d, L-kd]. \tag{164}$$

Since the intersection (164) is not easily computed for larger values of N, we consider the symmetry properties of the observable (156), which have not been exploited so far. To this end we rewrite $q_{N,k}$ as

$$q_{N,k}(r) = |S^-|^1 \int_{S_1} dX'\, \delta(N,k;r,X') = |S|^{-1} \int_{S_1 + D_1} dX'\, \delta(N,k;r-[N-k]d,X'). \tag{165}$$

The function $\delta(N,k;r,X')$ is invariant under all transformations $\mathscr{P}$ with $Pk = k$ and $PN = N$. Now let the transformations $\mathscr{P}_u$ be related to the permutations P_u with

$$\begin{aligned} &j(u) \in \{1,...,k-1\}, \quad \text{for } j = 1,...,k-1; \\ &k(u) = k; \\ &l(u) \in \{k+1,...,N-1\}, \quad \text{for } l = k+1,...,N-1; \\ &N(u) = N \end{aligned} \tag{166}$$

(the first relation has to be dropped for $k = 1$ and the third one for $k = N-1$). Then

$$\delta(N,k;r,X) = \delta(N,k;r,\mathscr{P}_u X), \tag{167}$$

$$\begin{aligned} q_{N,k}(r) &= |S|^{-1} \frac{1}{(k-1)!(N-1-k)!} \times \sum_u^{(N,k)} \int_{S_1 + D_1} dX' \times \delta(N,k;r-[N-k]d,\mathscr{P}_u X') \\ &= |S|^{-1} \frac{1}{(k-1)!(N-1-k)!} \times \int_{V_1^{(N,k)}} dX'\, \delta(N,k;r-[N-k]d,X'), \end{aligned} \tag{168}$$

with

$$V_1^{(N,k)} = \bigcup_u^{(N,k)} \mathscr{P}_u (S_1 + D_1). \tag{169}$$

The closure of the set (169) is

$$\overline{V}_1^{(N,k)} = \left\{ X' \,\middle|\, \begin{array}{l} x_k \leqslant x_N \leqslant x_k + L_d \\ x_N - L_d \leqslant x_j \leqslant x_k, \quad \text{for } j = 1,...,k-1 \\ x_k \leqslant x_l \leqslant x_N, \quad \text{for } l = k+1,...,N-1 \end{array} \right\}. \tag{170}$$

To verify that the rhs of (170) is indeed the closure of the rhs of (169) assume first that $X' \in \mathscr{P}_u(S_1 + D_1)$. Then $\mathscr{P}_u^{-1} X' \in S_1 + D_1$, i.e., $x_{1(u)} < x_{2(u)} < \cdots < x_{N(u)} < x_{1(u)} + L_d$. But this implies (i) $x_k < x_N < x_k + L_d$, (ii) $X_N - L_d < x_{j(u)} < x_k$, for $j(u) = 1,...,k-1$, and (iii) $x_k < x_{l(u)} < x_N$, for $l(u) = k+1,...,N-1$; hence $X' \in \overline{V}_1^{(N,k)}$. Next assume $X' \in \overline{V}_1^{(N,k)}$, i.e., $x_k \leqslant x_N \leqslant x_k + L_d$, etc. Now order the variables x_j and x_l according to their magnitude: $x_N - L_d \leqslant x_{1(u)} \leqslant x_{2(u)} \leqslant \cdots \leqslant x_{(k-1)(u)} \leqslant x_k \leqslant x_{(k+1)(u)} \leqslant \cdots \leqslant x_{(N-1)(u)} \leqslant x_N$ with $j(u) \in \{1,...,k-1\}$ and $l(u) \in \{k+1,...,N-1\}$. The vectors X' for which at least one equality holds form a set of measure zero. Removing this set from $\overline{V}_1^{(N,k)}$ we see that for each of the remaining elements of $\overline{V}_1^{(N,k)}$ there exists a unique permutation P_u of the form (166) such that $\mathscr{P}_u^{-1} X' \in S_1 + D_1$.

To calculate the integral (168) we pass from the orthogonal coordinates $\bar{x}_1,...,\bar{x}_{N-1}$ (cf. Sec. II B) to the nonorthogonal coordinates

$$y_m = x_N - x_m, \quad m = 1,...,N-1. \tag{171}$$

The functional determinant is

$$\left| \frac{\partial(y_1,...,y_{N-1})}{\partial(\bar{x}_1,...,\bar{x}_{N-1})} \right| = \left| \frac{\partial(y_1,...,y_{N-1},\bar{x}_N)}{\partial(\bar{x}_1,...,\bar{x}_{N-1},\bar{x}_N)} \right| = \left| \frac{\partial(y_1,...,\bar{x}_N)}{\partial(x_1,...,x_N)} \right| = \sqrt{N}. \tag{172}$$

Recalling that $(\sqrt{N}|S|)^{-1} = (N-1)! L_d^{1-N}$ we finally arrive at

$$\begin{aligned} q_{N,k}(r) &= \frac{(N-1)!}{L_d^{N-1}(k-1)!(N-1-k)!} \times \int_0^{L_d} dy_k\, \delta(y_k - r + [N-k]d) \times \left[\int_{y_k}^{L_d} dy_j\right]^{k-1} \left[\int_0^{y_k} dy_l\right]^{N-1-k} \\ &= \frac{(N-1)!}{L_d(k-1)!(N-1-k)!} \theta(L - kd - r) \times (1 - (r - [N-k]d)/L_d)^{k-1} \times \theta(r - [N-k]d), \\ &\quad \times ((r - [N-k]d)/L_d)^{N-1-k}, \end{aligned} \tag{173}$$

where

$$\theta(x) = \begin{cases} 1 & \text{for } x > 0, \\ 0 & \text{for } x < 0. \end{cases} \tag{174}$$

The functions (173) are the so-called β distributions.[7] Then,

$$\lim_{r \downarrow d} q_{N,k}(r) = \delta_{k,N-1}[(N-1)/L_d]. \tag{175}$$

It follows from the definition (155) that g is an even function,

$$g(r) = g(-r). \tag{176}$$

If $r > 0$ then $q_{j,k}(r) = 0$ for $j < k$, as may be deduced from (163) and (164); this and the second relation (161) yield

$$g(r) = \frac{1}{N} \sum_{n=0}^{N-2} (N-1-n) q_{N,N-1-n}(r), \quad \text{for } r > 0. \tag{177}$$

Combining (177), (173), and (176) we finally obtain

$$\begin{aligned} g(r) &= \frac{(N-1)!}{L_d N} \sum_{n=0}^{N-2} \frac{(N-1-n)}{n!(N-2-n)!} \times \theta(|r| - [n+1]d) \theta(L_d + [n+1]d - |r|) \\ &\quad \times \left(\frac{|r| - [n+1]d}{L_d}\right)^n \left(1 - \frac{|r| - [n+1]d}{L_d}\right)^{N-2-n}, \end{aligned} \tag{178}$$

$$\lim_{r\downarrow d} g(r) = (N-1)^2/NL_d$$
$$= [(N-1)/N] \lim_{d\downarrow 0} q_{k+1,k}(r),$$
$$\text{for} \quad 1 \leqslant k \leqslant N-1. \tag{179}$$

In the thermodynamic limit ($N\uparrow\infty$, $N/L=\rho=\text{const}$; $L_d\uparrow\infty$, $N/L_d=\rho_d=\text{const}$; r finite) the functions $q_{j,1}$ become the Poisson distributions

$$q_{j,1}(r) = \theta(r-[j-1]d)\rho_d[1/(j-2)!] \times\{\rho_d[r-(j-1)d]\}^{j-2} \times\exp\{-\rho_d[r-(j-1)d]\} \tag{180}$$

and

$$g(r) = \rho_d \sum_{n=0}^{\infty} \theta(|r|-[n+1]d)\frac{1}{n!} \times\{\rho_d[|r|-(n+1)d]\}^n \times\exp\{-\rho_d[|r|-(n+1)d]\}, \tag{181}$$

$$\lim_{r\downarrow d} g(r) = \rho_d = \lim_{r\downarrow d} q_{k+1,k}(r), \quad \text{for } 1\leqslant k<\infty. \tag{182}$$

The pair distribution function (181) has been already derived by several authors, both heuristically[8,9] and formally.[10]

The above results simplify for point particles ($d=0$). Splitting the sum into two parts, the summation over n may be carried out and g becomes

$$g(r) = \theta(L-|r|)\left(\frac{N-1}{L}\right)\left[\frac{N-2}{N}\left(1-\frac{|r|}{L}\right)+\frac{1}{N}\right], \quad \text{for } N<\infty, \tag{183}$$

$$g(r) = \rho, \quad \text{for } N\uparrow\infty. \tag{184}$$

The arguments used to derive the polynomial representation (178) of the pair distribution function may be generalized to derive similar formulas for multiplet distribution functions.[11] However, to illustrate the difference between finite systems and their thermodynamic limit it is quite sufficient to see how the function (178) varies as the size of the system increases. This can be seen in Fig. 14 of I which suggests that the first few peaks of the pair distribution function for $N\uparrow\infty$, Eq. (181), agree well with those obtained for rather small systems ($N\approx 20$).

VIII. THE EQUATION OF STATE

If one wants to calculate the pressure of a one-dimensional hard rod system, the problem consists less in evaluating a given formula than in finding an appropriate definition of this quantity. Different authors have used different definitions of the pressure (see, e.g., Refs. 2, 11, and 25 of I) and it is by no means obvious that they all lead to the same equation of state. Here we start with a definition in terms of point mechanics: what we call "pressure" is the time average of the force exerted onto a tagged particle from its left neighbor. After having found the corresponding phase space function we calculate its expectation value for the various distributions considered in this paper. The results for the microcanonical and canonical ensembles are then compared with those derived from the corresponding partition functions and found to coincide or at least to approach each other in the thermodynamic limit.

The equivalence of the "mechanical" and the "thermodynamical" definitions of the pressure for systems with fixed walls, onto which the particles are elastically reflected, has been proven by Presutti.[12] However his proof does not include the model considered here in which the positions of the walls, even if they were completely specified, change in the course of time. The following calculations may therefore be regarded as a simple variant of Presutti's general considerations.

The observable measuring the change of momentum of a tagged particle, which is caused by collisions with its left neighbor during the time interval $(0,T)$, is

$$m(j,T;X,V) = \int_0^T dt\, m(j;X_t[X,V],V_t[X,V]), \tag{185}$$

$$m(j;X,V) = m(j;\mathscr{P}X,\mathscr{P}V),$$

$$X'\in S_1: \quad m(j;X,V) = \lim_{\tau\downarrow 0} \delta(x_{j,-\tau}[X,V] - x_{j-1+N\delta_{j1},-\tau}[X,V] - d + L\delta_{j1}) \times|v_j - v_{j-1+N\delta_{j1}}|m(v_j - v_{j-1+N\delta_{j1}}). \tag{186}$$

In interpreting Eqs. (185) and (186) one should take the following into account: (i) what is measured for the particle with number j, if $X'\in S_1$, is measured for the jth particle counted from the left, in general; (ii) the left neighbor of the first particle is the Nth particle; and (iii) if one assumes that exactly one collision takes place in $(0,T)$, then $\int dt\,\delta(\cdots)|v_j\cdots| = 1$ and $m(v_j - v_{j-1})$ is the transferred momentum since v_j is the velocity of particle j after and v_{j-1} the velocity before the collision.

For stationary distribution functions $f_\sigma(X,V) = f_\sigma(X_{-t}[X,V], V_{-t}[X,V])$ so that

$$\langle f_\sigma, m(j,T)\rangle = T\langle f_\sigma, m(j)\rangle. \tag{187}$$

The expectation value of the average force exerted onto the jth particle by its left neighbor is therefore

$$p_{j,\sigma} = (1/T)\langle f_\sigma, m(j,T)\rangle = \langle f_\sigma, m(j)\rangle. \tag{188}$$

The vectors X, V in (186) contain the positions and velocities at an infinitesimal time after the collision where $x_j + \delta_{j1}L = x_{j-1+N\delta_{j1}} + d$. For sufficiently small times after a collision all particles move freely. Thus defining $\sigma_j[V']$ by

$$\sigma_j[V'] = \operatorname{sgn}(v_j - v_{j-1+N\delta_{j1}}), \tag{189}$$

we may write (186) as

$$X'\in S_1: \quad m(j;X,V) = \lim_{\tau\downarrow 0} \delta(x_j - \tau v_j - x_{j-1+N\delta_{j1}} + \tau v_{j-1+N\delta_{j1}} - d + L\delta_{j1}) \times m\sigma_j[V'](v_j - v_{j-1+N\delta_{j1}})^2$$
$$= \lim_{\xi\downarrow 0} \delta(x_j - x_{j-1+N\delta_{j1}} - d$$

$$+L\delta_{j1}-\sigma_j[V']\xi)$$
$$\times m\sigma_j[V'](v_j-v_{j-1+N\delta_{j1}})^2. \tag{190}$$

For the distribution function f^1_σ, given by Eq. (117) with $s=1$, one obtains, using (159), (164), and (175),

$$p^1_{j,\sigma}=\langle f^1_\sigma, m(j)\rangle=\int dV'\,\varphi_\sigma(V')\lim_{\xi\downarrow 0} q_{jj-1+N\delta_{j1}}$$
$$\times(d-L\delta_{j1}+\sigma_j[V']\xi)$$
$$\times\sigma_j[V']m(v_j-v_{j-1+N\delta_{j1}})^2$$
$$=\int dV'\,\varphi_\sigma(V')\theta(v_j-v_{j-1+N\delta_{j1}})$$
$$\times((N-1)/L_d)m(v_j-v_{j-1+N\delta_{j1}})^2. \tag{191}$$

In the following we drop the superscript 1 because the same integral is obtained starting from any f^s_σ. Because $\varphi_\sigma(V')=\varphi_\sigma(\mathscr{P}V')$ (see Sec. V) $\varphi_\sigma(V')$ may be replaced by $\frac{1}{2}[\varphi_\sigma(V')+\varphi_\sigma(\mathscr{P}_{(j,j-1+N\delta_{j1})}V')]$ in the integral (191). The two contributions may be combined to obtain

$$p_{j,\sigma}=\frac{N-1}{L_d}\int dV'\,\varphi_\sigma(V')\frac{m}{2}(v_j-v_{j-1+N\delta_{j1}})^2. \tag{192}$$

Instead of $\varphi_\sigma(V')$ we may write $(1/N!)\Sigma_r\varphi_\sigma(\mathscr{P}_rV')$ without altering the result. Substituting $V'\to\mathscr{P}_r^{-1}V'$ in each of the $N!$ integrals, we see that it is also possible to replace the function $(v_j-v_{j-1+N\delta_{j1}})^2$ in (192) by its symmetrized mean

$$\frac{1}{N(N-1)}\sum_{j,k}(v_j-v_k)^2=\frac{2}{N-1}(V')^2; \tag{193}$$

this implies $p_{j,\sigma}=p_{k,\sigma}=p_\sigma$. Carrying out the integration one finds

$$p_{\{v_{i0}\}}=\frac{m}{L_d}\sum_j v_{j0}^2, \tag{194}$$
$$p_E=2E/L_d, \tag{195}$$
$$p_T=(N-1)kT/L_d. \tag{196}$$

It is instructive to obtain the pressure p_T from (192) using the weight function $\varphi_T(V)\delta(V'')$ and integrating over all velocities v_i. For $N>2$,

$$\int dV\left(\frac{\gamma}{\pi}\right)\frac{N-1}{2}\exp[-\gamma(v_1^2+\cdots+v_N^2)]$$
$$\times\sqrt{N}\,\delta(v_1+\cdots+v_N)\delta(v_j-v)\delta(v_{j-1+N\delta_{j1}}-w)$$
$$=(\gamma/\pi)\exp[-\gamma(v^2+w^2)]F_{N,\gamma}(v+w), \tag{197}$$
$$F_{N,\gamma}(z)=[N/(N-2)]^{1/2}\exp[-\gamma z^2/(N-2)], \tag{198}$$

and therefore

$$p_T=\sqrt{\frac{\gamma}{\pi}}\int dv\,e^{-\gamma v^2}p_{N,\gamma}(v), \tag{199}$$
$$p_{N,\gamma}(v)=\sqrt{\frac{\gamma}{\pi}}\int dw\,e^{-\gamma w^2}\frac{m}{2}(v-w)^2$$
$$\times\rho_d\left(1-\frac{1}{N}\right)F_{N,\gamma}(v+w). \tag{200}$$

Here, $p_{N,\gamma}(v)$ may be interpreted as the pressure onto a particle with velocity v moving in a system of N particles with Maxwellian velocity distributions. The factor $(1-1/N)F_{N,\gamma}$ results from the assumption that the center of mass is at rest. The influence of this constraint decreases with increasing size of the system and vanishes completely in the thermodynamic limit.

In the thermodynamic definition of the pressure we start from a partition function Z, build the thermodynamic potential $k\ln Z$, and obtain from it the equation of state, i.e., the pressure as a function of the variables characterizing the underlying ensemble. For the microcanonical ensemble (of systems with resting center of mass and fixed order, e.g., $X'\in S_1$),

$$Z(E,L,N)=\left(\frac{m}{h}\right)^{N-1}\int_{S_1}dX'\int_{\Delta V'(E)}dV'$$
$$=\left(\frac{m}{h}\right)^{N-1}|S|\,|\Delta V'(E)|, \tag{201}$$

where

$$\Delta V'(E)=\{V'|2E<m(V')^2<2(E+\Delta E)\} \tag{202}$$

is an $(N-1)$-dimensional spherical sheet of radius $\sqrt{2E/m}$ and thickness $\Delta E/\sqrt{2mE}$ ($\Delta E\ll E$). Its volume $|\Delta V'(E)|$ can be calculated by means of spherical coordinates[13]; it is approximately

$$|\Delta V'(E)|\approx\frac{\sqrt{\pi}}{(N-1)/2}\left(\frac{2\pi}{m}\right)^{(N-2)/2}E^{(N-4)/2}\Delta E. \tag{203}$$

Hence

$$p_E=p(E,L,N)=\frac{(\partial/\partial L)k\ln Z(E,L,N)}{(\partial/\partial E)k\ln Z(E,L,N)}$$
$$=\frac{N-1}{N-4}\frac{2E}{L_d}. \tag{204}$$

This expression is, of course, meaningless for $N<5$ but approaches the mechanical pressure (195) with increasing N.

For the canonical ensemble (with constraint $X''=0$, $V''=0$, and order $X'\in S_1$),

$$Z(T,L,N)=\left(\frac{m}{h}\right)^{N-1}\int_{S_1}dX'\int dV'\exp\left[-\left(\frac{m}{kT}\right)(V')^2\right]$$
$$=\left(\frac{2\pi mkT}{h^2}\right)^{(N-1)/2}\frac{L_d^{N-1}}{\sqrt{N}(N-1)!}, \tag{205}$$

so that

$$p_T=p(T,L,N)=-T\frac{\partial}{\partial L}k\ln Z(T,L,N)$$
$$=\frac{(N-1)kT}{L_d}. \tag{206}$$

That this coincides with the mechanical pressure (196) even for arbitrarily small systems ($N\geqslant 2$) seems to be fortuitous in view of the difference between (204) and (195). It is, however, well known that the thermodynamical pressure (206) is related to the pair potential and the pair distribution function through the virial theorem. For a one-dimensional system with potential

$$\phi_N(x_1,\ldots,x_N;L)$$
$$=\sum_{j<k}[\phi(|x_j-x_k|)+\phi(L-|x_j-x_k|)],$$

ϕ = smooth function with $\lim_{x\downarrow 0}\phi(x) = +\infty$, (207)

this relation reads

$$\frac{p_T}{kT} = \frac{N-1}{L} + \frac{N}{L}\int_0^L dr[\, g(r) + g(L-r)] \times e^{\beta\phi(r)} r \frac{d}{dr} e^{-\beta\phi(r)}. \quad (208)$$

Assuming as usual that the functions $g(r)\exp[\beta\phi(r)]$ and $g(L-r)\exp[\beta\phi(r)]$ are continuous also in the limit of hard rod potential,

$$\exp[-\beta\phi(r)] \to \theta(r-d), \quad (209)$$

one finds

$$\frac{p_T}{kT} = \frac{N-1}{L} + \frac{Nd}{L}\lim_{\delta\downarrow 0}[\, g(d+\delta) + g(L-d-\delta)], \quad (210)$$

in agreement with (206), (179), and

$$\lim_{r\uparrow(L-d)} g(r) = (1/N)\lim_{r\uparrow(L-d)} q_{N,1}(r) = (N-1)/NL_d \quad (211)$$

[cf. (177) and (173)].

[1] P. Kasperkovitz and J. Reisenberger, Phys. Rev. A **31**, 2639 (1985).
[2] P. Kasperkovitz, J. Phys. C **9**, 217, 227, 4069 (1976).
[3] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1978), pp. 271–290.
[4] M. V. Berry, AIP Conf. Proc. **46**, 16 (1978).
[5] W. Maak, *Fastperiodische Funktionen* (Springer, Berlin, 1967), 2nd ed.
[6] D. S. Jones, *Generalized Functions* (McGraw, New York, 1966).
[7] W. Feller, *An Introduction to Probability Theory and its Applications* (Wiley, New York, 1966), Vol. II, p. 49.
[8] F. Zernike and J. A. Prins, Z. Phys. **41**, 184 (1927).
[9] J. Frenkel, *Kinetic Theory of Liquids* (Oxford U.P., New York, 1940), p. 126.
[10] Z. W. Salsburg, R. W. Zwanzig, and J. G. Kirkwood, J. Chem. Phys. **21**, 1098 (1953).
[11] J. Reisenberger, Dissertation, Technische Universität Wien, 1983.
[12] E. Presutti, J. Stat. Phys. **13**, 301 (1975).
[13] E. Madelung, *Die Mathematischen Hilfsmittel des Physikers* (Springer, Berlin, 1964), p. 243.

Finite hard rod systems and their thermodynamic limit. III. Collision frequency, no-collision probability, and velocity autocorrelation function

P. Kasperkovitz and J. Reisenberger
Institut für Theoretische Physik, Technische Universität Wien, Karlspatz 13, A-1040 Wien, Austria

(Received 7 November 1984; accepted for publication 21 March 1985)

We use the methods developed for the one-dimensional hard rod gas in the preceding papers to derive exact expressions for the average number of collisions per unit time, the probability that a tagged particle does not collide with other ones within a given time interval, and the velocity autocorrelation function. It is discussed how these expectation values, obtained for finite systems and three kinds of ensembles (smallest stationary, microcanonical, and canonical), behave as the size of the system increases (thermodynamic limit).

I. INTRODUCTION

This is the third of a series of papers on hard rod systems. In the first two papers,[1,2] cited as I and II in the following, we stated and discussed our results and established notation. This paper contains the proofs of our results on time-dependent quantities. In Secs. II and III the collision frequency and the no-collision probability are derived; in the remaining part of the paper (Sec. IV) the velocity autocorrelation is calculated by two complementary methods that are specially suited to finding the form of this function for short and long times, respectively.

II. THE COLLISION FREQUENCY

If one considers the form of time-dependent expectation values for very short times a characteristic time unit emerges quite naturally. It is given by the inverse of the collision frequency which is defined as the average number of collisions a tagged particle undergoes per unit time. To calculate this quantity we start from the observable

$$c(j,T;X,V)=\int_0^T dt\, c(j;X_t[X,V],V_t[X,V]),$$

$$c(j;X,V)=c(j;\mathscr{P}X,\mathscr{P}V), \tag{1}$$

$$X'\in S_1:\quad c(j;X,V)=\lim_{\tau\downarrow 0}\delta(x_{j,-\tau}[X,V]-x_{j-1+N\delta_{j1},-\tau}[X,V]-d+L\delta_{j1})|v_j-v_{j-1+N\delta_{j1}}|. \tag{2}$$

The form of this observable and its interpretation are similar to the one used to calculate the pressure (II, Sec. VIII). Thus if the interval $(0,T)$ is so short that the jth particle, counted from the left, collides with its left neighbor only once, the integral (1) is just equal to 1.

The expectation value of (1) for a stationary distribution

$$f_\sigma(x,V)=f_\sigma(X_{-t}[X,V],V_{-t}[X,V]) \tag{3}$$

is

$$\langle f_\sigma,c(j,T)\rangle=T\langle f_\sigma,c(j)\rangle, \tag{4}$$

so that the average collision frequency of the jth particle is

$$\nu_{j,\sigma}=(1/T)\langle f_\sigma,c(j,T)+c(j+1-N\delta_{jN},T)\rangle=\langle f_\sigma,c(j)+c(j+1-N\delta_{jN})\rangle. \tag{5}$$

We assume again that the particles are ordered in such a way that their number index increases from the left to the right. Then $f_\sigma=f^1_\sigma$, Eq. (117) of II, and

$$\langle f^1_\sigma,c(j)\rangle=\langle f^1_\sigma,c(j+1-N\delta_{jN})\rangle=\tfrac{1}{2}\nu^1_\sigma, \tag{6}$$

because of the cyclic symmetry of this distribution function [cf. the first of Eqs. (131) of II]. Equations (1) and (6) entail $\nu^1_\sigma=\nu^s_\sigma$ for $s\neq 1$ so that we drop the superscript indicating the order of the particles. Following the calculation of p_σ [II, Eqs. (189)–(191)] we find

$$\begin{aligned}\nu_\sigma&=2\int dV'\,\varphi_\sigma(V')\lim_{\xi\downarrow 0}q_{j,j-1+N\delta_{j1}}(d+\sigma_j[V']\xi-L\delta_{j1})\times|v_j-v_{j-1+N\delta_{j1}}|\\&=2\int dV'\,\varphi_\sigma(V')\theta(v_j-v_{j-1+N\delta_{j1}})\left(\frac{N-1}{L_d}\right)\times|v_j-v_{j-1+N\delta_{j1}}|\\&=\frac{N-1}{L_d}\int dV'\,\varphi_\sigma(V')|v_j-v_{j-1+N\delta_{j1}}|.\end{aligned} \tag{7}$$

For the smallest stationary ensemble with $\varphi_{\{v_0\}}$ given by Eq. (119) of II the integration is easily carried out to yield

$$\nu_{\{v_0\}}=\frac{1}{NL_d}\sum_{jk}|v_{j0}-v_{k0}|. \tag{8}$$

To obtain the collision frequency for the microcanonical ensemble the frequency (8) has to be averaged over all initial velocities V'_0 with $(V'_0)^2=v_E^2=2E/m$ [cf. II, Eqs. (120)–(125)]. To this end we introduce polar coordinates $v,\omega_1,\ldots,\omega_{N-2}$ for the vectors $V'=(\bar v_s)\in R'$ such that[3]

$$\begin{aligned}\bar v_1&=v\cos\omega_{N-2},\\ \bar v_2&=v\sin\omega_{N-2}\cos\omega_{N-3},\\ \bar v_3&=v\sin\omega_{N-2}\sin\omega_{N-3}\cos\omega_{N-4},\\ &\vdots\\ \bar v_{N-2}&=v\sin\omega_{N-w}\cdots\sin\omega_2\cos\omega_1,\\ \bar v_{N-1}&=v\sin\omega_{N-2}\cdots\sin\omega_2\sin\omega_1;\end{aligned} \tag{9}$$

$$0<\omega_1<2\pi,\quad 0<\omega_l<\pi,\quad \text{for } l=2,\ldots,N-2; \tag{10}$$

$$d\omega\,\mu(\omega)=\prod_{k=1}^{N-2}d\omega_k\sin^{k-1}\omega_k, \tag{11}$$

$$dV'=dv\,v^{N-2}\,d\omega\,\mu(\omega).$$

 0022-2488/85/102617-10$02.50

Without loss of generality the first of the Cartesian coordinates $\bar{v}_s$ may be chosen to be

$$\bar{v}_1 = (1/\sqrt{2})(v_j - v_{j-1+N\delta_{j1}}). \tag{12}$$

Inserting (12), (9), and (11) in (7) and using the integrals[4]

$$\int_0^\pi d\omega \sin^n \omega = \frac{\Gamma((n+1)/2)}{\Gamma((n+2)/2)}\sqrt{\pi}, \tag{13}$$

one finally obtains

$$\nu_E = \frac{2(N-1)}{L_d}\frac{\Gamma((N-1)/2)}{\Gamma(N/2)}\sqrt{\frac{E}{m\pi}}. \tag{14}$$

The collision frequency for the canonical ensemble can be obtained from (14) by integrating ν_E over E with the weight function $h_T(E)$ given in II, Eq. (129). The result is

$$\nu_T = [2(N-1)/L_d]\sqrt{kT/m\pi}. \tag{15}$$

It can also be obtained directly from (7), (12), and Eq. (126) of II, and be written as

$$\nu_T = \sqrt{\frac{\gamma}{\pi}}\int dv\, e^{-\gamma v^2}\nu_{N,\gamma}(v), \tag{16}$$

$$\nu_{N,\gamma}(v) = \sqrt{\frac{\gamma}{\pi}}\int dw\, e^{-\gamma w^2}|v-w| \times \rho_d\left(1-\frac{1}{N}\right)F_{N,\gamma}(v+w) \tag{17}$$

[cf. II, Eqs. (197)–(199)]. The thermodynamic limit of these equations is the collision frequency of Lebowitz and Percus.[5]

III. THE NO-COLLISION PROBABILITY

In this section we calculate the probability for a specified particle to move freely during the time interal $(0,t)$. This probability has already been termed "no-collision probability" and denoted by $W(t)$ in paper I. If the first particle is considered the corresponding probability $W_1(t)$ is the expectation value of the observable

$$\begin{aligned} &w(1,t;X,V) = w(1,t;\mathscr{P}X,\mathscr{P}V),\\ &X' \in S_1: \quad w(1,t;X,V) = \chi^{Q_1}(X' + D_1 + V't), \end{aligned} \tag{18}$$

where χ^{Q_1} is the characteristic function of the parallelepiped Q_1 defined in Sec. II E of II. To see that $\chi^{Q_1} = 1$ if the first particle did not collide within $(0,t)$, and $\chi^{Q_1} = 0$ if it did, one has to recall the evolution of the system as described in Sec. III of II. The ray $X' + D_1 + V't, t \geqslant 0$, describes the evolution in the extended configuration space, starting with $X_0 = X' \in S_1$ and $V_0 = V'$. A collision happens whenever the point of the ray passes from one simplex to a neighboring one and the reflection transforming the two adjacent simplices into each other indicates which pair of particles did just collide. If the first particle is not allowed to collide the admissible transformations are of the form $\mathscr{P}^{(1)}$ [II, Eq. (22)] and the region reached from $\bar{S}_1 + D_1$ by these reflections is just the parallelepiped $\bar{Q}_1$. These facts are easily visualized for $N = 3$, where $\bar{Q}_1$ consists of two triangles (see Sec. 4 of I, especially Figs. 7, 8, and 11). The situation for $N = 3$ suggests that, if the jth particle is considered instead of the first one, the appropriate observable is

$$\begin{aligned} &w(j,t;X,V) = w(j,t;\mathscr{P}X,\mathscr{P}V),\\ &X' \in S_1: \quad w(j,t;X,V) = \chi^{Q_j - U_j}(X' + D_1 + V't), \end{aligned} \tag{19}$$

where $Q_j = \mathscr{P}_3^{j-1}Q_1$ and

$$U_j = \sum_{k=1}^{j-1} A_k = \sum_{k=1}^{j-1} \mathscr{P}_3^{k-1}A_1 \quad \text{for } 2 \leqslant j \leqslant N, \tag{20}$$

i.e., $Q_j + U_j$ is obtained from Q_1 by $j-1$ applications of the transformation $X' \to \mathscr{P}_3 X' - A_1$. In fact this may be proven for arbitrary N using Eqs. (30) and (32) of II.

The no-collision probability is the expectation value of (19) for one of the stationary distributions defined in Sec. V of II:

$$W^s_{\sigma,j}(t) = \langle f^s_\sigma, w(j,t)\rangle. \tag{21}$$

Because of the first of Eqs. (19) and the relation between distribution functions belonging to different values s $W^s_{\sigma,j}$ is the same function for all values of s. Moreover the symmetry properties of f^1_σ [II, Eqs. (130) and (131)] and definition of $\bar{Q}_j + U_j$ entail $W^1_{\sigma,j} = W^1_{\sigma,1}$. We therefore write

$$W_\sigma(t) = W^s_{\sigma,j}(t) \tag{22}$$

and calculate $W^1_{\sigma,1}(t)$. Noting that

$$\varphi_\sigma(V') = \varphi_\sigma(\mathscr{P}V'), \tag{23}$$

$$G \subset R': \quad \chi^G(\mathscr{P}^{-1}[X' - Y']) = \chi^{\mathscr{P}G + Y'}(X'), \tag{24}$$

$$\chi^{Q_1}(\mathscr{P}^{(1)}X') = \chi^{Q_1}(X'), \tag{25}$$

we find

$$\begin{aligned} W_\sigma(t) &= W^1_{\sigma,1}(t)\\ &= \int dX'\,dV'|S|^{-1}\chi^{\bar{S}_1}(X')\\ &\quad\times\varphi_\sigma(V')\chi^{Q_1}(X' + D_1 + V't)\\ &= \int dV'\,\varphi_\sigma(V')|S|^{-1}\\ &\quad\times\int dX'\,\chi^{\bar{S}_1 + D_1}(X')\chi^{Q_1}(X' + V't)\\ &= \int dV'\,\varphi_\sigma(V')|S|^{-1}\frac{1}{(N-1)!}\\ &\quad\times\sum_{\mathscr{P}^{(1)}}\int dX'\,\chi^{\bar{S}_1 + D_1}(\mathscr{P}^{(1)}X')\\ &\quad\times\chi^{Q_1}(\mathscr{P}^{(1)}[X' + V't])\\ &= \int dV'\,\varphi_\sigma(V')|Q|^{-1}\\ &\quad\times\int dX'\,\chi^{Q_1}(X')\chi^{Q_1 - V't}(X')\\ &= \int dV'\,\varphi_\sigma(V')|Q|^{-1}|(Q_1 + V't)\cap Q_1|. \end{aligned} \tag{26}$$

For fixed V' the volume of the intersection of $Q_1 + V't$ and Q_1 decreases monotonically from $|Q|$ to 0 as $|t|$ increases. Since φ_σ is a normalized measure, $W_\sigma(t)$ therefore decreases monotonically from 1 to 0 for increasing t (>0). To compute $|(Q_1 + V't)\cap Q_1|$ we introduce suitable (nonorthogonal) coordinates in R':

$$\bar{z}_j = x_{j+1} - x_1, \quad j = 1,\ldots,N-1, \tag{27}$$

$$\left|\frac{\partial \bar{z}_j}{\partial y_k}\right| = \sqrt{N}. \tag{28}$$

Then

$$\bar{Q}_1 + Y' = \{X \,|\, X'' = 0;\ 0 \leqslant x_{j+1} - x_1 - y_{j+1} + y_1 \leqslant L_d\} \tag{29}$$

and

$$|(Q_1 + V't) \cap Q_1| = \frac{1}{\sqrt{N}} \prod_{j>1} (L_d - |v_j - v_1|t)_+, \tag{30}$$

where

$$a_+ = \tfrac{1}{2}(a + |a|). \tag{31}$$

Inserting (30) into (26) one obtains for a smallest stationary ensemble

$$W_{\{v_{j0}\}}(t) = \frac{1}{NL_d^{N-1}} \sum_j \prod_{k \neq j} (L_d - |v_{k0} - v_{j0}|t)_+. \tag{32}$$

The $W_{\{v_{j0}\}}$ vanishes for $t \geqslant t_{\max}$ where $t_{\max} = \max t_j$ and $t_j = L_d$ (max of $|v_{k0} - v_{j0}|$ for fixed j)$^{-1}$. For $t \in (0, t_{\max})$ $W_{\{v_{j0}\}}(t)$ is a polynomial of degree $N-1$, but the form of this polynomial changes whenever t transcends one of the t_j's. The first polynomial, representing $W_{\{v_{j0}\}}$ for $0 \leqslant t \leqslant \min t_j$ is obtained from (32) putting $a_+ = a(>0)$ everywhere. Collecting the terms according to the powers of t, one finds

$$W_{\{v_{j0}\}}(t) = 1 - \nu_{\{v_{j0}\}} t + O(t^2), \tag{33}$$

in accordance with the interpretation of W and ν.

The no-collision probability W_E is obtained from $W_{\{v_{j0}\}}$ by averaging over all initial velocities with $|V_0| = |V_0'| = v_E$. This is practicable only for sufficiently small times where the functions $W_{\{v_{j0}\}}$ are all represented by their first polyomial. This is the case if

$$t \leqslant L_d / v_E \sqrt{2} = (L_d/2)\sqrt{m/E}, \tag{34}$$

because $(1/\sqrt{2})|v_{j0} - v_{k0}| \leqslant |V_0'| = v_E$. If we use the notation of II, Eqs. (122)–(124), note that

$$\int d\omega\, \mu(\omega) g(E(\omega)) = \int d\omega\, \mu(\omega) g(\mathscr{P} E(\omega)), \tag{35}$$

and define the unit vectors $D_{jk} \in R'$ by

$$D_{jk} = (1/\sqrt{2})(\delta_{ij} - \delta_{ik}), \tag{36}$$

we have

$$\begin{aligned} W_E(t) &= |K|^{-1} \int d\omega\, \mu(\omega) W_{\{v_E e_j(\omega)\}}(t) \\ &= |K|^{-1} \int d\omega\, \mu(\omega) L_d^{1-N} \\ &\quad \times \prod_{j>1} (L_d - |D_{j1} \cdot E(\omega)| \sqrt{2} v_E t) \\ &= 1 - (\sqrt{2} v_E t / L_d) |K|^{-1} \\ &\quad \times \int d\omega\, \mu(\omega) \sum_{j>1} |D_{j1} \cdot E(\omega)| + \cdots \\ &\quad + (-\sqrt{2} v_E t / L_d)^{N-1} |K|^{-1} \\ &\quad \times \int d\omega\, \mu(\omega) \prod_{j>1} |D_{j1} \cdot E(\omega)| \\ &= 1 + \sum_{n=1}^{N-1} \left(-\frac{\sqrt{2} v_E t}{L_d}\right)^n \binom{N-1}{n} |K|^{-1} \\ &\quad \times \int d\omega\, \mu(\omega) \prod_{j=2}^{n+1} |D_{j1} \cdot E(\omega)|. \end{aligned} \tag{37}$$

To calculate the integrals appearing in (37) it is advantageous to rotate the coordinate system in such a way that the vector D_{j1} is transformed into a vector C_{j-1}, the components of which vanish for $s \leqslant N - j$:

$$\begin{aligned} &C_j = (\bar{c}_{js}) = (\bar{c}_{j1}, \ldots, \bar{c}_{j(N-1)}) \in R', \\ &\bar{c}_{js} = 0, \quad \text{for } s \leqslant N - 1 - j. \end{aligned} \tag{38}$$

This convention and the relations

$$\begin{aligned} C_j \cdot C_k &= D_{(j+1)1} \cdot D_{(k+1)1} \\ &= \tfrac{1}{2}(1 + \delta_{jk}), \quad j,k = 1, \ldots, N-1, \end{aligned} \tag{39}$$

fix the vectors C_j up to a common sign:

$$\bar{c}_{js} = \begin{cases} 0 & \text{for } s < N - j, \\ [(j+1)/2j]^{1/2} & \text{for } s = N - j, \\ [2(N-s)(N-s+1)]^{-1/2} & \text{for } s > N - j. \end{cases} \tag{40}$$

Now let the unit vector $E(\omega) \subset K$ be parametrized in the same way as the vector $V' \in R'$ in Eq. (9) before. The components of interest are then

$$\begin{aligned} \bar{e}_s(\omega) &= e^j_{1-N+j+s}(\omega_1, \ldots, \omega_{j-1}) \prod_{l=j}^{N-1} \sin \omega_l \\ &\quad \text{for } N - j \leqslant s \leqslant N - 1, \end{aligned} \tag{41}$$

the quantities e^j being the components of a j-dimensional unit vector [cf. Eq. (9)], and

$$\begin{aligned} &e^1_1 = 1; \quad e^2_1(\omega_1) = \cos \omega_1, \quad e^2_2(\omega_1) = \sin \omega_1; \\ &e^3_1(\omega_1, \omega_2) = \cos \omega_1, \quad e^3_2(\omega_1, \omega_2) = \sin \omega_1 \cos \omega_2, \\ &e^3_3(\omega_1, \omega_2) = \sin \omega_1 \sin \omega_2; \quad \text{etc.} \end{aligned} \tag{42}$$

Using (10), (11), (13) one finds for $n = 1$,

$$\int d\omega\, \mu(\omega) |C_1 \cdot E(\omega)| = \frac{2\pi^{(N-2)/2}}{\Gamma(N/2)}, \tag{43}$$

and for $n \geqslant 2$,

$$\begin{aligned} &\int d\omega\, \mu(\omega) \prod_{j=1}^{n} |C_j \cdot E(\omega)| \\ &\quad = \pi^{(N-1-n)/2} \frac{(n-1)!}{\Gamma((N-1+n)/2)} 2n H_n, \\ &H_n = \frac{1}{2n} \int d\omega^n\, \mu(\omega^n) F_n(\omega^n), \end{aligned} \tag{44}$$

where $\omega^n = (\omega_1, \ldots, \omega_{n-1})$ are the angular variables of n-dimensional polar coordinates [cf. (9), (10)] and

$$\begin{aligned} F_n(\omega^n) = \prod_{j=1}^{n} \Bigg| &\sum_{q=1}^{j} \bar{c}_{j,N-1-j+q} e^j_q(\omega_1, \ldots, \omega_{j-1}) \\ &\times \prod_{l=j}^{n-1} \sin \omega_l \Bigg|. \end{aligned} \tag{45}$$

The intregal H_n can be calculated analytically for $n = 2$, while for $n > 2$ this has to be done numerically:

$$H_2 = \sqrt{3}/4 + \pi/24 \approx 0.56, \quad H_3 \approx 0.25, \quad \text{etc.} \tag{46}$$

Combining these results we finally arrive at

$$W_E(t) = 1 - \nu_E t + O(t^2)$$
$$= \sum_{n=0}^{N-1} \alpha_n (-\nu_E t)^n, \quad \text{for } \nu_E t \leqslant \frac{2\Gamma((N+1)/2)}{\sqrt{\pi}\Gamma(N/2)}, \tag{47}$$

with

$$\alpha_n = \frac{(N-1)!\Gamma((N-1)/2)}{(N-1-n)!\Gamma((N-1+n)/2)} \left[\frac{\Gamma(N/2)}{2\Gamma((N+1)/2)}\right]^n H_n,$$
$$H_0 = H_1 = 1. \tag{48}$$

We did not attempt to calculate $W_E(t)$ for $\nu_E t > 2\Gamma((N+1)/2)/[\sqrt{\pi}\Gamma(N/2)]$; we only recall that W_E is a monotonically decreasing function. A short reflection shows that

$$W_E(t) = 0, \quad \text{for } \nu_E t \geqslant \frac{4\Gamma((N+1)/2)}{\Gamma(N/2)} \left(\frac{N-1-\epsilon_N}{2}\right)^{1/2},$$
$$\epsilon_N = (1/2N)[1 + (-1)^N], \tag{49}$$

because the largest diameter of Q_j is $L_d(N - 1 - \epsilon_N)$.

In the thermodynamic limit $\alpha_n = H_n$ and (47) becomes

$$W_{\infty,\gamma}(t) = \sum_{n=0}^{\infty} H_n (-\nu_{\infty,\gamma} t)^n, \tag{50}$$

with $\gamma = m/(4\epsilon)$ and $\nu_{\infty,\gamma}$ being the limit of (14) for $L_d = N/\rho_d, E = \epsilon N, N \uparrow \infty$ [cf. I, Eq. (131)]. Equation (50) is valid for all finite t since the domain of the first polynomial increases as $\sqrt{N}$.

It is not possible to obtain the no-collision probability W_T by integrating over the polynomial (47) or the first polynomial $W_{\{v_0\}}$ because the domains of these polynomials depend on $|V'|$ which varies within the canonical ensemble. It is, however, possible to derive upper and lower bounds for W_T which approach each other for $t \downarrow 0$ and in the thermodynamic limit. For the lower bound note that $W_E(t) \geqslant 0$; hence

$$\overline{W}_T(t) = \int_0^{E_2} dE\, h_T(E) W_E(t) \tag{51}$$

is a lower bound for W_T for every $E_2 > 0$. For the upper bound note that $|(Q_1 + V't) \cap Q_1| < |(Q_1 + W't) \cap Q_1|$ for $|V'| > |W'|$, whence $W_{\{v_0\}}(t) < W_{\{w_0\}}(t)$ for $|V'| > |W'|$ and $W_E(t) < W_{E'}(t)$ for $E > E'$. Furthermore integral calculus tells us that

$$W_T(t) = \int_0^{\infty} dE\, h_T(E) W_E(t)$$
$$= \int_0^{E_2} dE\, h_T(E) W_E(t) + \int_{E_2}^{\infty} dE\, h_T(E) W_E(t)$$
$$= W_{E_1}(t) \int_0^{E_2} dE\, h_T(E) + W_{E_3}(t) \int_{E_2}^{\infty} dE\, h_T(E), \tag{52}$$

with $0 < E_1 < E_2 < E_3 < \infty$. Therefore $W_{E_3} < W_{E_1}$ and

$$W_{E_1}(t) = \frac{\int_0^{E_2} dE\, h_T(E) W_E(t)}{\int_0^{E_2} dE\, h_T(E)} \tag{53}$$

is an upper bound of W_T. If the denominator of (53) is expressed in terms of gamma functions[6] the bounds of W_T are

$$\overline{W}_T(t) \leqslant W_T(t) \leqslant \frac{\Gamma((N-1)/2)}{\gamma((N-1)/2, E_2/kT)} \overline{W}_T(t). \tag{54}$$

Up to now E_2 has been an arbitrary positive number. We now choose

$$E_2 = mL_d^2/4t^2 = mN^2/4\rho_d^2 t^2, \tag{55}$$

so that $W_E(t)$ in (51) can be represented by (47). For fixed t the ratio before $\overline{W}_T$ in (54) tends to 1. This follows from $\gamma(a,x) < \gamma(a,y) < \Gamma(a)$, for $0 < x < y < \infty$, $E_2 = bN^2/4 > bM^2$, where $M = (N-1)/2$, and

$$\gamma(M, bM^2)/\Gamma(M) = 1 - O(b^M M^M \exp[-bM^2 + M]). \tag{56}$$

The thermodynamic limit of W_T can therefore be obtained by first integrating each term of (51), multiplied with the weight function $h_T(E)$, from 0 to E_2 and then performing the limit $N \uparrow \infty$. The result has exactly the form (50) except that now $\gamma = m/(2kT)$.

Lebowitz and Percus[5] found an integral representation of the thermodynamic limit of W_T:

$$W_{\infty,\gamma}(t) = \sqrt{\frac{\gamma}{\pi}} \int dv \exp(-\gamma v^2) W_{\infty,\gamma}(v,t), \tag{57}$$
$$W_{\infty,\gamma}(v,t) = \exp[-\nu_\gamma(v)t], \tag{58}$$
$$\nu_\gamma(v) = \lim_{N \uparrow \infty} \nu_{N,\gamma}(v) = \sqrt{\frac{\gamma}{\pi}} \int dw \exp(-\gamma w^2) |v - w|_{\rho_d}$$
$$= \rho_d [v \operatorname{erf}(\sqrt{\gamma} v) + (1/\sqrt{\pi\gamma}) \exp(-\gamma v^2)]. \tag{59}$$

This integral representation can also be obtained from (32) if all particles are assumed to have Maxwellian velocity distributions and the thermodynamic limit is performed. Expanding the exponential (58) and integrating each term in (57), one obtains the power series

$$W_{\infty,\gamma}(t) = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \sqrt{\frac{\gamma}{\pi}} \int dv \exp(-\gamma v^2) [\nu_\gamma(v)]^n. \tag{60}$$

Comparing (60) with (50) we see that

$$\frac{1}{n!} \sqrt{\frac{\gamma}{\pi}} \int dv \exp(-\gamma v^2) [\nu_\gamma(v)]^n$$
$$= \frac{1}{n!} \sqrt{\frac{\gamma}{\pi}} \left(\rho_d \sqrt{\frac{\gamma}{\pi}}\right)^n \int dv\, dw_1 \cdots dw_n$$
$$\times |v - w_1| \cdots |v - w_n| \exp[-(v^2 + w_1^2 + \cdots + w_n^2)]$$
$$= \nu_{\infty,\gamma}^n H_n. \tag{61}$$

This identity may be checked by introducing polar coordinates in the integration domain R^{n+1}. Instead of the original definition in terms of angular variables, Eqs. (40)–(45), Eqs. (61) and (59) with $\rho_d = \gamma = 1$ may be used to calculate numerically the integrals H_n, which are also needed for the first polynomial of W_E if the system is finite [cf. I, Eqs. (100)–(103)].

IV. THE VELOCITY AUTOCORRELATION FUNCTION

A. Definition of the function

The velocity autocorrelation function (VAF) is, up to a scaling factor, the expectation value of the observable

$$g(t;X,V) = \sum_j g(j,t;X,V), \tag{62}$$

$$g(j,t;X,V) = g(j,t;\mathscr{P}X,\mathscr{P}V), \tag{63}$$

$$X' \in S_1: \quad g(j,t;X,V) = v_j v_{jt}[X,V], \quad \text{if } V_t[X,V] \text{ is defined,}$$
$$= 0, \quad \text{if } V_t[X,V] \text{ is undefined.} \tag{64}$$

The subset of the phase space K where $V_t[X,V]$ is undefined for a given instant t has dimension $2N-1$ and may be neglected for the distribution functions considered here. It follows from the evolution of the system as discussed in Sec. III of II that $v_{jt}[X,V]$ may be expressed as follows:

$$X' \in S_1: \quad v_{jt}[X,V] = \sum_{i,B} v_i \chi^{Q_i^* + U_j + B}(X' + D_1 + V't), \tag{65}$$

$$Q_i^* = Q_i \setminus (Q_i \cap E). \tag{66}$$

The vectors U_j have already been defined in (20).

It follows from (62), (63), and $f_\sigma^s(\mathscr{P}^{-1}X,\mathscr{P}^{-1}V) = f_\sigma^{Ps}(X,V)$ that

$$\langle f_\sigma^s, g(t)\rangle = \langle f_\sigma^1, g(t)\rangle. \tag{67}$$

Moreover the symmetry properties of f_σ^1, Eqs. (131) and (132) of II, and the transformation properties

$$g(j,t;\mathscr{P}_3 X - (L/L_d)A_1, \mathscr{P}_3 V) = g(j-1+N\delta_{j1};X,V), \tag{68}$$

resulting from Eqs. (81) and (11) of II, entail

$$\langle f_\sigma^1, g(j,t)\rangle = \langle f_\sigma^1, g(1,t)\rangle. \tag{69}$$

These relations allow us to define the VAF in several equivalent ways:

$$\psi_\sigma(t) = \frac{\langle f_\sigma^s, g(t)\rangle}{\langle f_\sigma^s, g(0)\rangle} = \frac{\langle f_\sigma^1, g(1,t)\rangle}{\langle f_\sigma^1, g(1,0)\rangle}. \tag{70}$$

The denominators of (70) are

$$\langle f_\sigma^s, g(0)\rangle = N\langle f_\sigma^1, g(1,0)\rangle = \int dV' \varphi_0(V') V'^2. \tag{71}$$

For the calculation of the numerator note that $g(1,t;\mathscr{P}^{(1)}X,V) = g(1,t;X,\mathscr{P}^{(1)-1}V)$ because of (63). Recalling the definitions of $\mathscr{P}^{(1)}$ and $\overline{Q}_1$, Eqs. (22) and (23) of II, one therefore has

$$\begin{aligned}\langle f_\sigma^1, g(1,t)\rangle &= \frac{1}{(N-1)!}\sum_r^{(1)} \langle f_\sigma^r, g(1,t)\rangle \\ &= \int dX'\, dV' |Q|^{-1} \chi^{\overline{Q}_1}(X') \varphi_\sigma(V') \\ &\quad \times \sum_{j,B} v_1' v_j' \chi^{Q_j^* + B}(X' + V't) \\ &= \sum_j \int dV' \varphi_\sigma(V') v_1' v_j' |Q|^{-1} \\ &\quad \times \sum_B |(Q_1 + V't) \cap (Q_j + B)|. \end{aligned} \tag{72}$$

For the following it will be useful to transcribe (72). If V' is replaced by $\mathscr{P}_3^{1-i}V'$ in (72) then

$$\langle f_\sigma^1, g(1,t)\rangle = \sum_j \int dV' \varphi_\sigma(V') v'_{i(N)} v'_{j+1-1(N)} |Q|^{-1} \times \sum_B |(Q_{i(N)} + V't) \cap (Q_{j+i(N)} + B)|, \tag{73}$$

$$n(N) = n \text{ (modulo } N) \in (1,\ldots,N), \tag{74}$$

because $(\mathscr{P}_3^{1-i}V')_j = v_{j+i-1(N)}$, $\mathscr{P}_3^{1-i}Q_j = Q_{j+i-1(N)}$, and $|\mathscr{P}G| = |G|$. Summing (73) over all values of i one finally obtains

$$\langle f_\sigma^1, g(t)\rangle = \int dV' \varphi_\sigma(V') G(t,V'), \tag{75}$$

$$G(t,V') = \sum_{i,j} v_i' v_j' |Q|^{-1} \sum_B |(Q_i + V't) \cap (Q_j + B)|, \tag{76}$$

which is the starting point for all further calculations.

B. Series expansion for short times

If we take into account that $\cup_{i,B}(\overline{Q}_i + B) = R'$, and hence

$$\sum_B |(Q_i + V't) \cap (Q_j + B)| = |Q| - \sum_{B,\,\overline{A} \neq i)} |(Q_i + V't) \cap (Q_j + B)|, \tag{77}$$

we may write (76) as

$$G(t,V') = \sum_i v_i'^2 + \sum_{i \neq j} (v_i' v_j' - v_i'^2)|Q|^{-1} \times \sum_B |(Q_i + V't) \cap (Q_j + B)|. \tag{78}$$

The task is now to compute the volumes of the various intersections appearing in (78). If t is sufficiently small, i.e.,

$$|V'|t < L_d/\sqrt{2}, \tag{79}$$

then only those parallelepipeds $Q_j + B$ will have a possibly nonempty intersection with a given $Q_i + V't$, which are adjacent to Q_i. The common boundaries of $\overline{Q}_i$ and $\overline{Q}_j + B$ with $\overline{Q}_i \cap (\overline{Q}_j + B) \neq \varnothing$ are again parallelepipeds spanned by vectors of length $|A_i| = L_d\sqrt{(N-1)/N}$ but of dimension $N-1-n$, $n = 1,\ldots,N-1$. These dimensions induce a natural classification scheme. If i and j are fixed there exists only a finite number of B's such that $\overline{Q}_i \cap (\overline{Q}_j + B)$ is nonempty. The various possibilities may be analyzed noting that $X' \in \overline{Q}_i$ if, and only if, $X' = \Sigma_k \alpha_k A_k$ with $0 \leqslant \alpha_k \leqslant 1$ and $\alpha_i = 0$. For X' to belong to the boundary of $\overline{Q}_i$ one or more of the α_k's with $k \neq i$ have to take the value $\alpha_k = 0$ or $\alpha_k = 1$. The result of these geometrical considerations is summed up in Table I. If now Q_i is shifted by $V't$, it penetrates either Q_j or $Q_j + B_{ij}$ with which it has common surfaces of dimension $N-2$. If $v_i > v_j$ the nonempty intersection is $(Q_i + V't) \cap Q_j$, which is contained in the intersection of the sheet $\{X' | 0 < x_i$

TABLE I. Intersection of parallelepipeds Q_i and $Q_j + B$.

class	(1)	(2)	(3)	etc.
dimension	$N-2$	$N-3$	$N-4$	
vectors B	$0, B_{ij}$	B_{jk}, B_{ki} $i \neq k \neq j$	$B_{kl}, B_{ji} + B_{kl}$ $i \neq k \neq j$, $i \neq l \neq j$, $k \neq l$	
possibilities	2	$2(N-2)$	$2(N-2)(N-3)$	

$-x_j<(v_i-v_j)t\}$ and Q_j; likewise if $v_i<v_j$ then $(Q_i+V't)\cap(Q_j+B_{ji})\subset\{X'|L_d<x_j-x_i<L_d+(v_j-v_i)t\}\cap(Q_j+B_{ij})$. Therefore in any case $|(Q_i+V't)\cap(Q_j+B)|=O(t^1)$ if the pair i,j belongs to class (1), and, generalizing this argument, one sees that $|(Q_i+V't)\cap(Q_j+B)|=O(t^n)$ if i,j belongs to class (n). Thus the sum over B is split according to

$$\sum_B=\sum_B^{(1)}+\sum_B^{(2)}+\cdots+\sum_B^{(N-1)}. \tag{80}$$

The partial sums $\Sigma^{(n')}$ with $n'>n$ all may be neglected when the coefficient of t^n is computed. To compute the value of this coefficient we have to compute, however, all contributions of the classes $n'\leqslant n$ up to the order t^n.

To see how this can be done in a systematic way, consider, for instance, class (1) and $(Q_i+V't)\cap Q_j=J$. Let the domains K and L_k, $k\neq i,j$, be defined by $K=Q_j\cap\{X'|0<x_i-x_j<(v_i-v_j)t\}$ and $L_k\{X'|x_k-x_i-(v_k-v_i)t\leqslant 0$ or $\geqslant L_d\}$. Then

$$|J|=|K|-\sum_{(k)}|K\cap L_k|+\sum_{(k,k')}|K\cap L_k\cap L_{k'}|$$
$$-\sum_{(k,k',k'')}|K\cap L_k\cap L_{k'}\cap L_{k''}|+\cdots,$$

where $\Sigma_{(k,k',\ldots)}$ means that terms where two or more indices coincide have to be excluded. The first term is the volume of a parallelepiped containing J. The remaining terms allow for the difference between the flat parallelepiped K and the disklike domain J along the narrow parts of the surfaces. The second term in the sum counts all points inside K outside of a sheet L_k formed by two parallel planes contributing to the surface of $Q_i+V't$. Since these sheets are considered independently of each other there is an overcounting of points which lie outside of two or more sheets. For points lying outside of only two sheets this error is removed by the third term in the sum. It is not difficult to verify that if the whole sum (ending with N-fold intersections) is performed, each interior point of J is counted exactly once while the contributions of points $X'\in K\setminus J$ vanish. The various volumes appearing in the sum may be explicitly calculated if suitable nonorthogonal coordinates are introduced [cf. Eq. (27)]. The advantage of this iterative method to compute $|J|$ lies in the fact that $|K|=O(t^1)$, $|K\cap L_k|=O(t^2)$, $|K\cap L_k\cap L_{k'}|=O(t^3)$, etc., so that it is clear how many terms have to be taken into account for a given order of t. This convenient feature is also present in those modifications of the method which apply for intersections belonging to higher classes.

This outline of the method clearly shows that the effort to compute the coefficient of t^n increases drastically with the order n. Even for $n=2$ the calculations become so lengthy that we refrain from reproducing them here and only state their result:

$$\sum_B^{(1)}|(Q_i+V't)\cap Q_j|$$
$$=t\frac{L_d^{N-2}}{\sqrt{N}}-\frac{t^2}{2}\frac{L_d^{N-3}}{\sqrt{N}}\sum_k\left[|v_i'-v_k'|\,|v_i'-v_j'|\right.$$
$$+|v_i'-v_j'|\,|v_k'-v_j'|-2(v_i'-v_k')_+(v_k'-v_j')_+$$
$$\left.-2(v_k'-v_i')_+(v_j'-v_k')_+\right]+O(t^3),$$

$$\sum_B^{(2)}|(Q_i+V't)\cap Q_j|$$
$$=\frac{t^2}{4}\frac{L_d^{N-3}}{\sqrt{N}}\sum_k\left[(v_i'-v_k')_+(v_j'-v_k')_+\right.$$
$$-(v_i'-v_k')_+(v_j'-v_i')_+$$
$$-(v_i'-v_j')_+(v_j'-v_k')_++(v_k'-v_i')_+(v_k'-v_j')_+$$
$$-(v_k'-v_i')_+(v_i'-v_j')_+$$
$$\left.-(v_j'-v_i')_+(v_k'-v_j')_+\right|+O(t^3), \tag{81}$$

$$\sum_B^{(n)}|(Q_i+V't)\cap Q_j|=O(t^n),\quad \text{for } n=3,\ldots,N-1.$$

Inserting this into (78) we obtain, after some algebraic manipulations,

$$G(t,V')=\sum_i v_i'^2-\frac{t}{2L_d}\sum_{i,j}|v_i'-v_j'|^3$$
$$+\frac{t^2}{4L_d^2}\sum_{i,j,k}\left[3(v_i'-v_j')|v_i'-v_j'|\right.$$
$$\times(v_i'-v_k')|v_i'-v_k'|$$
$$\left.-(v_i'-v_j')^2(v_i'-v_k')^2\right]+O(t^3)$$
$$\text{for } t\,|V'|<L_d/\sqrt{2}. \tag{82}$$

If this is averaged with $\varphi_{\{v_0\}}$ one obtains the leading terms of the first polynomial representing $\langle f^1_{\{v_0\}},g(t)\rangle$:

$$\langle f^1_{\{v_0\}},g(t)\rangle=G(t,V_0'). \tag{83}$$

It is also possible to average (82) over φ_E since the time up to which the first polynomial represents $V\cdot V_t$ is the same for all members of the microcanonical ensemble. The normalization constant (71) is easily calculated in this case and the average of each term of (82) may be calculated analytically if polar coordinates are introduced for V' (cf. Sec. III):

$$\psi_E(t)=1-4\frac{N-1}{L_d}\frac{\Gamma((N-1)/2)}{\Gamma(N/2)}\sqrt{\frac{E}{m\pi}}\,t$$
$$+9\sqrt{3}\frac{(N-2)N}{(N+1)L_d^2}\frac{E}{m\pi}t^2+O(t^3)$$
$$=1-2\nu_E t+O(t^2),\quad \text{for } 4Et^2<mL_d^2, \tag{84}$$

the thermodynamic limit of (84) is

$$\psi_{\infty,\gamma}(t)=1-2\nu_{\infty,\gamma}t+(9\sqrt{3}/8)(\nu_{\infty,\gamma}t)^2+O(t^3), \tag{85}$$

with $\gamma=m/(4\epsilon)$, $\epsilon=\lim E/N$. By a different method, described in Sec. IV C below, the following series representation of $\psi_{\infty,\gamma}$ may be derived:

$$\psi_{\infty,\gamma}(t)=\sum_{n=0}^{\infty}I_n(-\nu_{\infty,\gamma}t)^n, \tag{86}$$

$$I_n=\frac{n+1}{\sqrt{\pi}}\int dv\,\exp(-v^2)$$
$$\times\sum_l\frac{[f(v)]^{n-2l}}{(n-2l)!}\left[\frac{g^2(v)}{4}\right]^l\frac{1}{(l!)^2}, \tag{87}$$

$$g^2(v)=f^2(v)-(\pi/2)v^2, \tag{88}$$

$$f(v)=\frac{1}{\sqrt{2}}\int dw\,\exp(-w^2)|v-w|$$

$$= \sqrt{\frac{\pi}{2}}\,[v\,\mathrm{erf}\,v + (1/\sqrt{\pi})\exp(-v^2)]. \tag{89}$$

Using tabulated integrals[4] the first three terms of (86) can be shown to coincide with those of (85). Equations (84) and (86) form a base from which one can speculate about the coefficient of t^n for finite N. We already know that the coefficient β_n of $(\nu_E t)^n$ has to satisfy three properties: (i) the limit of β_n for $N \uparrow \infty$ is I_n; (ii) the coefficients with $n \leqslant 3$ must be of such a form that (84) is obtained if ν_E is expressed by Eq. (14); and (iii) $\beta_n = 0$ if $n \geqslant N$ because the definition (76) shows that $V \cdot V_t$ is represented by polynomials of degree $N-1$. Our guess of the first polynomial of ψ_E is

$$\psi_E(t) = 1 + \sum_{n=1}^{N-1} \beta_n(-\nu_E t)^n \quad \text{for } \nu_E t < \frac{2\Gamma((N+1)/2)}{\Gamma(N/2)}, \tag{90}$$

$$\beta_n = \frac{N!}{2(N-1-n)!}\,\frac{\Gamma((N-1)/2)}{\Gamma((N+1+n)/2)} \times \left[\frac{\Gamma(N/2)}{2\Gamma((N+1)/2)}\right]^n I_n. \tag{91}$$

The first polynomial calculated according to Eqs. (90) and (91) has been compared with the function ψ_E obtained by a different method [see Eq. (105) below]. Since we found complete coincidence for $N \leqslant 6$ we believe that (91) holds true for arbitrary N.

The limited validity of the first polynomial prevents $\langle f_T^1, g(t)\rangle$ from being obtained from $\langle f_E^1, g(t)\rangle$ by integration. It is, however, possible to obtain bounds of $\langle f_T^1, g(t)\rangle$ similar as it has been done before for the no-collision probablity (cf. Sec. III). To this end note that $|V \cdot V_t| \leqslant V^2$ implies $|\psi_\sigma(t)| \leqslant 1$. Since the normalization constants (71) are $\langle f_E^1, g(0)\rangle = 2E/m$ and $\langle f_T^1, g(0)\rangle = (N-1)kT/m$, respectively, the following inequality holds true:

$$\left|(N-1)\frac{kT}{2}\psi_T(t) - \int_0^{E_2} dE\, h_T(E) E \psi_E(T)\right| \leqslant \int_{E_2}^{\infty} dE\, h_T(E) E = (N-1)\frac{kT}{2}\left[1 - \frac{\gamma((N+1)/2, E_2/kT)}{\Gamma((N+1)/2)}\right]. \tag{92}$$

The right-hand side vanishes in the thermodynamic limit if E_2 is chosen as in Eq. (55) (cf. Sec. III). Therefore ψ_T is also given by (86) in this limit but $\nu_{\infty,\gamma}$ is now the limit of the collision frequency (15).

C. Series expansion for arbitrary times

If we want to know the VAF for arbitrary times we have to start afresh from Eq. (76). We write

$$\sum_B |(Q_i + V't) \cap (Q_j + B)| = \int dX'\, \chi^{Q_i + V't}(X')\tilde{\chi}^j(X'), \tag{93}$$

$$\tilde{\chi}^j(X') = \sum_B \chi^{Q_j + B}(X'), \tag{94}$$

and emphasize that $\tilde{\chi}^j$ is a periodic function, the periods being given by the lattice $\{B\}$. Therefore, $\tilde{\chi}^j$ may be represented by the (Fourier) series

$$\tilde{\chi}^j(X') = \sum_A \chi^j_A \exp(i\alpha A \cdot X'), \tag{95}$$

with $A \in \{A\}$ (see Sec. II C of II),

$$\alpha = 2\pi/L_d^2, \tag{96}$$

and (Fourier) coefficients

$$\chi^j_A = |C|^{-1} \int_C dX'\, \tilde{\chi}^j(X')\exp(-i\alpha A \cdot X') = |C|^{-1} \int_{Q_j} dX' \exp(-i\alpha A \cdot X'). \tag{97}$$

The integrals (96) are easily calculated if coordinates of the form (27) are used:

$$\chi^j_A = N^{-1}\beta^{1-N}\exp(-i\beta a_j)\prod_{k\neq j} a_k^{-1}\sin\beta a_k, \quad \text{for } A = (a_l), \tag{98}$$

$$\beta = \pi/L_d. \tag{99}$$

If $\tilde{\chi}^j$ is represented by the series (95), Eq. (93) becomes

$$\sum_B |(Q_i + V't) \cap (Q_j + B)| = \sum_A \chi^j_A \int_{Q_i} dX' \exp[i\alpha A \cdot (X' + V't)] = |C| \sum_A \chi^i_A * \chi^j_A \exp(i\alpha A \cdot V't). \tag{100}$$

This leads us to the series representation

$$G(t,V') = N \sum_{\substack{A\neq 0 \\ i\neq j}} \chi^i_A * \chi^j_A \exp(i\alpha A \cdot V't)(v_i' v_j' - v_i'^2), \tag{101}$$

which follows from (78), (100), $\chi^k_0 = 1/N$, and $\Sigma_i v_i' = 0$. Equation (101) shows explicitly that the VAF (83) is quasi-periodic since every V_0 may be approximated with any desired accuracy by a multiple of a vector $B \in \{B\}$ (cf. Sec. VI of II).

The products $\chi^i_a * \chi^j_A$ appearing in (101) can be written as

$$\chi^i_A * \chi^j_A = \chi^i_A \chi^j_A * = N^{-2}\beta^{2-2N}(\sin^2\beta a_1)^{N-1} a_i a_j \prod_{k=1}^{N} a_k^{-2}. \tag{102}$$

In (102) we have used the fact that $\sin^2\beta a_i = \sin^2\beta a_1$ because the definition of the vectors A implies $\beta(a_i - a_j) =$ integer multiple of π (see Sec. II C of II). Inserting (102) into (101) one obtains

$$G(t,V') = N^{-1}\beta^{2-2N} \sum_{A\neq 0} (\sin^2\beta a_1)^{N-1} \times \left(\sum_{k=1}^{N} a_k^{-2}\right)(A \cdot V')\exp(i\alpha A \cdot V't) = -\frac{\partial^2}{\partial(\alpha t)^2} N^{-1}\beta^{2-2N} \sum_{A\neq 0} (\sin^2\beta a_1)^{N-1} \times \left(\sum_{k=1}^{N} a_k^{-2}\right)\exp(i\alpha A \cdot V't), \tag{103}$$

the last form being especially suited for the calculation of ψ_E and ψ_T. Averaging the exponential over the microcanonical ensemble yields integrals of the form (13) and integral representations of Bessel functions[4]:

$$|K|^{-1}\int d\omega\,\mu(\omega)\exp[i\alpha A\cdot v_E E'(\omega)]$$
$$=\frac{\Gamma((N-1)/2)}{\sqrt{\pi}\Gamma((N-2)/2)}\int_0^\pi d\vartheta\sin^{N-3}\vartheta$$
$$\times\exp(i\alpha|A|v_E t\cos\vartheta)$$
$$=2^{(N-3)/2}\Gamma((N-1)/2)(\alpha|A|v_E t)^{(3-N)/2}$$
$$\times J_{(N-3)/2}(\alpha|A|v_E t). \tag{104}$$

The derivatives of the Bessel functions can be transformed by means of the well known recursion formulas. We have

$$\psi_E(t)=N^{-1}\beta^{2-2N}2^{(N-3)/2}\Gamma((N-1)/2)$$
$$\times\sum_{A\neq 0}(\sin^2\beta a_1)^{N-1}\left(\prod_{k=1}^N a_k^{-2}\right)|A|^2$$
$$\times[(z_{E,A}t)^{(3-N)/2}J_{(N-3)/2}(z_{E,A}t)$$
$$-(N-2)(z_{E,A}t)^{(1-N)/2}J_{(N-1)/2}(z_{E,A}t)|, \tag{105}$$
$$z_{E,A}=\alpha|A|v_E=(2\pi|A|)/L_d^2)\sqrt{2E/m}. \tag{106}$$

The vectors $A=(a_i)$ may be parametrized by $N-1$ integers $n_1,\ldots,n_{N-1}$ such that $a_i=L_d(n_i-n/N)$, $n_N=0$, $n=\Sigma_i n_i$. The resulting formula can be found in I [Eqs. (116)–(121)], where the form of the function ψ_E and its asymptotic decay have been discussed.

The VAF for the canonical ensemble can be obtained by averaging $\langle f_E^1, g(t)\rangle=(2E/m)\psi_E(t)$ with the weight function $h_T(E)$, Eq. (129) of II, or simply by averaging (103) with $\varphi_T(V')=(\gamma/\pi)^{(N-1)/2}\exp(-\gamma V'^2)$, $\gamma=m/(2kT)$,

$$\psi_T(t)=[(N-1)N]^{-1}\beta^{2-2N}$$
$$\times\sum_{A\neq 0}(\sin^2\beta a_1)^{N-1}\left(\prod_{k=1}^N a_k^{-2}\right)A^2$$
$$\times(1-z_{T,A}^2t^2)\exp(-\tfrac{1}{2}z_{T,A}^2t^2), \tag{107}$$

$$z_{T,A}=\alpha|A|/\sqrt{2\gamma}=(2\pi/L_d^2)|A|\sqrt{m/kT}. \tag{108}$$

Since $A^2=a_1^2+\cdots+a_N^2$ and $z_{T,A}$ is a symmetric function of the a_i's, one can replace the factor $A^2a_1^{-2}\cdots a_N^{-2}$ in (107) by $Na_1^{-2}\cdots a_{N-1}^{-2}$. Having done this it is convenient to parametrize the A's in the way described subsequent to Eq. (106). This gives

$$\psi_T(t)$$
$$=\frac{1}{N-1}\sum_{\underline{n}}{}'\left(\frac{1}{\pi}\sin\frac{n\pi}{N}\right)^{2N-2}\left[\prod_{k=1}^{N-1}\left(n_k-\frac{n}{N}\right)^{-2}\right]$$
$$\times(1-z_{T,\underline{n}}^2t^2)\exp(-\tfrac{1}{2}z_{T,\underline{n}}^2t^2), \tag{109}$$

$$\sum_{\underline{n}}{}'=\sum_{n_1,\ldots,n_{N-1}}\quad\text{with } Nn_k\neq n, \tag{110}$$

$$n=n_1+\cdots+n_{N-1}, \tag{111}$$

$$z_{T,\underline{n}}=\frac{2\pi}{L_d}\sqrt{\frac{kT}{m}}\left(n_1^2+\cdots+n_{N-1}^2-\frac{n^2}{N}\right)^{1/2}. \tag{112}$$

Equation (109) has been used to discuss the form of ψ_T for finite systems, especially for small ones (see Sec. 5 of I).

Because of the rapidly increasing number of terms the numerical evaluation of the series (109) is limited to systems with $N\lesssim 10$. Larger systems call for a transformation of (109), which will now be performed in several steps. We first change the time variable from t to $\nu_T t$:

$$\psi_T(t)=\Psi(\tau),\quad \tau=\nu_T t, \tag{113}$$

$$\Psi(\tau)=\phi(\tau)+\tau\frac{\partial}{\partial\tau}\phi(\tau), \tag{114}$$

$$\phi(\tau)=\frac{1}{N-1}\sum_{\underline{n}}{}'\left(\frac{1}{\pi}\sin\frac{n\pi}{N}\right)^{2N-2}\left[\prod_{k=1}^{N-1}\left(n_k-\frac{n}{N}\right)^{-2}\right]$$
$$\times\exp[-(n_1^2+\cdots+n_{N-1}^2-(n^2/N))\eta], \tag{115}$$
$$\eta=\pi^3\tau^2/2(N-1)^2. \tag{116}$$

Next we use

$$\delta_{n_1+\cdots+n_{N-1}+l,0}$$
$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}d\xi\exp[i\xi(n_1+\cdots+n_{N-1}+l)] \tag{117}$$

to replace the $(N-1)$-fold sum in (115) by an N-fold one:

$$\phi(\tau)=\frac{1}{N-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}d\xi\sum_{\substack{l\neq 0\\ n_1,\ldots,n_{N-1}}}\exp\left[-\eta\left(\frac{l}{N}\right)^2+i\xi\left(\frac{l}{N}\right)\right]\left(\frac{1}{\pi}\sin\pi\frac{l}{N}\right)^{2N-2}$$
$$\times\prod_{k=1}^{N-1}\left\{\exp\left[-\eta\left(n_k+\frac{l}{N}\right)^2+i\xi\left(n_k+\frac{l}{N}\right)\right]\left(n_k+\frac{l}{N}\right)^{-2}\right\}$$
$$=\frac{1}{2\pi(N-1)}\int_{-\pi}^{\pi}d\xi\sum_{l\neq 0}\exp\left[-\eta\left(\frac{l}{N}\right)^2+i\xi\left(\frac{l}{N}\right)\right]\left(\frac{1}{\pi}\sin\pi\frac{l}{N}\right)^{2N-2}$$
$$\times\left\{\sum_{n_2}\exp\left[-\eta\left(n_2+\frac{l}{N}\right)^2+i\xi\left(n_2+\frac{l}{N}\right)\right]\left(n_2+\frac{l}{N}\right)^{-2}\right\}^{N-1}. \tag{118}$$

The sum in the curly bracket is periodic in l with period N. This allows us to introduce two new summation indices n_1 and p by

$$l=Nn_1+p,\quad p=1,\ldots,N-1. \tag{119}$$

Using the definitions

$$S_1(\xi,\eta,\lambda)=\sum_n(n+\lambda)^2\exp[-\eta(n+\lambda)^2+i\xi(n+\lambda)], \tag{120}$$

$$S_2(\xi,\eta,\lambda) = \sum_n \exp[-\eta(n+\lambda)^2 + i\xi(n+\lambda)], \tag{121}$$

$$S_3(\xi,\eta,\lambda) = \sum_n (n+\lambda)^{-2} \exp[-\eta(n+\lambda)^2 + i\xi(n+\lambda)], \tag{122}$$

we may then write ϕ and Ψ as

$$\phi(\tau) = \frac{1}{(N-1)2\pi} \int_{-\pi}^{\pi} d\xi \sum_{p=1}^{N-1} \left(\frac{1}{\pi} \sin \pi \frac{p}{N}\right)^{2N-2} \times S_2(\xi,\eta,p/N) S_3^{N-1}(\xi,\eta,p/N), \tag{123}$$

$$\Psi(\tau) = \frac{1}{(N-1)2\pi} \int_{-\pi}^{\pi} d\xi \sum_{p=1}^{N-1} \left(\frac{1}{\pi} \sin \pi \frac{p}{N}\right)^{2N-2} \times \{S_2(\xi,\eta,p/N) S_3^{N-1}(\xi,\eta,p/N) - 2\eta [S_1(\xi,\eta,p/N) S_3^{N-1}(\xi,\eta,p/N) + (N-1) S_2^2(\xi,\eta,p/N) S_3^{N-2}(\xi,\eta,p/N)]\}. \tag{124}$$

Equation (124) is suited to compute the VAF ψ_T for systems with up to 100 particles. The resulting curves are shown in Fig. 17 of I.

It is also possible to derive from (124) the thermodynamic limit of ψ_T found by Lebowitz and Percus[5] almost twenty years ago. To this end we have to find approximations of the series $S_i(\xi,\eta,\lambda)$ for large N. It is sufficient to find approximations for $S_2(\xi,\eta,\lambda)$ and $S_3(\xi,0,\lambda)$ since

$$S_3(\xi,\eta,\lambda) = S_3(\xi,0,\lambda) - \int_0^{\eta} d\eta' \, S_2(\xi,\eta',\lambda), \tag{125}$$

and S_1 does not appear in ϕ. An obvious approximation for S_2 is

$$S_2(\xi,\eta,\lambda) \approx \int dz \exp(-\eta z^2 + i\xi z) = \left(\frac{\pi}{\eta}\right)^{1/2} \exp\left(-\frac{\xi^2}{4\eta}\right). \tag{126}$$

The series $S_3(\xi,0,\lambda)$ can be summed up without any approximations. Consider the function f_λ defined for $0<\lambda<1$ by

$$f_\lambda(\vartheta) = f_\lambda(\vartheta + 2\pi) = \sum_n \frac{1}{n+\lambda} e^{-in\vartheta} = \frac{\pi}{\sin \pi\lambda} e^{i\lambda(\vartheta-\pi)} \quad \text{for } \vartheta \in (0,2\pi). \tag{127}$$

Now if $\xi \in (-\pi,\pi)$ then

$$\int_0^{2\pi} d\vartheta \, f_\lambda^*(\vartheta) f_\lambda(\vartheta - \xi) = 2\pi \sum_n \frac{1}{(n+\lambda)^2} e^{in\xi} = \frac{\pi^2}{\sin^2 \pi\lambda} e^{-i\lambda\xi} [2\pi + 2i\xi \sin \pi\lambda \exp(i\pi\lambda \operatorname{sgn} \xi)] \tag{128}$$

so that

$$S_3(\xi,0,\lambda) = \frac{\pi^2}{\sin^2 \pi\lambda} \left[1 - \frac{1}{\pi} |\xi| \sin^2 \pi\lambda + \frac{i}{\pi} \xi \sin \pi\lambda \cos \pi\lambda\right]. \tag{129}$$

Inserting (129) and (126) into (125) one obtains

$$S_3(\xi,\eta,\lambda) \approx (\pi/\sin \pi\lambda)^2 [1 + (i/2\pi) \sin 2\pi\lambda] - 2\sqrt{\pi\eta} \exp(-\xi^2/4\eta) - \pi\xi \operatorname{erf}(\xi/2\sqrt{\eta}). \tag{130}$$

Besides (126) and (130) we make the following substitutions and approximations:

$$x = \left(\frac{2\pi}{N}\right) p, \quad \sum_{p=1}^{N-1} \to \left(\frac{N}{2\pi}\right) \int_0^{2\pi} dx, \tag{131}$$

$$y = \xi/(2\sqrt{\eta}), \tag{132}$$

and $N \approx N-1$. This yields

$$\phi(\tau) \approx \frac{1}{2\pi\sqrt{\pi}} \int_{-\pi/2\sqrt{\eta}}^{\pi/2\sqrt{\eta}} dy \int_0^{2\pi} dx \exp(-y^2) \times \{1 + \sqrt{\eta}[i(y/\pi)\sin x - \pi^{-3/2}(1-\cos x) \times \exp(-y^2) - \pi^{-1}(1-\cos x) y \operatorname{erf} y]\}^{N-1}. \tag{133}$$

For fixed τ,

$$\sqrt{\eta} = \frac{\tau\pi}{N-1} \sqrt{\frac{\pi}{2}} = O\left(\frac{1}{N-1}\right), \tag{134}$$

so that for $N \uparrow \infty$ the curly bracket can be approximated by an exponential and the integration over y may be extended from $-\infty$ to $+\infty$. In this approximation

$$\phi(\tau) = \frac{1}{2\pi\sqrt{\pi}} \int_{-\infty}^{\infty} dy \exp[-y^2 - \tau f(y)] \times \int_0^{2\pi} dx \exp\left[\tau f(y) \cos x + i\tau \sqrt{\frac{\pi}{2}} y \sin x\right], \tag{135}$$

where f is the function (89). If we put

$$\tau f(y) = a, \quad \tau\sqrt{\pi/2}\, y = b, \quad b/a = \tanh s, \tag{136}$$

then

$$a^2 > b^2, \quad a = (a^2 - b^2)^{1/2} \cosh s,$$
$$b = (a^2 - b^2)^{1/2} \sinh s,$$
$$\exp(a \cos x + ib \sin x) = \exp[(a^2 - b^2)^{1/2} \cosh(s + ix)]. \tag{137}$$

Now the integral representation of the modified Bessel function of order zero,[6]

$$\int_0^{2\pi} dx \exp[r \cosh(s + ix)] = i \int_{\mathscr{C}} dz \exp[r \cos z] = 2\pi I_0(r), \tag{138}$$
$$\mathscr{C}: \quad (\infty, -i\pi) \to (0, -i\pi) \to (0, i\pi) \to (\infty, i\pi),$$

can be used to carry out one integration in (135). The final form of ϕ is then

$$\phi(\tau) = \frac{1}{\sqrt{\pi}} \int dy \exp[-y^2 - \tau f(y)] I_0[\tau g(y)], \tag{139}$$

$$g(y) = [f^2(y) - (\pi/2) y^2]^{1/2} > 0. \tag{140}$$

From this, we get

$$\Psi(\tau) = \frac{1}{\sqrt{\pi}} \int dy \exp[-y^2 - \tau f(y)] \times \{[1 - \tau f(y)] I_0[\tau g(y)] + \tau g(y) I_1[\tau g(y)]\}. \quad (141)$$

This is the VAF found by Lebowitz and Percus[5] by a different method. If the exponential and the modified Bessel functions[4] are represented by power series and equal powers of τ are collected one obtains the series representation (86) of $\psi_{\infty,\gamma}(t) = \Psi(\nu_{\infty,\gamma} t)$.

We conclude this section by discussing the asymptotic form of the VAF's for $t \uparrow \infty$. For finite systems the dominant contributions to ψ_E and ψ_T are terms with $A = \pm A_j$, $j = 1,...,N$; the corresponding formulas are Eqs. (121) and (124) of I. It is clear from the form of the VAF's for larger systems that the long-time behavior $\psi_{\infty,\gamma}$ emerges from the short-time behavior of ψ_E and ψ_T for large N (see Figs. 16 and 17 of I). The short-time behavior of ψ_E is given by its first polynomial, which in the limit $N \uparrow \infty$ becomes the power series (86). To deduce the asymptotic form of this function it is more convenient to start from the integral representation (135). Changing the integration variable from y to

$$w = y\tau, \quad (142)$$

we obtain

$$\phi(\tau) = \frac{1}{2\tau\pi\sqrt{\pi}} \int_0^{2\pi} dx \exp\left[\left(\frac{\tau}{\sqrt{2}}\right)(\cos x - 1)\right] \times \int_{-\infty}^{\infty} dw\, F\left(\frac{w}{\tau}\right) \exp\left(iw\left(\frac{\pi}{2}\right)^{1/2} \sin x\right), \quad (143)$$

$$F\left(\frac{w}{\tau}\right) = \exp\left\{-\left(\frac{w}{\tau}\right)^2 + \tau(\cos x - 1)\left[f\left(\frac{w}{\tau}\right) - \frac{1}{\sqrt{2}}\right]\right\} = \sum_{k=0}^{\infty} c_k \left(\frac{w}{\tau}\right)^{2k} = 1 - \frac{1}{2}(2 + 2c)\left(\frac{w}{\tau}\right)^2 + (1/4!)(12 + 28c + 12c^2)(w/\tau)^4 - (1/6!)(120 + 504c + 480c^2 + 120c^3)(w/\tau)^6 + (1/8!)(1680 + 11\,664c + 18\,704c^2 + 10\,080c^3 + 1680c^4)(w/\tau)^8 - \cdots,$$

$$c = (\tau/\sqrt{2})(1 - \cos x). \quad (144)$$

The integral over w may be performed for each term of (144) seperately if the result is considered as the generalized function

$$\int dw\, w^{2k} \exp(-iwq) = 2\pi(-1)^k \frac{d^{2k}}{dq^{2k}} \delta(q). \quad (145)$$

To perform the integration over x we shift the boundaries and split the integral:

$$\int_0^{2\pi} dx \cdots = \int_{-\pi/2}^{\pi/2} dx \cdots + \int_{\pi/2}^{3\pi/2} dx \cdots,$$

$$\phi(\tau) = \phi_1(\tau) + \phi_2(\tau). \quad (146)$$

For the first integral we use $\sin x$ instead of x as integration variable,

$$z = \sin x = \sqrt{2/\pi}\, q, \quad (147)$$

so that

$$\int_{-\pi/2}^{\pi/2} dx \exp\left[\left(\frac{\tau}{\sqrt{2}}\right)(\cos x - 1)\right] \times \int dw\, w^{2k} \exp\left(iw\sqrt{\frac{\pi}{2}} \sin x\right) = (-1)^k \sqrt{2\pi}(-2/\pi)^k \exp(-\tau/\sqrt{2}) \times \left| \frac{d^{2k}}{dz^{2k}} \frac{\exp(\tau\sqrt{(1-z^2)}/2}{(1-z^2)^{1/2}} \right|_{z=0} \quad (148)$$

and

$$\phi_1(\tau) = \frac{1}{\tau}\frac{\sqrt{2}}{\pi} + \frac{1}{\tau^3}\frac{2\sqrt{2}}{\pi^2}\left(1 - \frac{5}{2\pi}\right) + \frac{1}{\tau^4}\frac{8}{\pi^3}\left(3 - \frac{8}{\pi}\right) + \frac{1}{\tau^5}\frac{\sqrt{2}}{\pi^3}\left(18 + \frac{36}{\pi} - \frac{461}{2\pi^2}\right) + \cdots. \quad (149)$$

If the second integral (146) is calculated in an analogous way one finds

$$\phi_2(\tau) = O(\tau^{-1} \exp[-\tau\sqrt{2}]), \quad (150)$$

whence these contributions of ϕ may be neglected. From (149) and (114), we get the asymptotic expansion

$$\Psi(\tau) \sim 2\sqrt{2}(2\pi - 5)(\pi\tau)^{-3} - 24(3\pi - 8)(\pi\tau)^{-4} - 2\sqrt{2}(36\pi^2 + 72\pi - 461)(\pi\tau)^{-5} - \cdots. \quad (151)$$

The leading term was first reported by Jepsen in his fundamental paper.[7]

[1] P. Kasperkovitz and J. Reisenberger, Phys. Rev. A **31**, 2639 (1985).
[2] P. Kasperkovitz and J. Reisenberger, J. Math. Phys. **26**, 2601 (1985).
[3] E. Madelung, *Die Mathematischen Hilfsmittel des Physikers* (Springer, Berlin, 1964), p. 243.
[4] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products* (Academic, New York, 1965), pp. 369, 933, 940, 307, 953, 961.
[5] J. L. Lebowitz and J. K. Percus, Phys. Rev. **155**, 122 (1967).
[6] G. N. Watson, *Theory of Bessel Functions* (Cambridge U.P., Cambridge, 1922), p. 181.
[7] D. W. Jepsen, J. Math. Phys. **6**, 405 (1965).

Brownian dynamics approximation to the decay rate of a metastable state

D. L. Weaver
Department of Physics, Tufts University, Medford, Massachusetts 02155

(Received 27 March 1985; accepted for publication 17 May 1985)

An integrodifferential equation for the time dependence of the unabsorbed fraction in a finite diffusion system is derived and solved approximately. The lowest-order approximate solution is equivalent to the first-passage time approximation to the decay rate of the system whether or not the absorbing boundary probability of absorption varies with time.

I. INTRODUCTION

The decay rate of a metastable state is an important parameter in the study of a number of physical systems involved in first-order phase transitions.[1] In these systems, one assumes that the underlying process has a stochastic nature and is describable, at least approximately, as the motion of a classical Brownian particle, the decay rate then being found by calculating and inverting the first passage time[2] for the physical system in question, for example, the switching of states for a homogeneously broadened two-mode laser characterized by a bistable potential,[3] or the desorption from a surface.[4] The method of calculation used in these examples is to set up a totally or partially absorbing barrier and to then find the mean time for absorption to occur by the standard methods for calculating the first-passage time.[2,5–8]

In the present paper, a point of view complementary to the first-passage time approximation to the decay rate is introduced. It is to derive an integrodifferential equation for the time dependence of the unabsorbed fraction $N(t)$ at time t [the integral over the (finite) diffusion space of the position probability density of the Brownian particles]. The equation can, then, be solved approximately. One of the approximations is identical to the first-passage time approximation to the decay rate.

The derivation is of interest both as an alternative view of the first-passage time approximation, and as a possible method to be used when the first-passage time approximation is invalid because of time-dependent boundary parameters. The derivation is carried out in the next section, and the final section is devoted to a discussion of the results with a possible application to more complicated boundaries.

II. THE DERIVATION

A. Boundary conditions with constant coefficients

The approach to be taken below is to derive from the diffusion equation and boundary conditions an integrodifferential equation for the unabsorbed fraction $N(t)$ and then to analyze the conditions that lead to useful approximations. For illustrative purposes, a one-dimensional derivation will suffice with uniform initial density.

Consider, therefore, the one-dimensional diffusion equation (diffusion constant D)

$$\frac{\partial \rho}{\partial t} = D\frac{\partial^2 \rho}{\partial x^2} \tag{1}$$

for probability density $\rho(x,t)$ defined in the space $x = 0$ to $x = b$ at time t. At $x = b$, assume that no absorption occurs and, therefore, that the boundary condition is for total reflection

$$\frac{\partial \rho}{\partial x} = 0, \quad x = b. \tag{2}$$

Recall that $-D\,\partial\rho/\partial x$ is the flux of probability at x, t. At $x = 0$, where a partially absorbing barrier is located, the boundary condition is

$$\frac{\partial \rho}{\partial x} = \kappa\rho, \quad x = 0, \tag{3}$$

with κ a constant determined by the absorption mechanism. In general, $\kappa = \beta/l$, where l is a length parameter and β is an absorption probability parameter, both determined by the particular absorption mechanism of the barrier.

Integrating Eq. (1) twice with respect to x and using the boundary conditions, one obtains

$$\rho(x,t) = -\frac{1}{D\kappa}\int_0^b dy\,\frac{\partial \rho}{\partial t}(y,t) - \frac{1}{D}\int_0^x dz\int_z^b dy\,\frac{\partial \rho}{\partial t}(y,t). \tag{4}$$

Defining

$$N(t) = \int_0^b dx\,\rho(x,t), \tag{5}$$

interchanging the order of integration and differentiation in the first term on the right-hand side of Eq. (4), and then integrating Eq. (4) over the entire diffusion space, one obtains

$$N(t) = -\frac{b}{D\kappa}\frac{dN}{dt} - \frac{1}{D}\int_0^b dx\int_0^x dz\int_z^b dy\,\frac{\partial \rho(y,t)}{\partial t}. \tag{6}$$

A differential equation for $N(t)$ will be found by approximating the triple integral term in Eq. (6). Standard mathematical methods[9] lead to

$$\int_0^b dx\int_0^x dz\int_z^b dy\,\frac{\partial \rho}{\partial t}(y,t) = \frac{d}{dt}\int_0^b dx\,\rho(x,t)(x)\left(b - \frac{1}{2}x\right). \tag{7}$$

The mean value theorem allows one to rewrite the right-hand side of Eq. (7) as

$$\frac{d}{dt}\int_0^b dx\,\rho(x,t)x\left(b - \frac{1}{2}x\right) = \frac{d}{dt}(g(t)N(t)), \tag{8}$$

where $g(t)$ is the appropriate value of $x(b - \frac{1}{2}x)$ for time t to

 0022-2488/85/102627-05$02.50

give the correct value of the integral. Substituting Eq. (8) in Eq. (6), one obtains

$$\frac{d}{dt}\left\{\left(\frac{b}{D\kappa}+\frac{g(t)}{D}\right)N(t)\right\} = -N(t). \tag{9}$$

If $g(t)$ were known for all t, then $N(t)$ could be calculated exactly. Conversely, approximating $g(t)$ in some reasonable way will produce an approximate solution for $N(t)$. For example, suppose that $g(t)$ is constant as an approximation. If $\rho(x,t)$ is known for a particular value of t, say t', then the integral in Eq. (8) may be carried out to determine $g(t')$. In fact, the initial probability distribution $\rho(x,0)$ is known to be uniform in this example, that is, $\rho(x,0) = 1/b$. Thus, $N(0) = 1$, and

$$g(0) = b^2/3, \tag{10}$$

so that Eq. (9) becomes

$$\frac{dN}{dt} \cong \frac{-N(t)}{b/D\kappa + b^2/3D}. \tag{11}$$

The result for $N(t)$ is

$$N(t) \cong N(0)e^{-t/\tau}, \tag{12}$$

that is, approximating $g(t) \cong g(0)$ results precisely in the first-passage time approximation to the decay of $N(t)$, where $\tau \equiv b/D\kappa + g(0)/D$ is the first-passage time known from other methods.[5,6]

The analysis of Eqs. (1)–(12) works equally well when there is a potential energy function in Eq. (1). Then, one is dealing with the Smoluchowski equation. The details are presented in Appendix A for the equivalent one-dimensional problem with potential; the conclusions being the same as those found above, namely, that using $g(t) \cong g(0)$ leads to the first-passage time approximation to $N(t)$, including potential energy effects in τ (see Refs. 5, 6, and 8).

The first-passage time approximation is known to be accurate for all but the shortest times (compared to τ) whenever $b/D\kappa > g(0)/D$. An example is shown in Fig. 1, which illustrates the goodness of the approximation. Even in the worst case, which occurs in one-dimensional problems when $g(0)/D \gg b/D\kappa$, the first passage time approximation is quite good. This is illustrated in Fig. 2 in which the numerical

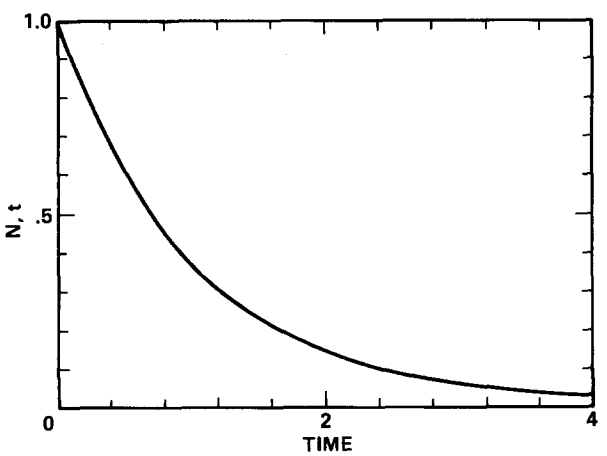


FIG. 1. Plot of $N(t)$, the fraction unabsorbed, versus time in units of $\tau = b/D\kappa + g(0)/D$ with $b\kappa = 10^{-6}$. Both the exact numerical solution and the exponential approximation of Eq. (12) are represented by the curve.

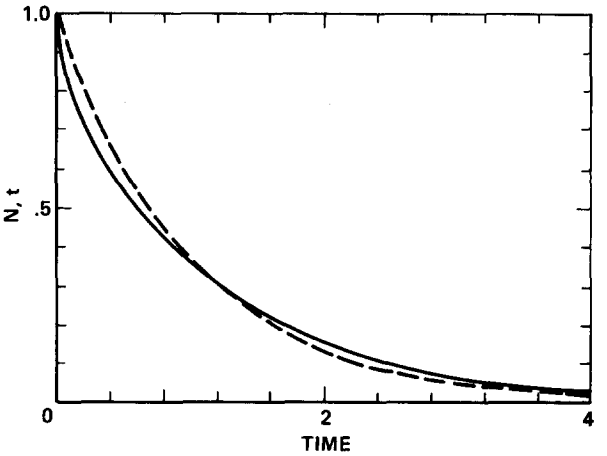


FIG. 2. Plot of $N(t)$, the fraction unabsorbed, versus time in units of $b^2/3D$ for the example in which $b/D\kappa$ is neglected in Eq. (9). The solid curve is the exact numerical solution and the dashed curve is the exponential approximation of Eq. (12) with $N(0) = 1$ and $\tau = b^2/3D$.

value of $N(t)$ calculated from Eq. (B2) in Appendix B is compared to the approximation of Eq. (12) using $\tau = b^2/3D$. For this case, $g(t)$ has also been calculated numerically according to the equation

$$g(t) = \frac{D}{N(t)}\int_t^\infty dt'\, N(t'), \tag{13}$$

using the results of Appendix B. The result for $g(t)$ is shown graphically in Fig. 3. As illustrated, $g(t)$ starts at $t = 0$ with the value $g(0) = b^2/3 = D\tau$ and rises rapidly to a limiting value of $g(\infty) = 4b^2/\pi^2 = D/\lambda_0$, where λ_0 is the smallest eigenvalue of Eq. (1) for the given boundary conditions.

To improve the first-passage time approximation to $N(t)$ defined by $g(t) \cong g(0)$, one might take a two-exponential approximation

$$N(t) \cong A_0 e^{-\lambda_0 t} + (1 - A_0)e^{-\lambda t}, \tag{14}$$

with A_0 and λ_0 being defined in Appendix B. This approximation satisfies the initial condition $N(0) = 1$, and also the sum rule $\int_0^\infty dt\, N(t) = \tau$ when one chooses for λ the value

$$\lambda = (1 - A_0)/(\tau - A_0/\lambda_0). \tag{15}$$

In the particular example evaluated in Appendix B, $A_0 = 8/$

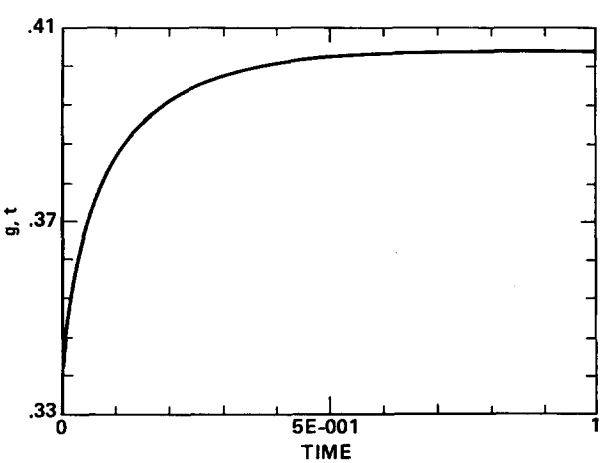


FIG. 3. Plot of $g(t)$, defined by Eq. (13), versus time in units of b^2/D.

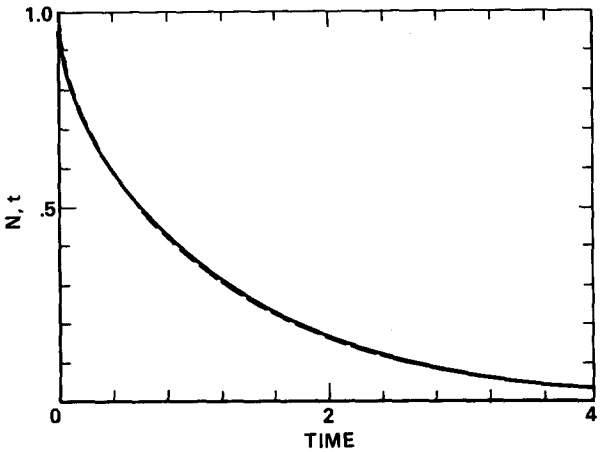


FIG. 4. Plot of $N(t)$, the fraction unabsorbed, versus time in units of $b^2/3D$ for the example in which $b/D\kappa$ is neglected in Eq. (9). The solid curve is the exact numerical solution and the dashed curve is the exponential approximation of Eq. (15) with $\tau = b^2/3D$, $\lambda = \pi^4/(12\tau)$, $g(0) = b^2/3$, and $g(\infty) = 4b^2/\pi^2$.

$\pi^2 = 0.8106$. The fit using Eq. (14) is shown in Fig. 4 and is seen to be indistinguishable from the numerical evaluation of $N(t)$. It should be emphasized, however, that in many cases of physical interest, the approximation $g(t) \cong g(0)$ is very good, as long as κ is independent of time.

B. Boundary conditions with time-dependent coefficients

The analysis leading to Eq. (9) may be extended to include physical situations in which the partially absorbing boundary, described by Eq. (3), has a time-dependent coefficient $\kappa(t)$. Such a situation could arise because of a well-defined physical or chemical modulation of the absorption probability at the boundary. To treat this possibility, Eq. (9) must be modified to

$$\frac{b}{D\kappa(t)}\frac{dN}{dt} + \frac{d}{dt}\left(\frac{g(t)}{D}N(t)\right) = -N(t). \tag{16}$$

As a first approximation to solving Eq. (16) for $N(t)$, one may replace $g(t)$ by its initial value $g(0)$ as proved successful in the preceding case of constant κ. With this approximation, Eq. (16) becomes

$$\left[\frac{b}{D\kappa(t)} + \frac{b^2}{3D}\right]\frac{dN}{dt} = N(t). \tag{17}$$

Equation (17) may be integrated in cases of interest to give an approximation to $N(t)$,

$$N(t) \cong \exp\left(-\int_0^t \frac{dt'}{\hat{\tau}(t')}\right), \tag{18}$$

where

$$\hat{\tau}(t) \equiv b/D\kappa(t) + b^2/3D, \tag{19}$$

the time-dependent analog of the first-passage time found in the preceding section. Comparison of numerical solutions of the diffusion equation for $N(t)$ with analytical integration of Eq. (18) have produced excellent agreement (details to be published elsewhere[10]), thus indicating the validity of the $g(t) \cong g(0)$ approximation even for time-dependent $\kappa(t)$.

If the absorption probability modulation is periodic, one may use the time integral in Eq. (18) to define a time-average passage time $\langle\hat{\tau}\rangle$ by

$$\langle\hat{\tau}\rangle \equiv \frac{t}{\int_0^t dt'[1/\hat{\tau}(t')]}, \tag{20}$$

where t is an integer multiple of the period of $\beta(t)$ in the definition $\kappa(t) = \beta(t)/l$ as discussed above. For a square wave absorption probability, that is $\beta = 1$ for a time t_1 followed by $\beta = 0$ for time t_0, the cycle repeating with period $t_1 + t_0$, one obtains

$$\langle\hat{\tau}\rangle_{SW} = T(1 + 1/\gamma)(1 + t_0/t_1), \tag{21}$$

where $\gamma \equiv b/(3l)$ and $T \equiv b^2/3D$. For a sawtoothed absorption probability, that is, β rises linearly to one from zero in time t_1, and subsequently falls linearly to zero in time t_0, with periodic repetition, one finds that

$$\langle\hat{\tau}\rangle_{SAW} = T/[1 - (1/\gamma)\ln(1 + \gamma)]\,. \tag{22}$$

For a sine-squared absorption probability, that is, $\beta(t) = \sin^2(\alpha t)$ with $\alpha = 2\pi/P$ and P the period of modulation, one obtains

$$\int_0^t \frac{dt'}{\hat{\tau}(t')} = \frac{1}{T}\left[t - \frac{1}{2\alpha\sqrt{1+\gamma}}\tan^{-1}(\sqrt{1+\gamma}\tan(\alpha t))\right] \tag{23}$$

and

$$\langle\hat{\tau}\rangle_{SIN} = T/[1 - 1/2\sqrt{1+\gamma}]. \tag{24}$$

Clearly the time-averaged passage time only gives a reasonable approximation to the overall time dependence of $N(t)$ when the periodicity of the absorption probability is either smaller than or not much bigger than the characteristic diffusion time T.

III. DISCUSSION

In this paper, a differential equation has been derived for the unabsorbed fraction in a finite diffusion system with a partially or totally absorbing boundary. The equation has been solved approximately both for boundaries with time-independent parameters, in which case it represents the first-passage time approximation, and for boundaries with time-dependent parameters.

Although the analysis has been restricted to one dimension, it may be readily extended to more complicated situations, including two- and three-dimensional spherically symmetrical situations, and motion on the surface of a sphere.

This work is somewhat complementary to the recent study of Bunde and Gouyet,[11] in which the modulation of various parameters from Brownian motion in a double-well potential was investigated in the WKB approximation. Related work has also been carried out recently by Northrup, Zanin, and McCammon[12] and by Szabo *et al.*[13] on the binding rate of ligands to proteins in which the reactivity of the binding site fluctuates in time. However, the latter studies[13] involve absorption probabilities that fluctuate stochastically in time rather than in the well-defined manner considered above, and both studies are applied to biomolecular diffu-

sion-influenced reactions, rather than the unimolecular, finite diffusion-space cases discussed in the present paper.

ACKNOWLEDGMENTS

Funds for support of this study have been allocated by NASA Ames Research Center, Moffett Field, California, under Interchange No. NCA2-OR785-201. The author is grateful to the Tufts University Faculty Research Awards Committee for a Biomedical Research Grant in 1982–83.

APPENDIX A: THE DERIVATION WITH EXTERNAL POTENTIAL

The one-dimensional Smoluchowski equation is

$$\frac{\partial \rho}{\partial t} = D\frac{\partial}{\partial x}\left(\frac{\partial \rho}{\partial x} + \rho\frac{dV}{dx}\right) = -\frac{\partial}{\partial x}J(x,t), \tag{A1}$$

where the units are such that $k_B T = 1$ with k_B the Boltzmann constant and T the absolute temperature. The boundary conditions are

$$J(b,t) = 0 \tag{A2}$$

and

$$\frac{\partial \rho}{\partial x} + \rho\frac{dV}{dx} = \kappa\rho, \quad x = 0. \tag{A3}$$

Integrating Eq. (A1) twice with respect to x and using the boundary conditions one obtains

$$\rho(x,t) = -\frac{1}{D\kappa}e^{V(0)-V(x)}\int_0^b dz\,\frac{\partial \rho(z,t)}{\partial t} - \frac{e^{-V(x)}}{D}\int_0^x dy\,e^{V(y)}\int_y^b dz\,\frac{\partial \rho(z,t)}{\partial t}. \tag{A4}$$

Using the definition in Eq. (5) for $N(t)$, interchanging the order of integration and differentiation in the first term on the right-hand side of Eq. (A4), and then integrating Eq. (A4) over the entire diffusion space, one obtains

$$N(t) = -\int_0^b dx\,\frac{e^{V(0)-V(x)}}{D\kappa}\frac{dN}{dt} - \frac{1}{D}\int_0^b dx\,e^{-V(x)}\int_0^x dy\,e^{V(y)}\int_y^b dz\,\frac{\partial \rho(z,t)}{\partial t}. \tag{A5}$$

Further integration of $N(t)$ over all time leads to the first passage time including potential energy effects.[5,6,8] The result for an initial condition in which all the probability density is at one point $x = x_0$ is

$$\tau(x_0) = \frac{1}{D\kappa}\int_0^b dz\,e^{V(0)-V(z)} + \frac{1}{D}\int_0^{x_0} dy\,e^{V(y)}\int_y^b dz\,e^{-V(z)}. \tag{A6}$$

Rewriting the last term on the right-hand side of Eq. (A5) as

$$\int_0^b dx\,e^{-V(x)}\int_0^x dy\,e^{V(y)}\int_y^b dz\,\frac{\partial \rho^{(z,t)}}{dt} = \frac{d}{dt}\int_0^b dx\,\rho(x,t)f(x), \tag{A7}$$

where

$$f(x) = \int_0^x dy\,e^{V(y)}\int_x^b dz\,e^{-V(z)} + \int_0^x dy\,e^{-V(y)}\int_0^y dz\,e^{V(z)}, \tag{A8}$$

one uses the mean-value theorem to obtain

$$\frac{d}{dt}\int_0^b dx\,\rho(x,t)f(x) = \frac{d}{dt}(\hat{g}(t)N(t)), \tag{A9}$$

where $\hat{g}(t)$ is the extension to include potential energy effects of the quantity defined in Eq. (8), that is, the appropriate value of $f(x)$ for time t to give the correct value of the integral. Substituting Eq. (A9) in Eq. (A5), one obtains

$$\frac{d}{dt}\left\{\left(\int_0^b dx\,\frac{e^{V(0)-V(x)}}{D\kappa} + \frac{\hat{g}(t)}{D}\right)N(t)\right\} = -N(t). \tag{A10}$$

If one chooses to approximate $\hat{g}(t)$ with $\hat{g}(0)$, one obtains the result

$$\hat{g}(0) = \int_0^{x_0} dy\,e^{V(y)}\int_y^b dz\,e^{-V(z)}, \tag{A11}$$

and, thus,

$$\tau(x_0) = \int_0^b dx\,\frac{e^{V(0)-V(x)}}{D\kappa} + \frac{\hat{g}(0)}{D}, \tag{A12}$$

so that there is a complete parallel with Eqs. (1)–(12) when potential energy effects are included in the analysis.

APPENDIX B: EXACT SOLUTION WITH $\kappa = 0$

Solving Eqs. (1)–(3) with $\rho = 0$ at $x = 0$ one obtains, using standard mathematical methods,

$$\rho(x,t) = \frac{4}{\pi b}\sum_{n=0}^{\infty}\frac{1}{2n+1}\sin\left(\frac{(2n+1)\pi x}{2b}\right) \times\exp\left(-\frac{D}{b^2}\left[\frac{(2n+1)\pi}{2}\right]^2 t\right), \tag{B1}$$

where a uniform initial distribution $\rho(x,0) = 1/b$ has been chosen. Using Eq. (5), one obtains

$$N(t) = \sum_{n=0}^{\infty}A_n e^{-\lambda_n t} = \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\exp\left(-\frac{D}{b^2}\left[\frac{(2n+1)\pi}{2}\right]^2 t\right), \tag{B2}$$

with A_n and λ_n [the eigenvalues of Eq. (1) for the given boundary conditions] being defined by Eq. (B2). From $N(t)$, one obtains the integral

$$\int_t^{\infty} dt'\,N(t') = \frac{32b^2}{\pi^4 D}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4} \times\exp\left(-\frac{D}{b^2}\left[\frac{(2n+1)\pi}{2}\right]^2 t\right). \tag{B3}$$

From Eqs. (B2) and (B3), one sees that

$$\lim_{t\to\infty} g(t) = \lim_{t\to\infty}\frac{D}{N(t)}\int_t^{\infty} dt'\,N(t') = \frac{4b^2}{\pi^2} = D/\lambda_0. \tag{B4}$$

[1]H. Haken, *Synergetics* (Springer, Berlin, 1978).
[2]G. H. Weiss, Adv. Chem. Phys. **13**, 1 (1967).
[3]R. Roy, R. Short, J. Durnin, and L. Mandel, Phys. Rev. Lett. **45**, 1486 (1980).
[4]D. L. Weaver, Phys. Lett. A **66**, 445 (1978).
[5]D. L. Weaver, Phys. Rev. B **20**, 2558 (1979).
[6]A. Szabo, K. Schulten, and Z. Schulten, J. Chem. Phys. **72**, 4350 (1980).
[7]D. L. Weaver, J. Chem. Phys. **72**, 3483 (1980).
[8]J. M. Deutsch, J. Chem. Phys. **73**, 4700 (1980).
[9]G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, New York, 1961).
[10]D. L. Weaver, to be published.
[11]A. Bunde and J. F. Gouyet, Z. Phys. B **42**, 169 (1981).
[12]S. H. Northrup, F. Zarrin, and J. A. McGammon, J. Phys. Chem. **86**, 2314 (1982).
[13]A. Szabo, D. Shoup, S. H. Northrup, and J. A. McCammon, J. Chem. Phys. **77**, 4484 (1982).

Global instability of a monoparametric family of vector fields representing the unfolding of a dissipative structure

Ariel Fernández
Chemical Physics Department, The Weizmann Institute of Science, Rehovot, Israel 76100

(Received 25 February 1985; accepted for publication 4 June 1985)

The continuous structural stability (CSS) of arcs of vector fields defined on a Banach space or, in general, on a normed space is introduced. It is shown that the nonequilibrium sensitivity of reaction-diffusion systems to external fields reflects itself in the lack of CSS for a family of arcs, which model the onset of two equally probable spatial organizations. It is therefore demonstrated that a universal model up to topological equivalence for the unfolding of a spatial dissipative structure does not exist. The results generalize some conclusions already available for bifurcation of stationary states in arcs of vector fields on compact manifolds.

I. INTRODUCTION

Perhaps the best studied example of a monoparametric family (arc) of vector fields, $\{N(\lambda)\}$, on a Banach space B of vector fields is the one provided by the Navier–Stokes equations.[1–3] The time evolution of a vector field $v \in B$ is given at a fixed value of the parameter λ (here, the Reynolds or the Rayleigh number) by $N(\lambda)v$:

$$\dot{v} = \frac{dv}{dt} = N(\lambda)v.$$

The notion of continuous structural stability (CSS) of arcs $\{N(\lambda)\}$ with λ in an interval I of the line was introduced initially for vector fields on a compact manifold.[4,5] It can be considered in the case of vector fields defined on a space of vector fields.[6]

We intend to formalize the notion of CSS for arcs $\{N(\lambda)\}$, where $N(\lambda)$ is a vector field (operator) on the normed space

$$C^r_{[-l,l]},\ r \geqslant 2.$$

II. THEORY

We are interested in arcs presenting a bifurcation at a critical value $\lambda = \lambda_c$ in the interior of the interval I. The results are applied to reaction-diffusion models operating far from the thermodynamic equilibrium regime.

The influence of external (gravitational or electric) fields in the selection of a dissipative structure over another in the critical region near λ_c (see Ref. 7) will reflect itself in the lack of CSS of the arc, which represents the case in which both dissipative structures are equally probable. In the present work, we shall be concerned with steady-state–spatial organizations. The notion of CSS naturally extends the concept of structural stability of a fixed vector field representing far-from-equilibrium kinetics[8,9] to the case in which a locally attractive and locally invariant center manifold appears in the phase space indicating that a dissipative structure has emerged.[10–12] Consider the arc $N(\ ): I \to \mathrm{Op}(C^r \to C^{r-2})$. This arc maps the real interval I on a space of operators on $C^r_{[-1,1]}$, $r \geqslant 2$. Each operator $N(\lambda)$ defines a vector field that assigns the vector $N(\lambda)f$ belonging to $C^{r-2}_{[-l,l]}$ to each vector f of $C^r_{[-l,l]}$. The associated dynamics is given by

$$\dot{f} = N(\lambda)f. \tag{1}$$

$C^p_{[-l,l]}$ is a normed space with the norm

$$||f||_p = \sup_{\substack{z \in [-l,l] \\ 0 < j < p}} \left(\left| \frac{d^j}{dz^j} f(z) \right| \right). \tag{2}$$

Operators such as d^2/dz^2 or d/dz are bounded seen as elements of $\mathrm{Op}(C^r \to C^{r-2})$ with the supreme norm since

$$\left|\left| \frac{df}{dz} \right|\right|_{r-2} \leqslant ||f||_r, \quad \text{for all } f \in C^r_{[-l,l]}, \tag{3}$$

and

$$\left|\left| \frac{d^2}{dz^2} f \right|\right|_{r-2} \leqslant ||f||_r.$$

Therefore

$$\left|\left| \frac{d}{dz} \right|\right| \equiv \sup_{f \in C^r} \left(\frac{||(d/dz)f||_{r-2}}{||f||_r} \right) \leqslant 1, \tag{4}$$

and also

$$\left|\left| \frac{d^2}{dz^2} \right|\right| \leqslant 1. \tag{5}$$

We shall also introduce the restriction that the map $(f, \lambda) \to N(\lambda)f$ is C^1. (Throughout this paper this will be referred to as restriction R.)

The distance d between arcs $M(\lambda)$ and $N(\lambda)$ is given by

$$d(M,N) = \sup_{\lambda \in I} (||M(\lambda) - N(\lambda)||). \tag{6}$$

Now we extend the concept of CSS: An arc $N(\lambda)$ satisfying the restriction R has CSS if there exists a number $\delta > 0$ such that for all arcs $M(\lambda)$ satisfying restriction R and

$$d(M,N) \leqslant \delta, \tag{7}$$

there exists a family of bijective bounded operators with bounded inverse $\{h(\lambda)\}$ operating on C^{r-2} (hence also on C^r), which make the diagram

$$\begin{array}{ccc} & N(\lambda) & \\ C^r & \longrightarrow & C^{r-2} \\ {\scriptstyle h(\lambda)}\downarrow & & \downarrow{\scriptstyle h(\lambda)} \\ C^r & \longrightarrow & C^{r-2} \\ & M(S(\lambda)) & \end{array} \tag{8}$$

commutative for a continuous parametrization $S: I \to I$. The family $h(\lambda)$ is also subject to restriction R. As a consequence,

 0022-2488/85/102632-02$02.50

if "Ker" denotes the kernel of an operator, we have that the restriction $h(\lambda)$:Ker $N(\lambda) \to$ Ker $M(S(\lambda))$ is a homeomorphism for every λ in I.

Precisely, consider the case in which Ker $N(\lambda)$ becomes nontrivial for $\lambda > \lambda_c$. (As we shall see, this corresponds in a reaction-diffusion model to the onset of dissipative structures in form of spatial–steady-state organizations.) This means that a bifurcation occurs at $\lambda = \lambda_c$. The fact that an arc has CSS or not depends on the type of bifurcation. For mathematical simplicity consider the arc

$$\lambda \to \left[D \frac{d^2}{dz^2} + F(\ ,\lambda) \right] = N(\lambda), \tag{9}$$

representing the behavior of the concentration f in a reaction-diffusion system operating far from equilibrium. Here D is Fick's diffusion coefficient and the nonlinear function $F(f,\lambda)$ expresses the kinetics. Equation (1) gives the time evolution of the concentration. Such models were considered by Kondepudi and Prigogine,[7] and a simple realization of them is the coupling of autocatalysis with rate constant λ of a single chemical to diffusive transpot to produce a "freezed wave" of concentrations in the Belousov–Zhabotinsky batch reactor of Busse.[13] Zero-flux boundary conditions are imposed:

$$\left. \frac{df}{dz} \right|_{z = \pm l} = 0. \tag{10}$$

The bifurcation of a spatial dissipative structure occurs at $\lambda = \lambda_c$ if an eigenvalue of the f-Fréchet derivative of $N(\lambda)$ $(\partial N(\lambda)/\partial f)$, which is negative for $\lambda < \lambda_c$, becomes positive for $\lambda > \lambda_c$. We recall that restriction R holds and therefore the center manifold theorem applies.[14]

Near criticality, the stationary states are

$$f_0(z) = C_0 + \alpha \sin(n_c \pi z/2l), \tag{11}$$

where C_0 is the stationary state of the kinetics $[F(C_0,\lambda) = 0]$, the wavenumber n_c is given by

$$-D\left(\frac{n_c \pi}{2l}\right)^2 + \left. \frac{\partial F}{\partial f} \right|_{f = C_0, \lambda = \lambda_c} = 0, \tag{12}$$

and the amplitude $\alpha = \alpha(\lambda)$ obeys the bifurcation equation of the form

$$B\alpha^3 + A_1(\lambda - \lambda_c)\alpha = 0. \tag{13}$$

Here B and A_1 are constants that can be obtained, using Sattinger's method, from group representation theory[15] and they depend on the f and λ derivatives of $N(\lambda)$ at criticality.

In general, each arc $N(\lambda)$ has, associated with it a bifurcation equation whose real roots $\alpha = \alpha(\lambda)$ give the amplitude of the attractors that form Ker $N(\lambda)$. The set of these roots is called the bifurcation set.

We have the following general result.

Lemma: If two arcs $N(\lambda)$, $M(\lambda)$ subject to restriction R are topologically equivalent in the sense that the family $\{h(\lambda)\}$ exists, then their respective bifurcation sets are homeomorphic.

This result follows from the fact that $h(\lambda)$:Ker $N(\lambda)$ $\to$ Ker $M(S(\lambda))$ is a homeomorphism for each $\lambda \in I$ and $\{h(\lambda)\}$ is subject to restriction R. We can now state and prove the following theorem.

Theorem: The arc $N(\lambda)$ given by Eqs. (9) and (10) is not continuously structurally stable. The arcs $M_g(\lambda)$:$\lambda \to D(d^2/dz^2) + F(\ ,\lambda) - \mu g(d/dz)$, which have CSS and correspond to the coupling of a reaction-diffusion system with an external field of intensity g can approximate $N(\lambda)$ as much as desired by regulating g but they are not topologically equivalent to $N(\lambda)$ in the sense described by Eq. (8). (Here μ represents the drift of the molecules, cf. Ref. 7.)

Proof: The bifurcation set for the arcs $M_g(\lambda)$ is the set of real solutions of the equation

$$B\alpha^3 + A_1(\lambda - \lambda_c)\alpha + A_2 g = 0, \quad A_2 \neq 0. \tag{14}$$

The corresponding curve in the α-λ plane is a Newton trident asymptotic to the λ axis with minimum separation $k_g = \frac{3}{2}[-4A_2 g/B]^{1/3}$ between the stable branches. So if $g \neq 0$, no bifurcation occurs, and the presence of the external field precludes it by "helping" the system to select one stable branch over the other.

Clearly, for any interval I containing λ_c in its interior, the bifurcation sets given by Eqs. (13) and (14) are not homeomorphic. In the first case both stable branches are equally probable and their amplitudes are proportional to $\pm [\lambda - \lambda_c]^{1/2}$, respectively. In the second case, however, the second stable nonthermodynamic branch emerges suddenly separated by a distance k_g from the first branch. Given $\delta > 0$, arbitrarily small, we can always find an arc M_g such that $d(M_g, N) \leqslant \delta$. It suffices to take the intensity of the field as $g = \delta/|\mu|$, since from Eq. (3)

$$||M_g - N|| = |\mu| g \left| \left| \frac{d}{dz} \right| \right| \leqslant g|\mu|. \tag{15}$$

The theorem is proved.

It has been established that the CSS condition holds in many cases of arcs of vector fields on compact manifolds presenting bifurcation of steady states. The reader can consult Ref. 2 for the saddle-node and "flip" cases. However, the present results show that this is not the case in the more general context where nonhomogeneous steady states can bifurcate from the homogeneous phase corresponding to the "thermodynamic branch."

[1] D. Ruelle and F. Takens, Commun. Math. Phys. **20**, 167, (1971).
[2] P. Brunovsky, "One parameter families of diffeomorphisms," presented at the Symposium on Differential Equations and Dynamical Systems, Warwick, 1968–1969.
[3] M. Hirsh, C. Pugh, and M. Schub, Bull. Am. Math. Soc. **76**, 1015, (1970).
[4] S. Smale, Bull. Am. Math. Soc. **66**, 43 (1960).
[5] S. Smale, Ann. Math. **74**, 199 (1961).
[6] J. Leray, Acta Math. **63**, 193 (1934).
[7] D. Kondepudi and I. Prigogine, Physica A **107**, 1 (1981).
[8] A. Fernández and O. Sinanoğlu, J. Math. Phys. **25**, 406 (1984).
[9] A. Fernández and O. Sinanoğlu, J. Math. Phys. **25**, 2576 (1984).
[10] A. Fernández, Ph D. thesis, Yale University, February, 1984.
[11] A. Fernández, and O. Sinanoğlu, Phys. Rev. A **29**, 2029 (1984).
[12] A. Fernández and O. Sinanoğlu, Phys. Rev. A **30**, 1522 (1984).
[13] H. Busse, J. Phys. Chem. **73**, 750 (1969).
[14] J. Marsden, Bull. Am. Math. Soc. **84**, 6, 1139 (1978).
[15] D. Sattinger, *Topics in Stability and Bifurcation Theory, Springer Mathematical Lectures*, Vol. 309 (Springer, New York, 1973).

Practical pretheories of QED. I. Properties of the interaction picture

M. C. Yoakam
Physics Department, Purdue University, West Lafayette, Indiana 47906

(Received 13 March 1984; accepted for publication 5 April 1985)

Discussions and generalizations of the usage of variational methodology for second-quantized quantum field theories are presented, including extensions to situations [which mimic quantum electrodynamics (QED)] in which the new canonical variables obtained for different pictures are not the simple images of the interaction picture canonical variables. Methods helpful for segregating the new canonical variables and for insuring Poincaré invariance of the resultant theories are included; the latter being necessary because most of the state-evolution equations are not explicitly invariant (covariant). A parametrization of the boson commutators is introduced which includes the Proca and Coester fields as special cases and which, in the zero-mass limit, includes the Yennie, Landau, and Feynman gauge fields. If the Coester fields satisfy the Lorentz condition in the weak-convergence limit, then it is shown that all fields parametrized in this fashion satisfy the Lorentz condition in the weak-convergence limit, and that invariance under field interchange of such heavy-meson field operators reduces to special gauge invariance in the zero-mass limit. Superfield formalism for the spinor fields is introduced in which none of the spinor fields are segregated as "physical" fields.

I. INTRODUCTION

This article and its sequel[1] are devoted to the discussion of a parametrization (based on gauge and renormalization parameters) of a subset of the neutral massive vector fields which reduces to quantum electrodynamics in the zero-mass limit. This particular subset is chosen such that all admissible vector fields satisfy the subsidiary condition in the weak-convergence limit (for restricted values of the commutator parameters), and such that field interchange within this subset reduces to special-gauge invariance in the zero-mass limit. Zero-mass limits will be specifically avoided so that the results will be infrared convergent.[2,3] Theories based on vector fields within this group, with interactions of the form ($j \cdot A$ + renormalization terms + whatever else is necessary), will be referred to as "pretheories of QED" (quantum electrodynamics).

The existence of an interaction picture will be assumed so that the commutation relations and field equations at all points in space-time will be available.[4] The existence of a constant-state picture in which the field equations and transformation properties of the fields are *a priori* known (i.e., Heisenberg picture) will not be assumed because such assumptions overspecify the pretheories.[5]

The ultraviolet divergences (mass, charge, wave function) will be controlled by introducing a sufficient but unspecified number of auxiliary fields.[6,7] This regulation prescription is simpler to implement within a variational formalism than some of the other methods of regulation (e.g., analytic regulation[8]), and alternate methods of regulation and renormalization do not appear to contribute additional information. Renormalization group methods[9] will not be incorporated at this time because the procedure is well defined only when the renormalization constants are finite, and because the procedure has been developed for the control of time-ordered exponentials based on interactions that do not, in general, yield all of the proper Heisenberg picture field equations even in the weak-convergence limit. An auxiliary-field-regulated pretheory, which assumes only the existence of an interaction picture, will be referred to as a "practical pretheory of QED."

Readers familiar with the ultraproduct construction from sentential logic (often associated with nonstandard extensions[10]) will recognize the possibility of defining QED as an ultraproduct based on the practical pretheories because only those results which are independent of both the commutator parameters and the renormalization parameters are to be retained. Ultraproducts may prove to be very useful in quantum field theory because they provide a mathematically well-defined procedure to replace (heuristically) the statement "for every $\cdots$" by the less restrictive statement "for almost every $\cdots$." A suitably restricted set of practical pretheories would form an acceptable base for the representation of QED as an ultraproduct. The C^* algebras,[11] used to extend the appearance of separability to situations that are not separable, may approximate (if not duplicate) the later use of a Fermi operator defined on one arbitrarily chosen space-time surface.

The construction of pretheories with stationary subsidiary conditions (as weak-convergence relations) in constant-state pictures based on ($j \cdot A + \cdots$)-type interactions has, in the absense of the Boson mass counterterms, been successfully completed.[5] Such pretheories are still ultraviolet divergent, but they supplement the zero-mass works[12–16] and the massive meson work[17] by the inclusion of the spinor mass splittings and by the emphasis on the properties of the constant-state pictures, which are really being used. Our primary focus is on the constant-state pictures, which can be connected with the interaction picture before the zero-mass limits are taken (i.e., adjusting the series "sum" before attempting term-by-term corrections), and on the modifica-

 0022-2488/85/102634-05$02.50

tions to the interaction operator, which are required to make the constant-state pictures which are accessible (when starting from the interaction picture) more closely approximate the desired Heisenberg picture. The boson mass counterterms have not yet been successfully incorporated: the implication is that the symmetry group for the pretheories does not yet simultaneously include the Poincaré, renormalization, and gauge symmetries.[5]

This article will focus on the restrictive role that the action plays for second quantized field theories, the problems associated with attempting to reach constant-state pictures which resemble the Heisenberg picture, and the selection of appropriate interaction picture fields.

II. CONSTANT-STATE PICTURES: INTERACTION PICTURES: INTERACTIONS

An interaction picture is a picture for which the spinor and boson fields satisfy the source-free field equations and commutation relations. The emphasis on interaction pictures is predicated both by the frequent use of these pictures as starting points, and the fact that a choice of picture specifies the action integral. By this last statement, it is implied that if $\mathsf{L}_{\text{free}}(\zeta,\partial\zeta)$ is the free-field Lagrangian density and if $\mathsf{L}_W(\zeta^W,\partial\zeta^W)$ is the Lagrangian density in any picture connected to the interaction picture via $\zeta^W = \mathsf{U}_W\zeta\mathsf{U}_W^{-1}$ and $\Omega_W = \mathsf{U}_W\Omega$, then

$$\mathsf{L}_W(\zeta^W,\partial\zeta^W) = \mathsf{L}_{\text{free}}((\mathsf{U}_W^{-1}\zeta^W\mathsf{U}_W),\ \partial(\mathsf{U}_W^{-1}\zeta^W\mathsf{U}_W)).$$

Variational principles which do not contain specific information within L concerning the commutation relations appear to retain sufficient freedom to reach almost any picture of interest because fewer questions are being answered. Variational principles, which contain specific information on the commutation relations in one picture (e.g., Schwinger's generalized action principle), or which have such information postulated externally (e.g., free-field commutation relations in the interaction picture), limit the collection of pictures which can be reached using canonical transformations of the form $\mathsf{U}_W\zeta\mathsf{U}_W^{-1}$. In discussions of the Feynman gauge massless fields,[12,18,19] the Proca fields,[20] and the pretheory fields,[5] reaching constant-state pictures in which the desired field equations (including gauge conditions) are at least satisfied in the weak-convergence limit requires sacrificing either the spinor transformation properties of the images of the Fermi fields or the vector transformation properties of the images of the boson fields. To avoid overspecifying the problem, one may postulate only two of the following: an interaction picture; a Heisenberg picture (constant-state picture satisfying particular properties); or the interaction operator defining the canonical transformation connecting the two. Only those pictures which may be connected to the interaction picture via a canonical transformation U_W will be considered. Therefore, the form of the interaction is assumed to be at least partially known, i.e.,

$$i\frac{d\Omega}{d\tau} = W(\tau)\Omega.$$

Under the assumption of integrability (path independence) of the Schrödinger equation,[5,20] a constant-state picture can always be reached via

$$i\frac{d}{d\tau}\Omega_W = \mathsf{U}_W(\tau,\tau_0)(W_W(\tau) + W(\tau))\mathsf{U}_W(\tau,\tau_0)^{-1}\Omega_W,$$

where $\Omega_W = \mathsf{U}_W(\tau,\tau_0)\Omega$ and $i(d/d\tau)\mathsf{U}_W = \mathsf{U}_W W_W$, by choosing $W_W(\tau) = -W(\tau)$. The relative field equations can be obtained via

$$\partial\mu\mathsf{O}(x,\tau)^W = \mathsf{U}_W(\tau,\tau_0)(\partial\mu\mathsf{O}(x,\tau) + in_\mu \lim_{x\to(x\in\tau)}([W_W(\tau),\mathsf{O}(x,\tau)]))\times\mathsf{U}_W(\tau,\tau_0)^{-1},$$

and the relative transformation properties can be obtained via

$$\delta\mathsf{O}(x,\tau)^W = \mathsf{U}_W(\tau,\tau_0)(\delta\mathsf{O}(x,\tau))\mathsf{U}_W(\tau,\tau_0)^{-1} + (\delta\mathsf{U}_W(\tau,\tau_0))\mathsf{O}(x,\tau)\mathsf{U}_W(\tau,\tau_0)^{-1} + \mathsf{U}_W(\tau,\tau_0)\mathsf{O}(x,\tau)(\delta\mathsf{U}_W(\tau,\tau_0)^{-1}).$$

The constant-state picture reached often only poorly approximates the desired Heisenberg picture.[5,20] One method for adjusting the form of W so that the constant picture reached more closely approximates the desired Heisenberg picture is the introduction of a Fermi operator.

III. FERMI OPERATORS

To minimize the difference between the desired Heisenberg picture and constant-state pictures which can be reached using a canonical transformation,[5,12,20] it is often useful to split the interaction U_H into two parts G and U_{IH} (i.e., $\mathsf{U}_H = \mathsf{G}^{IH}\mathsf{U}_{IH} = \mathsf{U}_{IH}\mathsf{G}$), where U_{IH} satisfies an integrable (path-independent) Schrödinger equation of the form

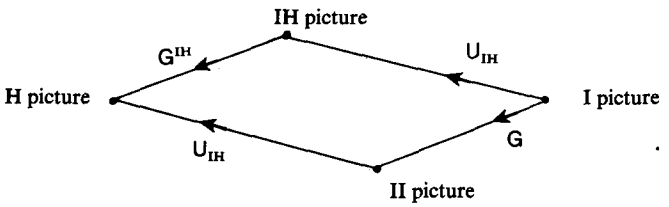
$$i\frac{d}{d\tau}\mathsf{U}_{IH} = \mathsf{U}_{IH}\mathsf{W}_{IH}$$

and G is an operator defined on a single spacelike surface τ, where $n\cdot x + \tau = 0$. Here, G will be called a Fermi operator (compare with "e^D" in Refs. 12 and 17). This splitting will not lead to a Lorentz-invariant theory unless the Fermi operator G can be specified in a Lorentz-invariant manner. Since all properties on a spacelike surface τ can be specified using Lorentz invariants and the timelike normal n which defines τ (i.e., $n\cdot x + \tau = 0$ implies $x\in\tau$), then the evolution of the state equations will proceed in a path-independent manner by assuming that the timelike normal n takes the same form (same coefficients) in all Lorentz frames.[5,20] The Fermi operator G is most often used to make the subsidiary condition stationary as time proceeds (i.e., G is used to maintain the gauge symmetry). In Ref. 5, it is proposed that G also be utilized to mediate the splitting of the boson masses into "interaction" and "field" masses (i.e., G also be used to help incorporate renormalization group symmetries).

The presence of G will, in general, allow the specification of the field equations on only one surface. It will then be necessary to check time derivatives of some of the field equations (notably, the gauge conditions) to insure that they do not drift as time progresses. Therefore, G represents a fundamental change in the interaction operator W_{IH} connecting the Heisenberg and interaction pictures unless $(d/d\tau)\mathsf{G} = 0$.

Two additional useful pictures are suggested by the

splitting of U_H. They will be designated as the "intermediate Heisenberg" (IH) and "intermediate interaction" (II) pictures, defined by $\zeta^{IH} = U_{IH}\zeta U_{IH}^{-1}$ and $\zeta^{II} = G\zeta G^{-1}$, respectively. These intermediate pictures are related to the interaction and Heisenberg pictures as shown in the following diagram:



The superscript on G implies that $G^{IH} = U_{IH}GU_{IH}^{-1}$. The choice of the operator W_{IH}, which appears in the Schrödinger equation for U_{IH}, reflects a "best guess" for the interaction.

The connection

$$L_H(\zeta^H,\partial\zeta^H) = L_{free}((U_H^{-1}\zeta^H U_H),\partial(U_H^{-1}\zeta^H U_H)) = L_{free}(\zeta,\partial\zeta)$$

implies that the action may be written as a composite of fields from different pictures. In particular, it has been shown[5] that for the class of meson fields which have been studied, L takes the familiar "Heisenberg" form[4] provided that the spinor fields are expressed in the IH picture while the vector fields are expressed in the Heisenberg picture.

IV. SPINOR FIELD CHOICE

N_S spinor fields ψ_j are chosen which satisfy the usual free-field commutation relations and field equations, i.e.,

$$(\gamma\cdot\partial + m_{S_j})\psi_j = 0,$$

$$\{\psi_j(x),\bar{\psi}_k(y)\} = i(\delta_{ij}/c_{S_j})S(x-y;m_{S_j}),$$

$$\{\psi_j,\psi_k\} = 0,$$

$$\bar{\psi}_j(x) = \psi_j(x)^\dagger\beta,$$

where $\beta = in\cdot\gamma$ and $n\cdot x + \tau = 0$. These may be generated from the Lagrangian density[4]

$$L_{S0} = \sum_{j=1}^{N_S} c_{S_j}\bar{\psi}_j(x)\left(-\frac{1}{2}\gamma\cdot\overleftrightarrow{\partial} + m_{S_j}\right)\psi_j(x).$$

To simplify the notation, it is convenient to introduce $4N_S\times 4N_S$ matrices

$$\gamma_\mu = \begin{bmatrix} \gamma_\mu & & & 0 \\ & \gamma_\mu & & \\ & & \ddots & \\ 0 & & & \gamma_\mu \end{bmatrix},$$

$$C_S = \begin{bmatrix} c_{S_1}1_4 & & & 0 \\ & c_{S_2}1_4 & & \\ & & \ddots & \\ 0 & & & c_{S_{N_S}}1_4 \end{bmatrix},$$

and

$$M_S = \begin{bmatrix} m_{S_1}1_4 & & & 0 \\ & m_{S_2}1_4 & & \\ & & \ddots & \\ 0 & & & m_{S_{N_S}}1_4 \end{bmatrix},$$

and $1\times 4N_S$ and $4N_S\times 1$ spinors

$$\Psi(x) = \begin{bmatrix} \psi_1(x) \\ \psi_2(x) \\ \vdots \\ \psi_{N_S}(x) \end{bmatrix}$$

and

$$\bar{\Psi}(x) = \Psi(x)^\dagger\beta = [\psi_1(x),\psi_2(x),\ldots,\psi_{N_S}(x)],$$

where $\beta = in\cdot\gamma$. The field equations, commutation relations, and Lagrangian density may be written as follows:

$$(\gamma\cdot\partial + M_S)\Psi(x) = 0, \tag{1}$$

$$\bar{\Psi}(x)(-\gamma\cdot\overleftarrow{\partial} + M_S) = 0, \tag{2}$$

$$L_{S0}(x) = \bar{\Psi}(x)C_S(-\tfrac{1}{2}\gamma\cdot\overleftrightarrow{\partial} + M_S)\Psi(x). \tag{3}$$

It will not be necessary to restrict C_S and M_S to the forms given previously, but

$$[C_S,\gamma] = [M_S,\gamma] = 0 \tag{4}$$

will be necessary to maintain the transformation properties of the bilinear covariants.

V. BOSON FIELD CHOICE

The boson fields are chosen by parameterization of the commutation relations such that the following conditions can be met.

(B1) The limit as $m_B \to 0$ is explicitly not taken.

(B2) The Yennie and Landau gauge fields may be obtained in the zero-mass limit.

(B3) The Proca and the Coester fields are included in the set.

(B4) The fields all satisfy the Lorentz condition in the weak-covergence limit (i.e., $\langle\partial\cdot(\text{boson field})\rangle = 0$).

(B5) Invariance under interchange of the fields reduces to special gauge invariance as a weak-convergence relation in the zero-mass limit. The justification for the above restrictions will be given before presenting the fields.

The zero-mass limits are avoided so that only infrared-convergent theories will be considered[4] to focus concentration solely on the ultraviolet divergences in theories that are otherwise finite. Splitting the masses into "interaction" masses and "field" masses to remove the infrared divergences from the propagators introduces both ultraviolet problems into the time-ordered exponential via the boson mass counterterms,[4] and serious problems with the time-independence of the subsidiary condition even in the weak-convergence limit.[5,20]

Fields which reduce to the Yennie and Landau gauge fields in zero-mass limits should be included to accommo-

date the following methodology: transform to the Yennie gauge to minimize the infrared divergences and remove them[2]; then transform to the Landau gauge to minimize the ultraviolet divergences and remove them.[21]

The Coester fields should be included because they are the only massive vector-meson fields with well-defined commutation relations in the zero-mass limit, and because they reduce to the Feynman-gauge photon fields in the zero-mass limit.[22] The Proca fields should be included because they give results which are equivalent to those obtained by using the Coester field,[23] and because they satisfy the Lorentz condition as an operator relation, and because they are the limiting case (for the parametrization to be considered here) such that "ghost fields" are not required.[5,24]

The Coester fields satisfy the Lorentz condition in the weak-convergence limit when indefinite-metric methods are used.[23,25] Indefinite-metric methods are desirable because they are relativistically invariant.[20] This is extremely important in a formalism that is not necessarily covariant.[5] Indefinite-metric methodology has the consequence of restricting all (except the Proca field) of the pretheory boson fields to satisfy the subsidiary condition as weak-convergence relations.

If all of the fields satisfy the subsidiary condition at least as weak-convergence relations, then the zero-mass limit of an interchange of fields in this set mimics special gauge invariance. Every example of special gauge invariance is not included; excluded are those fields whose commutator relations dictate that additional "ghost" fields be present.[24]

N_B vector fields $V_k(x;\eta_k)\mu, 1 \leqslant k \leqslant N_B$, are chosen such that they satisfy the following conditions:

$$(\Box - m_{B_k}^2) V_k(x;\eta_K)\mu = 0,$$

$$[V_K(x;\eta_k)\mu, V_1(y;\eta_1)_\nu] \tag{5}$$

$$= -i(\delta_{k1}/c_{B_k})(g_{\mu_\nu} - (\eta_k/m_{B_k}^2)\partial_\mu \partial_\nu)\Delta(x-y;m_{B_k}), \tag{6}$$

$$\langle \partial \cdot V_k(x;\eta_k) \rangle = 0. \tag{7}$$

The Proca and Coester fields are obtained when $\eta_k = 0$ and $\eta_K = 1$. The Feynman, Landau, and Yennie gauge fields are obtained in the zero-mass limits when $\eta_k = 0, -1, -2$. If $V_k(x;\eta_k)$ has only four independent canonical components, then $V_k(x;\eta_k)$ may be written as follows:

$$V_k(x;\eta_k)_\mu = C_k(x)_\mu + \epsilon_k \partial_\mu \partial \cdot C_k(x), \tag{8}$$

where

$$\epsilon_k = (-1 \pm \sqrt{1-\eta_k})/m_{B_k}^2.$$

The commutators take the following form:

$$[V_k(x;\eta_k)_\mu, V_1(x;\eta_1)_\lambda]$$

$$= -i(\delta_{k1}/c_{B_k})(g_{\mu_\lambda} - [(1-(1 + m_{B_k}^2\epsilon_k)^2)/m_{B_k}^2]\partial_\mu \partial_\lambda)\Delta(x-y;m_{B_k}). \tag{9}$$

If $V_k(x;\eta_k)$ remains uncharged ($\eta_k \leqslant 1$) and if $V_k(x;0)_\mu = C_k(x)_\mu$, (see Ref. 4) then

$$\epsilon_k = (-1 + \sqrt{1-\eta_k})/m_{B_k}^2. \tag{10}$$

Self-adjoint fields with $\eta > 1$ can be constructed in the following fashion:

$$V_k(x;\eta_k)_\mu = V_k(x;\eta_k')_\mu + \partial_\mu B_k(x),$$

where

$$\eta_k' \leqslant 1,$$

$$(\Box - m_{B_k}^2) B_k(x) = 0,$$

$$[B_k(x), V_1(y;\eta_1)\mu] = 0,$$

and

$$[B_k(x), B_1(y)] = -i(\delta_{k1}\delta_k/c_{B_k}m_{B_k}^2)\Delta(x-y;m_{B_k}).$$

The commutation relations imply

$$\eta_k = 1 - (1 + m_{B_k}^2\epsilon_k)^2 + \delta_k$$

and the subsidiary condition becomes

$$\partial \cdot V_k(x;\eta_k) = \sqrt{1-\eta_{kk}'}\,\partial \cdot C_k(x) + m_{B_k}^2 B_k(x).$$

Additional Fock space restrictions will be necessary to include the "ghost" fields $B_k(x)$ (see Ref. 24). Since limiting the number of independent canonical components to not more than four includes the required fields, the ghost fields will not be included.

When $\eta_k < 1$, the expression for $V_k(x;\eta_k)$ in terms of $C_k(x)$ can be inverted to give

$$C_k(x)_\mu = V_k(x;\eta_k)_\mu + [\epsilon_k/(1 + m_{B_k}^2\epsilon_k)]\,\partial_\mu\,\partial \cdot V_k(x;\eta_k). \tag{11}$$

Therefore, any of these fields except the Proca field can be used as a basis. The Proca fields will not form an acceptable basis for construction of the other fields because they have only three independent components, while the other fields have four independent canonical components.

The free-field Lagrangian density $\mathsf{L}_{B0}(x)$ may be chosen as

$$\mathsf{L}_{B0}(x) = \sum_{k=1}^{N_B} c_{B_k}\mathsf{L}_{B_k}(x), \tag{12}$$

where

$$\mathsf{L}_{B_k}(x) = (-\tfrac{1}{2}F_k(x)^{\mu\lambda}F_k(x)_{\mu\lambda} - \tfrac{1}{2}m_{B_k}^2 V_k(x;\eta_k)\cdot V_k(x;\eta_k) - \tfrac{1}{2}(\partial \cdot V_k(x;\eta_k))(\partial \cdot V_k(x;\eta_k))). \tag{13}$$

Up to a four-divergence, $\mathsf{L}_{B0}(x)$ does not depend upon the value of η_k because it is chosen to produce only one result[5]: namely,

$$(\Box - m_{B_k}^2) V_k(x;\eta_k)_\mu = 0.$$

VI. FINAL COMMENTS

More detail can now be given on the construction of QED as an ultraproduct since the pretheory boson fields have been presented. If η_k is restricted such that $\eta_k \leqslant 1$, the set of real values which η_k may take is still an infinite set which lacks a "$\leqslant$" lower bound contained in the set. Therefore, a mathematical filter F_η can be defined based on the acceptable values of η_k such that F_η is the set of all subsets f_{η_k} of the real numbers such that each f_{η_k} contains all real numbers less than or equal to η_k. If the set of all true statements about a neutral meson field of mass m_{B_k} and commutator parameter η_k is labeled as the structure Q_{η_k}, it is well

known that (in principle) it is possible to construct (using an ultrafilter containing F_η as a subset) a logically consistent structure (the ultraproduct) U such that U contains all statements which are true in "most" of the structures Q_{η_k} (see Ref. 10). The zero-mass limit of U should, if it exists, be a candidate for QED. Since the filter F_η is free, it will not necessarily be true that any single one of the structures reduces unambiguously to QED in the zero mass limit (that QED is not in the original collection of structures would be expected because the structure Q_{η_k} would contain η_k-dependent statements).

Heisenberg picture relations have the advantage of describing properties of the entire expansion for the interaction U_H. Some specific examples for the interaction operator $W(\tau)$ will be discussed in the next article.[1]

[1]M. C. Yoakam, "Practical pretheories of QED. II. Choosing the interaction," J. Math. Phys. **26**, 2639 (1985).

[2]H. M. Fried and D. R. Yennie, "New techniques in the Lamb shift calculation," Phys. Rev. **112**, 1391 (1958).

[3]D. Zwanziger, "Physical states in quantum electrodynamics," Phys. Rev. D **14**, 2570 (1976).

[4]See any text on quantum field theory, such as J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Springer, New York, 1980); or N. N. Bogoluibov and D. V. Shirkoff, *Introduction to the Theory of Quantized Fields* (Wiley, New York, 1980).

[5]M. C. Yoakam, Ph. D. thesis, Purdue University, 1983.

[6]W. Pauli and F. Villars, "On the invariant regularization in relativistic quantum theory," Rev. Mod. Phys. **21**, 434 (1949).

[7]S. N. Gupta, "Quantum electrodynamics with auxiliary fields," Proc. Phys. Soc. (London) Ser. A **66**, 129 (1953).

[8]E. R. Speer, "Analytic renormalization," J. Math. Phys. **9**, 1404 (1968).

[9]E. C. G. Stueckelberg and A. Petermann, "La normalisation des constantes dans la theorie des quanta," Helv. Phys. Acta **26**, 499 (1953).

[10]A. Robinson, *Non-Standard Analysis* (North-Holland, Amsterdam, 1966).

[11]A. Voros, "Introduction to nonstandard analysis," J. Math Phys. **14**, 292 (1973).

[12]K. Haller and L. F. Landovitz, "Subsidiary condition in quantum electrodynamics," Phys. Rev. **168**, 1749 (1968).

[13]K. Haller and L. F. Landovitz, "Equivalence of the Coulomb gauge and the reformulated Lorentz gauge," Phys. Rev. **182** 1922 (1969).

[14]K. Haller and L. F. Landovitz, "Renormalization constants, wave functions, and energy shifts in the Coulomb and Lorentz gauges," Phys. Rev. D **2**, 1498 (1970).

[15]K. Haller, "Gupta–Bleuler condition and infrared-coherent states," Phys. Rev. D **18**, 3045 (1978).

[16]K. Haller and R. B. Sohn, "Interaction of Schrödinger electrons and photons," Phys. Rev. A **20**, 1541 (1979).

[17]K. Haller, "Manifestly covariant and manifestly unitary formulations of massive vector-boson theory in indefinite-metric spaces," Phys. Rev. D **8**, 1796 (1973).

[18]J. Schwinger, "Quantum electrodynamics. I. A covariant formulation," Phys. Rev. **74**, 1439 (1948).

[19]J. Schwinger, "Quantum electrodynamics. II. Vacuum polarization and self-energy," Phys. Rev. **75**, 651 (1949).

[20]F. J. Belinfante, "The interaction representation of the Proca field," Phys. Rev. **76**, 66 (1949).

[21]L. D. Landau and I. M. Khalatnikoff, "The gauge transformation of the Green's functions for charged particles," Sov. Phys. JETP **2**, 69 (1956).

[22]F. Coester, "Quantum electrodynamics with nonvanishing photon mass," Phys. Rev. **83**, 798 (1951).

[23]E. C. G. Steuckelberg, "Die Wechselwirkungskrafte in der Elektrtrodynamik und in der Feldtheorie der Kernkrafte," Helv. Phys. Acta **11**, 225 (1938).

[24]K. Just, "The Lorentz condition in quantum theory," Nuovo Cimento **38**, 400 (1965).

[25]S. N. Gupta, "Theory of longitudinal photons in quantum electrodynamics," Proc. Phys. Soc. (London) **63**, 681 (1950).

Practical pretheories of QED. II. Choosing the interaction

M. C. Yoakam
Physics Department, Purdue University, West Lafayette, Indiana 47906

(Received 13 March 1984; accepted for publication 5 April 1985)

An interaction of the $j\cdot A$ form is introduced. The Coester transformation, which connects the Proca and the Coester fields, is extended to the pretheory boson fields when the boson mass counterterms are absent. The Fermi transformation, used to make the Lorentz condition stationary in the weak-convergence limit, is shown to be extendable to the pretheories (sans the freedom of commutator choice, which is usually associated with it). The familiar consequences of including a Fermi transformation (the Heisenberg picture images of the spinor fields do not transform as spinors) are retained, but a suitable limitation on the choice of the spinor charge and mass counterterms will allow the Heisenberg picture forms which are bilinear in the spinor operators to retain their transformation properties. In particular, familiar choices for L_{int} require that the spinor fields be expressed in the intermediate-Heisenberg picture. The introduction of boson mass counterterms into the practical pretheories is shown to give infinite gauge shifts which are independent of the zero-mass limit. Sufficient conditions for a modified Fermi operator are presented, and an "obvious" candidate is eliminated.

I. INTRODUCTION

In the first article of this series,[1] modifications to the variational principle were discussed, and the choices of the interaction picture pretheory field operators were presented. In this article, it will be shown that many of the early results for quantum electrodynamics (QED) are not only indigeneous to zero-mass theories, but are general properties of a wide class of heavy-meson theories. Interactions that are familiar in appearance, but that are not ultraviolet-convergent (because of the absence of boson mass counterterms), are considered first. The spinor mass counterterms are included since they introduce no additional complications.

The boson mass counterterms are then introduced and severe problems with the subsidiary condition[2] are shown to follow from the separation of the "physical" (Heisenberg picture) boson masses into "field" masses and "interaction" masses required for constructing an ultraviolet-convergent time-ordered exponential. This demonstrates that the renormalization and gauge symmetries may not be considered members of the symmetry group for Poincaré-invariant theories with interactions of the form

($j\cdot A$ + mass counterterms

+ multiples of free-field Lagrangians).

The difficulties associated with the inclusion of the boson mass counterterms have not yet been successfully resolved.[2] In this paper, the origin of the problem (subsidiary condition becomes time dependent) and the reason why a "standardized" Fermi operator cannot completely resolve the problem (no Bose–Bose terms) are demonstrated, and the restrictions which a modified Fermi operator must satisfy (if the Lagrangian approach to QED is to be successful) are presented. Ultimate success would imply that the symmetry group of the pretheories includes the Poincaré, gauge, and renormalization group symmetries so that a well-defined Schrödinger picture may be constructed. That would allow discussions of the scattering of free states into bound states to be treated on an equal footing with discussions of the scattering of free states into free states (usual scattering calculations) and bound-state energy-level shifts (Lamb shifts, anomalous magnetic moments).[3]

The notation and definitions presented in the first paper of this sequence (1) will be used throughout this paper. As before, it will be assumed that no zero-mass limits have been taken (i.e., the theory is infrared convergent) so that attention may be focused on the ultraviolet divergences.

II. THE INTERACTION

The operator $U_H(\tau,\tau_0)$ which connects the interaction picture with a constant-state picture is defined by $U_H(\tau,\tau_0) = G(\tau)^{IH} U_{IH}(\tau,\tau_0)$, where $U_{IH}(\tau,\tau_0)$ is a time-ordered exponential satisfying

$$i\frac{d}{d\tau}U_{IH}(\tau,\tau_0) = U_{IH}(\tau,\tau_0)W(\tau), \tag{1}$$

where

$$W(\tau) = \int_\tau d\sigma(j(x)\cdot C(x) - \bar{\Psi}(x)C_S(\delta M_S)\Psi(x)),$$

$$j(x) = -i\bar{\Psi}(x)C_S E\gamma\Psi(x),$$

and $G(\tau)^{IH}$ is an operator yet to be determined. The "charge" matrix E and the spinor mass-counterterm matrix δM_S are assumed to be arbitrary except that

$$[E,\gamma^\mu] = [\delta M_S,\gamma^\mu] = 0. \tag{2}$$

In particular, $\partial\cdot j(x) = 0$ is not required.

III. EXTENTION OF THE WORK OF COESTER AND STUEKELBERG TO THE PRETHEORY BOSON FIELDS

The interaction operator is expressed in terms of the Coester fields instead of the general pretheory boson fields. If a canonical transformation $G_C(\tau)$ is introduced by

$$U_{IH}(\tau,\tau_0) = G_C(\tau)^{IH} U'_{IH}(\tau,\tau_0),$$

then the Schrödinger equation for U'_{IH} is given by

 0022-2488/85/102639-04$02.50

$$i\frac{d}{d\tau}U'_{IH}(\tau,\tau_0)=U'_{IH}(\tau,\tau_0)W'(\tau)$$
$$=U'_{IH}(\tau,\tau_0)\Big(G_C(\tau)^{-1}W(\tau)G(\tau)$$
$$+iG_C(\tau)^{-1}\frac{d}{d\tau}G_C(\tau)\Big). \tag{3}$$

$W'(\tau)$ will take the form

$$W'(\tau)=\int_\tau d\sigma\Big(j(x)\cdot\sum_{k=1}^{N_B}c_{B_k}V_k(x;\eta_k)$$
$$-\bar\Psi(x)C_S\delta M_S\Psi(x)-\lambda n\cdot j(x)n\cdot j(x)\Big)$$

for the general pretheory boson fields when

$$G_C(\tau)=e^{i\Sigma(\tau)}, \tag{4}$$

where

$$\Sigma(\tau)=\int_\tau d\sigma\, n\cdot j(x)\sum_{k=1}^{N_B}c_{B_k}\epsilon_k\,\partial\cdot C_k(x),$$
$$\lambda=\frac{1}{2}\sum_{\substack{k=1\\ m_{B_k}\neq 0}}^{N_B}\frac{c_{B_k}\eta_k}{m_{B_k}^2}, \tag{5}$$
$$[E,(M_S+\delta M_S)]=0.$$

If Eq. (5) is not satisfied, then W' would be amended by the addition of the infinite series

$$-\sum_{n=1}^{\infty}\frac{(-i)^n}{n!}\bar\Psi(x)[E;(M_S+\delta M_S)]^n\Psi(x),$$

where $[A;B]^n=[A,[\cdots[A,[A,B]]\cdots]]$ (see Ref. 2). The four-spinor interaction is required to make Eq. (3) integrable.[2,4]

Since matrix elements are unaffected by cannonical transformations, nothing is lost by using the simple choice for $W(\tau)$. This is an extension of the connection between the Coester and Proca fields[5,6] ($\eta_k=1$) to a connection between the Coester and the general pretheory fields ($\eta_k\leqslant 1$).

IV. FAILURE IN THE FIELD EQUATIONS

If $G(\tau)=1$, then the IH and H pictures coincide. The spinor operators satisfy the following equations[2]:

$$(\gamma\cdot\partial+(M_S+\delta M_S))\Psi(x)^{IH}=iE\gamma\cdot C(x)^{IH}\Psi(x)^{IH}, \tag{6}$$
$$\bar\Psi(x)^{IH}(-\gamma\cdot\overleftarrow{\partial}+(M_S+C_S\delta M_S C_S^{-1}))$$
$$=i\bar\Psi(x)^{IH}(C_SEC_S^{-1})\gamma\cdot C(x)^{IH}, \tag{7}$$
$$\partial\cdot j(x)^{IH}=i\bar\Psi(x)^{IH}C_S[(M_S+\delta M_S),E]C_S^{-1}\Psi(x)^{IH}; \tag{8}$$

$\bar\Psi^{IH}$ and Ψ^{IH} transform as spinors, and if Eq. (5) is satisfied, then $\partial\cdot j(x)^{IH}$ is stationary,[2] i.e.,

$$\frac{d}{d\tau}(\partial\cdot j(x)^{IH})=n\cdot\partial(\partial\cdot j(x)^{IH})=0.$$

The difficulties which arise for the pretheory boson operators in the IH picture are the following[2]: $V_k(x;\eta_k)^{IH}$ is not a vector operator; $V_k(x;\eta_k)^{IH}$ does not satisfy a "recognizable" field equation;

$$\partial\cdot V_k(x;\eta_k)^{IH}=(\partial\cdot V_k(x;\eta_k))^{IH}-\epsilon_k\partial_\perp\cdot j(x)^{IH};$$
$$\langle\partial\cdot V_k(x;\eta_k)^{IH}\rangle^{IH}\neq\langle\partial\cdot V_k(x;\eta_k)\rangle;$$

and

$$\Big(\frac{d}{d\tau}\Big)(\partial\cdot V_k(x;\eta_k)^{IH})\neq 0.$$

The first two problems can be handled using methods analogous to the methods used for the Proca field.[4] If $A_k(x;\eta_k)^{IH}$ is defined by

$$A_k(x;\eta_k)^{IH}_\mu=V_k(x;\eta_k)^{IH}_\mu+\epsilon_k n_\mu n\cdot j(x)^{IH},$$
$$\epsilon_k=(-1+\sqrt{1-\eta_k})/m_{B_k}^2,$$

then $A_k(x;\eta_k)^{IH}$ is a vector operator and $A_k(x;\eta_k)^{IH}$ satisfies the following conditions:

$$(\Box-m_{B_k}^2)A_k(x;\eta_k)^{IH}_\mu-\partial_\mu\,\partial\cdot A_k(x;\eta_k)^{IH}$$
$$=-j(x)^{IH}_\mu-\sqrt{1-\eta_k}\,n_\mu n\cdot j(x)^{IH}, \tag{9}$$
$$\partial\cdot A_k(x;\eta_k)^{IH}=(\partial\cdot V_k(x;\eta_k))^{IH}-\epsilon_k\partial\cdot j(x)^{IH}, \tag{10}$$
$$\frac{d}{d\tau}(\partial\cdot A_k(x;\eta_k)^{IH})=n\cdot\partial(\partial\cdot A_k(x;\eta_k)^{IH})$$
$$=(n\cdot\partial\partial\cdot V_k(x;\eta_k))^{IH}$$
$$-(1+\epsilon_k\Box)n\cdot j(x)^{IH}. \tag{11}$$

Condition (11) is particularly troublesome because, although indefinite-metric methods may be applied on one timelike surface τ to fix the subsidiary condition on that surface, as time progresses the subsidiary condition will drift away from the weak-convergence Lorentz condition, introducing spurious interactions which are gauge dependent.[4,7,8] This problem requires the introduction of a nontrivial Fermi operator [i.e., $G(\tau)\neq 1$].

V. FERMI OPERATORS

The gauge problems for the boson operators can be corrected by introducing additional terms into the interaction with a Fermi operator $G(\tau)$ chosen to be the operator $G_F(\tau)$ defined as follows:

$$G_F(\tau)=e^{i\Sigma_F(\tau)}, \tag{12}$$

where

$$\Sigma_F(\tau)=\int_\tau d\sigma\, V_C(x)B(x),$$
$$V_C(x)=-\frac{1}{4\pi}\int_\tau d\sigma'\,\frac{n\cdot j(x')}{|x_\perp-x'_\perp|},$$
$$B(x)=\sum_{k=1}^{N_B}c_{B_k}(a_k\,\partial_\perp\cdot V_k(x;\eta_k)+b_k\,\partial\cdot V_k(x;\eta_k))$$
$$=-\sum_{k=1}^{N_B}c_{B_k}\,\partial_\perp\cdot C_k(x)=-\partial_\perp\cdot C(x).$$

The restriction on $B(x)$ must be stronger than the usual assumption[4] that b_k is arbitrary to make the results valid for all of the pretheory boson fields (i.e., for all $\eta_k\leqslant 1$).

The vector operator in the Heisenberg picture is now

$$V_k(x;\eta_k)^H=A_k(x;\eta_k)^{IH},$$

and $V_k(x;\eta_k)^H$ satisfies the following properties[2]:

$$(\Box-m_{B_k}^2)V_k(x;\eta_k)^H_\mu-\partial_\mu\,\partial\cdot V_k(x;\eta_k)^H$$
$$=-j(x)^H_\mu-\sqrt{1-\eta_k}\,n_\mu n\cdot j(x)^H, \tag{13}$$
$$\partial\cdot V_k(x;\eta_k)^H=(\partial\cdot V_k(x;\eta_k))^H-\epsilon_k\,\partial\cdot j(x)^H, \tag{14}$$
$$\frac{d}{d\tau}(\partial\cdot V_k(x;\eta_k)^H)=n\cdot\partial\,\partial\cdot V_k(x;\eta_k)^H$$
$$=(n\cdot\partial\,\partial\cdot V_k(x;\eta_k))^H. \tag{15}$$

Equations (14) and (15) guarantee that indefinite-metric methods may be used to fix the Lorentz condition in the weak-convergence limit for all times τ in both the Heisenberg and interaction pictures. The interaction operator $W(\tau)$ becomes

$$W(\tau)=\int_\tau d\sigma\Big\{j(x)\cdot C(x)-\bar\Psi(x)C_S\delta M_S\Psi(x)+\chi(x)\partial_\perp\cdot j(x)+V_C(x)n\cdot\partial\,\partial_\perp\cdot C(x)\Big\},\quad(16)$$

$$\chi(x)=-\frac{1}{4\pi}\int_\tau d\sigma'\,\frac{B(x')}{|x_\perp-x'_\perp|},$$

where $[E,M_S+\delta M_S]=0$ and $\Sigma_{k=1}^{N_B}c_{B_k}=0$ have been used.[2] The interaction has become nonlocal in order to obtain the proper boson field equations.

Fermi operators introduce minor problems for the pretheory spinor fields,[2] which are analogous to the ones that are introduced when only the Proca fields are considered.[4] The $\bar\Psi^{\rm H}$ and $\Psi^{\rm H}$ no longer transform as spinors, and they may be expressed as follows:

$$\bar\Psi(x)^{\rm H}=\bar\Psi(x)^{\rm IH}e^{-i\chi(x)^{\rm IH}(C_SEC_S^{-1})},$$

$$\Psi(x)^{\rm H}=e^{i\chi(x)^{\rm IH}E}\Psi(x)^{\rm IH}.$$

If $\Sigma_{k=1}^{N_B}c_{B_k}=0$ and $[E,M_S+\delta M_S]=0$ are used to make the field equations well defined [no $\delta(0)$ terms] and of finite length, then the following field equations are obtained on a single space-time surface τ:

$$(\gamma\cdot\partial+(M_S+\delta M_S))\Psi(x)^{\rm H}=iE\gamma\cdot(C(x)^{\rm H}+\partial\chi(x)^{\rm H})\Psi(x)^{\rm H},\quad(17)$$

$$\bar\Psi(x)^{\rm H}(-\gamma\cdot\overleftarrow{\partial}+(M_S+C_S\delta M_SC_S^{-1}))=\bar\Psi(x)^{\rm H}(i(C_SEC_S^{-1})\gamma\cdot(C(x)^{\rm H}+\partial\chi(x)^{\rm H})).\quad(18)$$

Equations (17) and (18) will not be valid in arbitrary Lorentz frames. This can be avoided, and an effective Lagrangian density which has a familiar appearance can be obtained with the choice

$$\mathsf{L}_{\rm free}(\Psi,\bar\Psi,V)=\mathsf{L}(\mathsf{U}_{\rm IH}^{-1}\Psi^{\rm IH}\mathsf{U}_{\rm IH},\mathsf{U}_{\rm IH}^{-1}\bar\Psi^{\rm IH}\mathsf{U}_{\rm IH},\mathsf{U}_{\rm H}^{-1}V^{\rm H}\mathsf{U}_{\rm H})=\mathsf{L}_{\rm familiar}(\Psi^{\rm IH},\bar\Psi^{\rm IH},V^{\rm H}),$$

i.e., the boson fields are expressed in the Heisenberg picture and the spinor fields are expressed in the IH picture.

The forms of $\bar\Psi^{\rm H}$ and $\Psi^{\rm H}$ imply that $\bar\Psi^{\rm H}\Gamma\Psi^{\rm H}=\bar\Psi^{\rm IH}\Gamma\Psi^{\rm IH}$ provided that C_S, M_S, δM_S, and E all commute with Γ. Therefore, a sufficient condition for the preservation of the transformation properties of a general bilinear covariant Γ is the commutation of these four matrices with γ_μ.

VI. MODIFICATIONS WHEN BOSON MASS COUNTERTERMS ARE PRESENT

The interaction operator $W(\tau)$ will be chosen as follows:

$$W(\tau)=\int_\tau d\sigma\Big(j(x)\cdot C(x)-\bar\Psi(x)C_S\delta M_S\Psi(x)-\sum_{k,l=1}^{N_B}c_{B_k}C_k(x)_\mu\delta M^2_{B_{kl}}C_l(x)^\mu\Big).$$

The results for the IH picture pretheory spinor operators are unchanged because the term added to $W(\tau)$ commutes with them and their derivatives, i.e., $\bar\Psi^{\rm IH}$ and $\Psi^{\rm IH}$ are spinors if $\bar\Psi$ and Ψ were spinors; $\partial\cdot j(x)^{\rm IH}$ is stationary if Eq. (5) is valid; and Eqs. (6)–(8) are still true.

The changes for the boson operators may be summarized by replacing $j(x)$ by an effective current $j_{\rm eff}(x)$ defined as

$$j_{\rm eff}(x)=j(x)-\sum_{l=1}^{N_B}T_{kl}C_l(x),$$

where

$$T_{kl}=(l/c_{B_k})(c_{B_k}\delta M^2_{B_{kl}}+c_{B_l}\delta M^2_{B_{lk}}).$$

The IH picture vector operator $A_k(x;\eta_k)^{\rm IH}$ is now given by

$$A_k(x;\eta_k)^{\rm IH}_\mu=V_k(x;\eta_k)^{\rm IH}_\mu+\epsilon_k n_\mu n\cdot j_{\rm eff}(x)^{\rm IH},$$

and $A_k(x;\eta_k)^{\rm IH}$ satisfies the following conditions:

$$(\Box-m^2_{B_k})A_k(x;\eta_k)^{\rm IH}_\mu-\partial_\mu\,\partial\cdot A_k(x;\eta_k)=-j_{\rm eff}(x)^{\rm IH}-\sqrt{1-\eta_k}\,n_\mu n\cdot j_{\rm eff}(x)^{\rm IH},\quad(9')$$

$$\partial\cdot A_k(x;\eta_k)^{\rm IH}=(\partial\cdot V_k(x;\eta_k))^{\rm IH}-\epsilon_k\,\partial\cdot j_{\rm eff}(x)^{\rm IH},\quad(10')$$

$$\frac{d}{d\tau}(\partial\cdot A_k(x;\eta_k)^{\rm IH})=n\cdot\partial(\partial\cdot A_k(x;\eta_k)^{\rm IH})=(n\cdot\partial\,\partial\cdot V_k(x;\eta_k))^{\rm IH}-(1+\epsilon_k\Box)n\cdot j_{\rm eff}(x)^{\rm IH}.\quad(11')$$

Equations (10') and (11') again imply that state restrictions such as those introduced by indefinite-metric methods[9] will not make the subsidiary condition stationary even in the weak-convergence limit. Equation (11') also implies that attempts to remove the regulation (i.e., taking the limit as $\delta M_B^2\to\infty$) will give infinite gauge shifts as time progresses. The IH picture theory then has no gauge symmetry, and a nontrivial Fermi operator $\mathsf{G}(\tau)$ will be necessary.

The situation will not improve substantially when $\mathsf{G}(\tau)=\mathsf{G}_{\rm F}(\tau)$ because $\mathsf{G}_{\rm F}(\tau)$ will remove the $j(x)$ terms while leaving the mass-counterterm contributions untouched. Therefore, it will be necessary for $\mathsf{G}(\tau)$ to be a nontrivial extension of $\mathsf{G}_{\rm F}(\tau)$.

VII. GENERAL FERMI OPERATORS WHEN BOSON MASS COUNTERTERMS ARE PRESENT

The Fermi operator method will work if a $\mathsf{G}(\tau)$ can be found such that the following conditions are satisfied by all of the pretheory fields:

$V_k(x;\eta_k)^{\rm H}$ is a vector operator if

$$V_k(x;\eta_k)\text{ is a vector operator;}\quad(19)$$

$$\partial\cdot V_k(x;\eta_k)^{\rm H}=(\partial\cdot V_k(x;\eta_k))^{\rm H};\quad(20)$$

$$\frac{d}{d\tau}(\partial\cdot V_k(x;\eta_k)^{\rm H})=\Big(\frac{d}{d\tau}(\partial\cdot V_k(x;\eta_k))\Big)^{\rm H}$$

or

$$n\cdot\partial\,\partial\cdot V_k(x;\eta_k)^{\rm H}=(n\cdot\partial\,\partial\cdot V_k(x;\eta_k))^{\rm H}.\quad(21)$$

Conditions (19)–(21) will be satisfied for all pretheory boson fields when they are satisfied for the Coester fields because $V_k(x;\eta_k)_\mu=C_k(x)_\mu+\epsilon_k\,\partial_\mu\,\partial\cdot C_k(x)$. Since $\mathsf{U}_{\rm H}$ is

a function of τ and τ_0 only, then $\partial_\perp U_H = (\partial + nn\cdot\partial)U_H = 0$ and (21) implies

$$\partial_\mu\, \partial\cdot C_k(x)^H = (\partial_\mu\, \partial\cdot C_k(x))^H.$$

Equations (9'), (10'), and (11') and the relationships $U_H = G^{IH}U_{IH} = U_{IH}G$ and $O^H = U_H O U_H^{-1}$ imply that the Fermi operator $G(\tau)$ will satisfy the following:

$$G(\tau)C_k(x)G(\tau)^{-1} = C_k(x), \tag{22}$$

$$G(\tau)(\partial\cdot C_k(x))G(\tau)^{-1} = \partial\cdot C_k(x), \tag{23}$$

$$G(\tau)(n\cdot\partial\,\partial\cdot C_k(x))G(\tau)^{-1} = n\cdot\partial\,\partial\cdot C_k(x) + n\cdot j_{\rm eff}(x). \tag{24}$$

If $G(\tau)$ is given by

$$G(\tau) = e^{i\Sigma(\tau)},$$

where $\Sigma(\tau) = \Sigma_F(\tau) + \Sigma_B(\tau)$ and $\Sigma_B(\tau)$ is bilinear in the boson operators, then Eqs. (22)–(24) become the following:

$$e^{i\Sigma_B(\tau)}C_k(x)e^{-i\Sigma_B(\tau)} = C_k(x), \tag{22'}$$

$$e^{i\Sigma_B(\tau)}(\partial\cdot C_k(x))e^{-i\Sigma_B(\tau)} = \partial\cdot C_k(x), \tag{23'}$$

$$e^{i\Sigma_B(\tau)}(n\cdot\partial\,\partial\cdot C_k(x))e^{-i\Sigma_B(\tau)} = n\cdot\partial\,\partial\cdot C_k(x) - n\cdot\sum_{l=1}^{N_B} T_{kl}C_l(x). \tag{24'}$$

It should be noted that assuming commutivity of Σ_F and Σ_B is not necessary to obtain Eqs. (22')–(24'), but it is necessary to avoid a proliferation of terms of different operator type (i.e., different products of creation and annihilation operators).[2] The problem is reduced to finding a $\Sigma_B(\tau)$, which commutes with $\partial_\perp\cdot C_k(x)$ and which satisfies (22')–(24'). An appropriate Σ_B has not yet been successfully constructed. One special case which is suggested by the results presented in this article will be discussed below.

The special case is

$$j(x)\to j_{\rm eff}(x) = j(x) - \sum_{l=1}^{N_B} T_{kl}C_l(x).$$

This substitution cannot be used directly in $\Sigma_F(\tau)$ because there is an unsummed boson index in $j_{\rm eff}$. However, if $\Sigma_B(\tau)$ has the form

$$\Sigma_B(\tau) = \int_\tau\int_\tau d\sigma\, d\sigma' \sum_{k,l=1}^{N_B} G_{kl}\left(\frac{n\cdot C_l(x')\partial_\perp\cdot C_k(x)}{|x_\perp - x'_\perp|}\right),$$

then $G_{kl} = -(1/4\pi)T_{kl}$ to satisfy (24'). This choice for $\Sigma_B(\tau)$ satisfies (22') as well and $[\Sigma_B(\tau), \Sigma_F(\tau)] = 0$, but it will not satisfy (23'). In fact, condition (23') will yield

$$G(\tau)(\partial\cdot C_k(x))G(\tau)^{-1} = \partial\cdot C_k(x) + \frac{1}{4\pi}\int_\tau d\sigma' \frac{\sum_{l=1}^{N_B} T_{kl}\,\partial'_\perp\cdot C_l(x')}{|x_\perp - x'_\perp|}.$$

This can be shown to make the gauge condition for even the Proca field nonstationary.[2] Therefore, $\Sigma_B(\tau)$ in this form is not sufficient.

VIII. SUMMARY

The Fermi-operator methodology was first applied to the general pretheory fields in the absence of the boson mass counterterms. It was shown that if the Lagrangian density L is to have the expected appearance with proper spinor and boson fields as the field variables, then the boson fields must be expressed in the Heisenberg picture to preserve the subsidiary condition and, by default, the spinor operators must be expressed in the IH picture. Therefore, canonical commutation relations can no longer be assumed in a single constant-state picture because the images of the interaction picture field variables are not the new canonical field variables in either the IH or the Heisenberg pictures. In particular, it is impossible to reach a constant-state picture from the interaction picture in which well-defined spinor and Bose variables simultaneously exist and satisfy the expected field equations. It was shown that this is not a serious liability because the spinor observables (bilinear covariants) can be maintained. This approximation of the practical pretheories is not ultraviolet convergent because of the absence of the boson mass counterterms.

Amending the IH picture interaction operator by adding boson mass counterterms is shown to cause the subsidiary condition to drift away from the Lorentz condition as time evolves, even in the weak-convergence limit. Removing the regulation will often (possibly always) make these terms infinite (indefinite), so the full theory includes infinite gauge shifts (i.e., gauge invariance is not a good symmetry). Since the gauge failure is independent of the physical masses, no improvement is obtained by taking zero-mass limits. This is a proof that QED is not well defined until the effects of mass splitting (e.g., gauge shifting) are removed by means external (at this time) to the variational methodology, i.e., the theory derived by variational means does not have the proper symmetry group. Conditions (22')–(24') will allow incorporation of the mass-splitting effects within the variational formalism while maintaining the gauge condition in the weak-convergence limit if a suitable Σ_B can eventually be constructed. Infinite gauge shifts in the Heisenberg picture preclude the possibility of constructing a Schrödinger picture because such shifts will appear in the Hamiltonian.[3] Further work will be necessary to determine whether or not these shifts are the sole source of the difficulties associated with simultaneously regulating both the interaction and Schrödinger pictures.

[1] M. C. Yoakam, "Practical pretheories of QED. I. Properties of the interaction picture," J. Math. Phys. **26**, 2634 (1985).

[2] M. C. Yoakam, "Reflections on the simultaneous existence of the Heisenberg and interaction pictures in practical pretheories of QED," Ph.D thesis, Purdue University, 1983.

[3] See any text on quantum field theory, such as J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Springer, New York, 1980) or N. N. Bogoluibov and D. V. Shirkoff, *Introduction to the Theory of Quantized Fields* (Wiley, New York, 1980).

[4] F. J. Belinfante, "The interaction representation of the Proca field," Phys. Rev. **76**, 66 (1949).

[5] F. Coester, "Quantum electrodynamics with nonvanishing photon mass," Phys. Rev. **83**, 798 (1951).

[6] E. C. G. Stueckelberg, "Die Wechselwirkungskrafte in der Elektrodynamik und in der Feldtheorie der Kernkrafte.," Helv. Phys. Acta **11**, 225 (1938).

[7] K. Haller and L. F. Landovitz, "Subsidiary condition in quantum electrodynamics," Phys. Rev. **168**, 1749 (1968).

[8] K. Haller and L. F. Landovitz, "Equivalence of the Coulomb gauge and the reformulated Lorentz Gauge," Phys. Rev. **182**, 1922 (1969).

[9] S. N. Gupta, "Theory of longitudinal photons in quantum electrodynamics," Proc. Phys. Soc. (London) **63**, 681 (1950).

Hopf bifurcation in Yang–Mills mechanics

Ulrich Kursawe
Institut für Theoretische Physik der Universität zu Köln, Zülpicherstrasse 77, D-5000 Koln 41, Federal Republic of Germany

Edward Malec
Institute of Nuclear Physics, Radzikowskiego 152, 31-342 Kraków, Poland

(Received 12 November 1984; accepted for publication 12 April 1985)

The periodic non-Abelian space-independent solutions to Yang–Mills SU(2) equations are studied. Hopf bifurcation is shown to appear. New analytic solutions generating a simple Abelian source are found and their stability is discussed. The source can be produced by a fermion field; a self-consistent solution of Yang–Mills–Dirac equations is presented.

I. INTRODUCTION

Our objective is to study solutions of the so-called Yang–Mills mechanics.[1] Up to now, some numerical investigations were made, which suggested the existence of at least a countable number of periodic oribits.[1,2] We will complement those studies and find analytic approximate and exact solutions.

The program of the paper is the following. At first, we will do a qualitative analysis, i.e., we will prove the existence of periodic solutions, using bifurcation theory. Next, some of these solutions are studied numerically. This is done in Secs. II and IV. The kind of bifurcation studied is known as Hopf bifurcation.

In Sec. III we present in detail an exact Weierstrass function solution, which corresponds to an Abelian source. Section V is devoted to stability analysis. Stationary solutions appear to be unstable, which is not surprising, since Yang–Mills mechanics is known to be chaotic.[1-3] What is new is the existence of such perturbations that do not lead to chaos.

The above solutions generate a static Abelian charge. In Sec. VI we present a fermion field, which generates just such a current.

The last section contains a short summary and speculations about the physical relevance of the obtained results.

II. BIFURCATION SOLUTIONS IN YANG–MILLS MECHANICS

The Yang–Mills SU(2) equations read

$$\partial_\mu F^{a\mu\nu} + f^{abc} A^b_\mu F^{c\mu\nu} = j^{a\nu}, \tag{1}$$

where $F^{a\mu\nu} = \partial^\mu A^{a\nu} - \partial^\nu A^{a\mu} + f^{abc} A^{b\mu} A^{c\nu}$, f^{abc} denotes the completely antisymmetric structure constants of the SU(2) algebra, Greek space-time indices range from 0 to 3 and Latin upper isospin indices $a,b,c,\ldots$ range from 1 to 3. Assuming $A^a_0 = 0, A^3_i = A^a_3 = 0$, defining

$$A^1_1 = s, \quad A^2_1 = z, \quad A^1_2 = u, \quad A^2_2 = w, \tag{2}$$

and supposing that gauge potentials A^a_μ depend only on time, Eqs. (1) can be reduced to the following system of ordinary differential equations:

$$s\dot{z} - \dot{s}z + u\dot{w} - \dot{w}u = \delta^{a3}\,\text{const}, \tag{3a}$$

$$\ddot{s} = -w(sw - uz), \quad \ddot{u} = z(sw - uz),$$

$$\ddot{z} = u(sw - uz), \quad \ddot{w} = -s(sw - uz). \tag{3b}$$

These equations describe the so-called Yang–Mills mechanics[1,2] [but note that in Ref. 2 Eq. (3a) is lacking].

In this section we will consider only the evolution equations (3b).

Further simplification follows after introducing new variables $\theta_1, \theta_2, r_1, r_2$ such that

$$\begin{aligned} (1/\sqrt{2})(s+w) &= r_1 \cos\theta_1, \\ (1/\sqrt{2})(u-z) &= r_1 \sin\theta_1, \\ (1/\sqrt{2})(u+z) &= r_2 \sin\theta_2, \\ (1/\sqrt{2})(s-w) &= r_2 \cos\theta_2. \end{aligned} \tag{4}$$

Inserting this into (3b) one obtains[2]

$$r_1^2\dot{\theta}_1 = \text{const} = L_1, \quad r_2^2\dot{\theta}_2 = \text{const} = L_2, \tag{5}$$

$$\ddot{r}_1 - L_1^2/r_1^3 + (r_1/2)(r_1^2 - r_2^2) = 0,$$

$$\ddot{r}_2 - L_2^2/r_2^3 + (r_2/2)(r_2^2 - r_1^2) = 0. \tag{6}$$

Remark 1: These equations possess a simple solution for $L_1 = L_2 = L$. It reads

$$r_1 = r_2 = (1/D)((D^2 t + C)^2 + L^2)^{1/2},$$

$$\theta = (L/D)\arctan((D^2/L^2)t + (C/L^2)),$$

where D and C are integration constants. However, it is rather trivial since its holonomy group is U(1). Thus it represents a solution of Maxwell theory embedded into SU(2).

Our objective is to find static solutions of Eqs. (6) and next to look for periodic solutions that bifurcate from these solutions. At first, note that Eqs. (6) do not possess any static solutions when both constants are nonzero ($L_1, L_2 \neq 0$). So, let us study Eqs. (6) under the assumption that $L_1 = 0$, $L_2 \neq 0$. Now they read

$$\ddot{r}_1 + (r_1/2)(r_1^2 - r_2^2) = 0, \tag{7}$$

$$\ddot{r}_2 - L_2^2/r_2^3 - (r_2/2)(r_1^2 - r_2^2) = 0. \tag{8}$$

There exists a static solution

$$r_1 = 0, \quad r_2 = 2^{1/6} L_2^{1/3}. \tag{9}$$

To save some writing we define

$$\alpha = 2^{1/6} L_2^{1/3}, \quad r_1 = x, \quad r_2 = \alpha + y. \tag{10}$$

With these conventions the equations of motion become

$$\ddot{x} + (x/2)(x^2 - (y+\alpha)^2) = 0, \tag{11a}$$

$$\ddot{y} - \alpha^6/2(y+\alpha)^3 - [(y+\alpha)/2](x^2 - (y+\alpha)^2) = 0. \tag{11b}$$

 0022-2488/85/102643-05$02.50

We now show the existence of nontrivial solutions satisfying periodic boundary conditions $x(0) = y(0) = x(w(\alpha)T) = y(\omega(\alpha)T) = 0$, where $\omega(\alpha)$ is a frequency and $t \in (0,2T)$. The period $2T$ is arbitrary.

Theorem 1: The system (11) possesses one and only one nontrivial solution which bifurcates from $(x = 0, y = 0)$ at values $\alpha_n = n\pi/\sqrt{3}T$ and satisfies periodic boundary conditions $x(0) = y(0) = x(\omega(\alpha)T = \pi) = y(\omega(\alpha)T = \pi) = 0$. For α sufficiently close to $\alpha_n = n\pi/\sqrt{3}T$, the bifurcating solution is given by the following formulas:

$$x(t,\alpha) = 0,$$

$$y(t,\alpha) = \mu \sin(\sqrt{3}\alpha_n t) + O(\mu^2), \tag{12}$$

$$\alpha^2(\mu) = n^2\pi^2/3T^2 + 2n\pi\mu/3^{3/2}T + O(\mu^2).$$

Proof: Let us note that the nonlinear part of (11) and their first derivatives with respect to x and y vanish at $x = y = 0$. The linear part of (11) is

$$\ddot{x} - (\alpha^2/2)x = 0, \tag{13a}$$

$$\ddot{y} + 3\alpha^2 y = 0, \tag{13b}$$

and there are nonzero solutions

$$\Sigma = \begin{pmatrix} 0 \\ \sin(\omega_n t) \end{pmatrix} \tag{14}$$

with frequencies $\omega_n = \sqrt{3}\alpha_n$ being equal to $n\pi/T$. All eigenvalues ω_n are simple. For definiteness we assume $n = 1$.

Now we can apply standard results of bifurcation theory (for instance, Theorem 5.1, p. 188 in Ref. 4), which assert that the above described properties are necessary and sufficient to guarantee the existence of new solutions of Eqs. (11). These nontrivial solutions read

$$x(t,\alpha) = O(\mu^2),$$

$$y(t,\alpha) = \mu \sin((\pi/T)t) + O(\mu^2), \tag{15}$$

$$\alpha^2(\mu) = \pi^2/3T^2 + K(\mu) + O(\mu^2).$$

Morever, all small solutions to (11) are either the trivial one $x = y = 0$ or these given by (15). [A solution $(x(\mu), y(\mu))$ is said to be small if it tends to (0,0) as μ tends to zero.] Now we are in a position to compute the factor K of μ in the last of formulas (15) and to find all possible solutions of the form already described. Let us insert (15) into (11), multiply the resulting equations by the transpose Σ^T of Σ and integrate over the half-period T. Integrating by parts and discarding boundary terms, using the fact that Σ solves the linear equation (13) and performing some calculations yields

$$3\mu(K(\mu)T/\pi - 2\mu/3^{3/2}) + O(\mu^3) = 0. \tag{16}$$

Thus we obtain

$$K = 2\pi\mu/3^{3/2}T + O(\mu^2).$$

Dividing (16) by μ gives

$$F \equiv K(\mu)T/\pi - 2\mu/3^{3/2} + O(\mu^2) = 0, \tag{17}$$

whereas

$$\left.\frac{\partial F}{\partial(K(\mu))}\right|_{\substack{\mu=0 \\ K(\mu)=0}} = \frac{T}{\pi} \neq 0;$$

therefore using the implicit function theorem argument, one obtains the existence of exactly one small solution of (16). Since the number of small solutions of the reduced bifurcation equation (16) is equal to the number of small solutions of the original equations (11), we get the desired uniqueness result already stated in Theorem 1. To complete it, let us note that Eqs. (11) are solved by $x(t,\mu) = 0$ and $y(t,\mu)$ being a solution of (11b), which becomes now

$$\ddot{y} - \alpha^6/2(y + \alpha)^3 + \tfrac{1}{2}(y + \alpha)^3 = 0. \tag{18}$$

Remark 2: The bifurcation we found is usually called Hopf bifurcation in western literature. The plane $x = \dot{x} = 0$ in the phase space of the system, forms a central manifold for the trivial solution. What was proved above, reflects simply the fact that perturbation within the central manifold gives rise to periodic orbits. Note also that our bifurcation is nongeneric, in the sense of Ref. 4, and that it exists above and below the bifurcation point α_n.

We made some numerical investigations of Eq. (18). Figure 1 shows a logarithmic ($\log_{10} D$) plot of the deviation of the numerical solution from the approximate solution $\mu \sin(\omega_1 t)$, i.e.,

$$D = |y_n - (e3^{3/2}/2\pi)\sin(\omega_1 t)|.$$

In computer calculations we used the initial momentum

$$Py = \frac{d}{dt}(\mu \sin(\omega_1 t))|_{t=0} = 3^{3/2}e/2,$$

the half-period $T = 1$, $\omega_1 = \pi$, and a new parameter

$$e \equiv 2\pi\mu/3^{3/2} = 10^{-5}, 10^{-3}, 10^{-1}, 1.$$

Transparently there is a quite good agreement of both solutions up to $e = 10^{-3}$, where the deviation is of order 10^6 smaller than the amplitude of y_n. Figures 2–4 show numerical solutions for values $e = 10^{-5}, 10^{-1}, 1.0$, respectively. One can see the increasing-with-"e" deviation from the value $2T = 2$. This is due to the fact that the frequency $\omega(\alpha)$ of a solution depends on the amplitude "e"; our solution is a 2π-periodic function of $\omega(\alpha)t$, not of t alone.

III. EXACT SOLUTION

Theorem 1 states that bifurcation solutions describe oscillations along the y axis whose x position is fixed and whose y position is governed by Eq. (18). Note that the original equations (11) conserve an energy

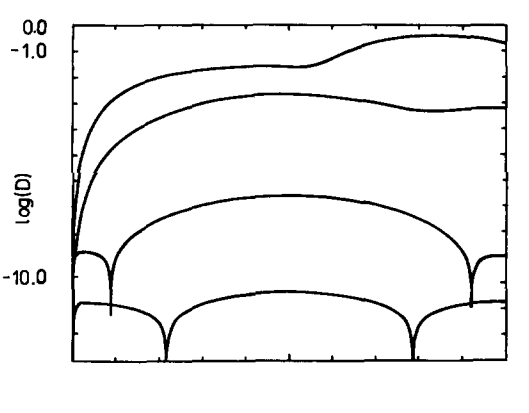


FIG. 1. The logarithmic plot of the deviation D of the numerical solution from the approximate sine solution for $e = 10^{-5}, 10^{-3}, 10^{-1}, 1.0$, respectively.

$$\epsilon = \frac{\dot{x}^2}{2} + \frac{\dot{y}^2}{2} + \frac{1}{8}(x^2 - (y+\alpha)^2)^2 + \frac{\alpha^6}{4(y+\alpha)^2} \tag{19}$$

and therefore Eq. (18) conserves

$$\epsilon = \dot{y}^2 + \tfrac{1}{8}(y+\alpha)^4 + \alpha^4/4(y+\alpha)^2.$$

Thus instead of (18), which is an equation of second order, one can write

$$\left(\frac{dz}{dt'}\right)^2 = 4z^3 - 32\epsilon z - 8\alpha^6 \tag{20}$$

[where $z = -(y+\alpha)^2$, $t' = t/2$], which is a first-order differential equation that can be obtained from the last expression for ϵ. Its solutions are the Weierstrass elliptic functions[5]

$$z(t) = P(t' + C, 32\epsilon, 8\alpha^6).$$

The constant C has to be determined from our periodicity condition and it should ensure that $z(t)$ is real negative and has no singularities for $t \in (-\infty, \infty)$. It follows from properties of the Weierstrass elliptic function that the imaginary part of the appropriate C must be equal to

$$\omega_2 = K(\sqrt{(e_1 - e_2)/(e_1 - e_3)})/\sqrt{e_1 - e_3}. \tag{21}$$

Here K is the complete elliptic integral and $e_1 > e_2 > e_3$ are the roots of the equation

$$4z^3 - 32\epsilon z - 8\alpha^6 = 0.$$

The roots are real in our case since the discriminant $\Delta = (32\epsilon)^3 - 27(8\alpha^6)^2$ is non-negative. The minimal value of ϵ is $\epsilon = \frac{3}{8}\alpha^4$ and then $\Delta = 0$. The real part of C is just the usual integration constant and is determined by the initial condition $z(0)$. One can check that for those α that are close to the already known critical value $\sqrt{3}\alpha = \pi/T$, and for energies ϵ, which differ slightly from the minimum $\frac{3}{8}\alpha^4$, this solution is close to our approximate sine solution. In particular the period

$$\omega_1 = 4K(\sqrt{(e_2 - e_3)/(e_1 - e_3)})/\sqrt{e_1 - e_3}$$

of the Weierstrass elliptic function $P(t' + C, 32\epsilon, 8\alpha^6)$ is close to $2T = 2$, which is the period of the approximate solution $\mu \sin(\omega_1 t) = \mu \sin(\pi t/T)$. Morever, for $\mu \to 0$ the Weierstrass function tends to $-\alpha^2$, which implies that $y = \sqrt{-P} - \alpha$ tends to 0; therefore $y = \sqrt{-P} - \alpha$ can be identified, using the uniqueness result of Theorem 1, with the already described bifurcation solution.

IV. BIFURCATION SOLUTIONS IN YANG–MILLS MECHANICS (CONTINUATION)

In this section we assume $L_1 = L_2 = 0$. In such a case Eqs. (3) reduce to those already studied in Ref. 1:

$$\ddot{s} + sw^2 = 0, \tag{22a}$$

$$\ddot{w} + ws^2 = 0. \tag{22b}$$

Here the notation is the same as in Sec. II. All components of potentials A^a_μ vanish, save A^1_1, A^2_2. These equations possess a static solution $s = 0$, $w = \alpha$. Define "v" by $w = v + \alpha$ and insert it into (22). Then we arrive at

$$\ddot{s} + \alpha^2 s + 2sv\alpha + sv^2 = 0, \tag{23a}$$

$$\ddot{v} + s^2\alpha + vs^2 = 0. \tag{23b}$$

Repeating the procedure already applied in the case $L_1 = 0$, $L_2 \neq 0$ and appealing to Ref. 4, one comes to the following result.

Theorem 2: The system (23) has a solution $(s(t,\mu), v(t,\mu))$, which bifurcates from $(s = 0, v = 0)$ at $\alpha_n = n\pi/T$, and

$$\begin{pmatrix} s \\ v \end{pmatrix} = \mu \begin{pmatrix} \sin \alpha_n t \\ 0 \end{pmatrix} + O(\mu^2),$$

where μ is related to α by $\alpha^2 = n^2\pi^2/T^2 + O(\mu)$. The coefficients of the leading terms of the expressions can be calculated analogously as in the proof of Theorem 1.

V. STABILITY

The solution of an evolution system is said to be stable if its perturbation $\delta x(t)$ is controllable, that is

$$\|\delta x(t)\| < C\,\|\delta x(t_0)\|, \tag{24}$$

where C is a constant independent of time as well as of $\delta x(t_0)$ and $\delta x(t_0)$ is an initial perturbation; the norms $\|\cdot\|$ are those of a suitable function space. Note that decreasing the initial perturbation $\|\delta x(t_0)\|$ yields $\|\delta x(t)\|$ tending to 0. It could be shown (e.g., Ref. 6) that if the evolution system is

$$\dot{x} = f(x), \tag{25}$$

with x, f belonging to a Banach space, f being nonlinear in x and satisfying certain smoothness conditions, then a solution x_0 of (25) is stable if the eigenvalues λ of

$$f'(x_0)\delta x = \lambda \delta x$$

have a negative real part.

Here $f'(x_0)$ denotes a linearization of f at x_0. And conversely, for the real part of λ being a positive number, the condition (24) is violated and the solution is unstable.

This criterion gives in our case the following result. The trivial solution (9) is unstable against perturbations in the plane $x - \dot{x}$, since Eq. (13a) has exponentially growing solutions (positive eigenvalues $\alpha^2/2$). This could be expected since Yang–Mills mechanics is known to be chaotic[1–3]; perturbing its initial state one cannot predict its further evolution. The notion of chaos is narrower than of instability, in the sense that chaos requires instability but not on the contrary. Thus it is of interest that the trivial solution appears to be stable under perturbations in the y-$\dot{y}$ plane, since all solutions of (13b) are bounded. This means that such perturbations do not imply chaotic behavior.

The analogous analysis for the Weierstrass elliptic function solution is much more complicated, since then the corresponding linearized equations contain double-periodic coefficients (Weierstrass elliptic functions). These are of the type known as Hill's equation.[7] We expect similar results as above, although the confirmation of this expectation would be troublesome.

VI. YANG–MILLS MECHANICS COUPLED TO A DIRAC FIELD

As we pointed out earlier, the solutions considered in Sec. II generate an Abelian source, which in explicit form reads

$$j^a_\mu = (\alpha^3/\sqrt{2})\delta^{a3}\delta_{\mu 0}.$$

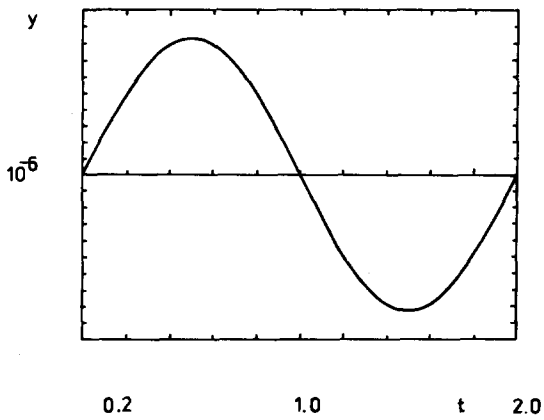


FIG. 2. Plot of the numerical solution for $e = 10^{-5}$.

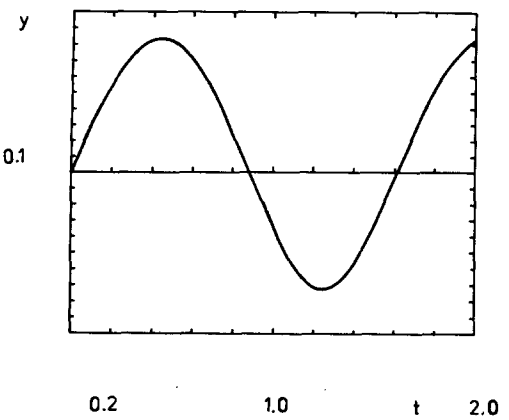


FIG. 4. Plot of the numerical solution for 1.0.

It is shown in Ref. 8 that this source is generated by a constant bispinor

$$\psi = \begin{pmatrix} g\sigma^2 + \chi\sigma^3\sigma^2 \\ g^*\sigma^2 + \chi^*\sigma^3\sigma^2 \end{pmatrix}, \tag{26}$$

where $\alpha^3 = -2\sqrt{2}(g^*\chi + g\chi^*)$ and $\arg g = \arg\chi \pm \pi$. The two previously found solutions, the trivial one and the Weierstrass function solution, and (26) form a self-consistent solution of Yang–Mills fields coupled to massless Dirac fields.

VII. SUMMARY

We proved the existence of new periodic solutions which bifurcate from a known static solution. Next, some of the solutions were found explicitly. In our original notation they read

$$A_0^a = 0, \quad A_i^3 = A_3^a = 0;$$

(i)

$$A_1^1 = -A_2^2 = \alpha' \cos(\alpha' t),$$

$$A_2^1 = A_1^2 = \alpha' \sin\alpha' t, \quad \alpha' = \alpha/\sqrt{2};$$

(ii)

$$A_1^1 = -A_2^2 = (1/\sqrt{2})(-P(t/2 + C, 32\epsilon, 8\alpha^6))^{1/2} \times \cos\theta(t),$$

$$A_2^1 = A_1^2 = (1/\sqrt{2})(-P(t/2 + C, 32\epsilon, 8\alpha^6))^{1/2} \times \sin\theta(t);$$

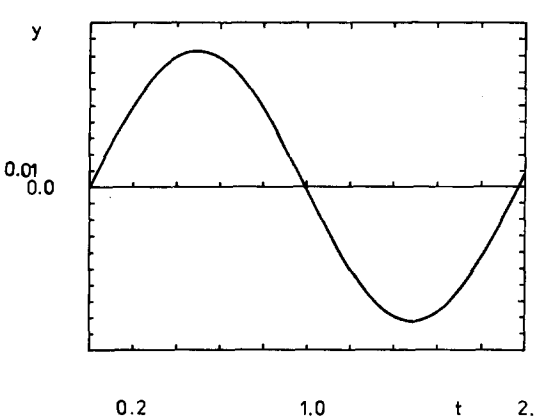


FIG. 3. Plot of the numerical solution for $e = 10^{-1}$.

where P is the Weierstrass elliptic function, C, ϵ, and α are defined as in Sec. III and

$$\theta(t) = \frac{-\alpha^3}{\sqrt{2}} \int_0^t \frac{dt'}{P(t'/2 + C, 32\epsilon, 8\alpha^6)}.$$

In the limit $\epsilon \to \frac{3}{8}\alpha^4$ the solution (ii) coincides with (i).

In the case $L_1 = 0, L_2 \neq 0$, these solutions describe oscillations along the y axis near $r_2 = \alpha_n = 2^{1/6} L_2^{1/3}$ for $\alpha_n = n\pi/\sqrt{3}T$, where T is (approximately) a half-period of the oscillations. The smallest value α at which bifurcation can occur is $\alpha_1 = \pi/\sqrt{3}T$, that is, coming back to our original notation, $(A_1^1)^2 + (A_1^2)^2 = \frac{1}{2}\alpha_1^2 = \pi^2/6T^2$. In the case $L_1 = L_2 = 0$, the new solutions describe rotations in the s-w plane around $s = 0$, $w = \alpha_n = n\pi/T$. The smallest value of potentials at which bifurcation can appear is $(A_1^1)^2 + (A_2^2)^2 = \alpha_1^2 = \pi^2/T^2$. Thus bifurcation solutions appear in the region of strong potentials. This is in accord with earlier results.[9]

The periodic-in-time solutions, which we found, could be physically relevant if the recent suggestion of Palumbo[10] that imposition on potentials of time-dependent boundary conditions will remedy QCD is true. The periodic boundary conditions cannot be imposed directly since they are inconsistent with hyperbolic equations (which Yang–Mills equations are). The natural way to implement them is to quantize QCD around solutions that are periodic in time. There are three candidates, that of Baseyan *et al.*[11] and those reproduced above [note that together with (26) they solve the Yang–Mills–Dirac system].

ACKNOWLEDGMENT

One of the authors (ME) is grateful to the ICTP, Trieste for financial support, which made possible his participation in the 1983 Summer School on Dynamical Systems as well as this joint work.

[1] S. G. Matinyan, G. K. Savvidi, and N. G. Ter-Arutyunyan-Savvidi, Zh. Eksp. Teor. Fiz. **80**, 830 (1981).
[2] I. Frøyland, Phys. Rev. D **27**, 943 (1983).
[3] A. Górski, Acta Phys. Pol. B **15**, 465 (1984).

[4]S. N. Chow and J. K. Hale, *Methods of Bifurcation Theory* (Springer, New York, 1982).
[5]E. T. Whittaker and C. N. Watson, *A Course of Modern Analysis* (Cambridge U. P., Cambridge, 1963) Vol. II.
[6]G. Ioos and D. D. Joseph, *Elementary Stability and Bifurcation Theory* (Springer, New York, 1980).
[7]E. Kamke, *Diffrentialgleichungen, Lösungsmethoden und Lösungen* (Leipzig, 1959).
[8]E. Malec, "Self-consistent solutions of Yang–Mills–Dirac equations," IFJ preprint No. 1251/P1, Krakow, 1984.
[9]R. Jackiw and P. Rossi, Phys. Rev. D **21**, 426 (1980); E. Malec, J. Phys. A **13**, 2609 (1980); Acta Phys. Pol. B **12**, 665 (1981); J. Math. Phys. **23**, 21 (1982).
[10]F. Palumbo, Phys. Lett. B **132**, 165 (1983).
[11]A. A. Baseyan, S. G. Matinyan, and G. K. Savvidi, Pis'ma Zh. Teor. Eksp. Fiz. **29**, 641 (1979).

Titchmarsh–Weyl theory and its relations to scattering theory: Spectral densities and cross sections; Theory and applications

Erkki Brändas[a)]
Quantum Theory Project, Williamson Hall, University of Florida, Gainesville, Florida 32611
Magnus Rittby and Nils Elander
Research Institute of Physics, Frescativägen 24, S-104 05 Stockholm, Sweden

(Received 5 December 1984; accepted for publication 12 April 1985)

The radial equation (or set of equations) derived in scattering theory is analyzed by means of Titchmarsh–Weyl theory for singular second-order differential equations. In particular we have focused on the spectral density concept and the corresponding relation to the scattering cross section. The method of complex deformations is brought in as a necessary ingredient in the evaluation of the underlying pole strings, which together with the background build up the actual dispersion relation data. The analysis is supported by numerical applications to a centrifugal family of simple potentials.

I. INTRODUCTION

In a scattering experiment the goal of the experimentalist is to determine the cross sections for the different processes that may occur. The information gained from such experiments, complemented by information from different types of spectroscopy, gives us a picture of the physics of the colliding system.

An important feature in the scattering experiments[1,2] is the occurrence of resonances in the observed cross sections, which are related to quasibound states formed by the colliding particles. Such resonances can be related to complex poles of the scattering matrix describing the collision. One should note in this context that the definition of a resonance is not unique. It can either be introduced in terms of its physical effects[3] or as an eigenvalue of an analytically continued Hamiltonian.[4,5] Even though for resonances with small widths the two points of view may be intimately related, they are different. We will here exploit the last definition. From a theoretical point of view the ability to look at the extension of scattering quantities to complex energy values is therefore of great interest. This extension provides the experimentalist with a possibility of interpreting the cross sections, labeling the resonance structure in terms of "generalized" quantum numbers, and unites the bound state and scattering theory in an attractive framework.

In this work we will return to the connection between Titchmarsh–Weyl (TW) theory for singular second-order differential equations[6,7] and the scattering formulation.[8,9] We will apply TW theory to the radial equation, or set of equations, derived in scattering theory and show that uniquely defined spectral densities can be obtained. In particular, we will concentrate on the spectral density as given by the boundary value problem and its associated expansion theorems. Furthermore, we will investigate the asymptotic behavior of the spectral density for large energy values. Stability questions in connection with the numerical determination of resonance information on the second sheet are analyzed in the light of complex deformations. The formulation is supplemented with some examples.

The paper is outlined as follows. In Sec. II we will emphasize the formulation embodying the case of a differential operator defined on an interval containing two singular endpoints. In our applications we have usually applied a one singular endpoint theory, i.e., on the interval $[a;\infty)$. The trick has then consisted of choosing a boundary condition at the point a, corresponding to a solution regular at origin. This would seem to introduce an energy-dependent boundary condition (b.c.), which is not in accordance with a proper definition of a self-adjoint differential operator. It is shown, however, that our "boundary-dependent" spectral density can be identified with the trace of the Coddington–Levinson spectral matrix[10] for a two singular endpoint interval, and hence our problem on $[a,\infty)$ is part of a more general formulation and can be imbedded in the theory for the two singular endpoint problem with energy-independent b.c.'s. Rigorous expansion theorems based on the Coddington–Levinson spectral density and in terms of general completeness relations are then obtained. Section III gives some examples of the spectral density function and comments on their displayed properties.

In Sec. IV we will consider the dependence of the spectral density for large values of $k=\sqrt{E}$. Many results on asymptotic forms of m functions exist, see, e.g., Akhiezer.[11,12] Rather than going into this general development, we will here present a simplified analysis derived from our physical assumptions and knowledge of the properties of the potential. Section V contains some additional numerical applications to scattering cross sections.

In Sec. VI we will present the idea of incorporating complex deformation techniques into the theory. This is necessary in order to understand the details of the spectral density or the scattering cross section. In particular, we will aim at the detection of complex poles and their physical occurrence in dispersion relation data. Numerical difficulties associated with complex pole evaluations are analyzed in Sec. VII. Some typical examples showing various types of pole

[a)] Permanent address: Quantum Chemistry Group for Research in Atomic, Molecular and Solid State Physics, University of Uppsala, S-751 20 Uppsala, Sweden.

 0022-2488/85/102648-11$02.50

strings are supplemented.

Section VIII, finally, contains a summary of the spectral density concept and the information obtained from the physical information on the real axis.

II. WEYL'S THEORY FOR A SECOND-ORDER DIFFERENTIAL OPERATOR

A. The two singular endpoint case

Weyl's theory[6,7,10] has been shown to be a useful tool in the investigation of the bound as well as the continuous spectrum. Applications not employing the complex deformation techniques described in Sec. VI have been made on the Stark effect on hydrogen,[13] predissociation by rotation.[9,14,15]

Weyl's theory, in the context of ordinary differential equations, refers to Weyl's extension of the Sturm–Liouville theory for a finite interval $[a,b]$ to an infinite interval $(0,+\infty)$ or $(-\infty,+\infty)$. The theory is also applicable to integration intervals containing points where the potential may be singular. For a detailed description the reader is referred to Everitt and Bennewitz.[12]

Since the extension problem of formally self-adjoint operators is vital to any quantum theoretical formulation, albeit seldom mentioned, we will here give a rather expository account of the two singular endpoint case. We will later see that this point of view is even more important in the context of spectral deformations. Note that modern treatments of functional analytic approaches[16–18] interpret mathematical physics to explain these things.

Consider the differential operator L, defined by

$$L(u(t)) = -(p(t)u(t)')' + q(t)u(t), \tag{2.1}$$

where $p(t)$ and $q(t)$ are assumed to be real and continuous and $p(t)>0$ on $[a,+\infty)$.

The Titchmarsh–Weyl m coefficient is related to the spectral density of L and the following theorem holds.

***m* theorem:** Let the interval $I=[a,+\infty)$ and the coefficients p and q be given as above and two initial value solutions be defined by $(0\leqslant\alpha<\pi/2)$

$$\varphi(a,\lambda)=\sin\alpha, \quad \psi(a,\lambda)=\cos\alpha,$$
$$p(a)\varphi'(a,\lambda)=-\cos\alpha, \quad p(a)\psi'(a,\lambda)=\sin\alpha.$$

Then there exists at least one pair of analytic functions (m_+,m_-), which have the following properties: (i) $m_\pm$ maps the complex number fields C onto itself, i.e. $m_\pm : C_\pm \to C_\pm$; (ii) $m_\pm$ belongs to the class of Cauchy analytic functions holomorphic in $C_\pm$, i.e., $m_\pm \in H(C_\pm)$; (iii) $\chi_\pm(\lambda)=\varphi(\lambda)+m_\pm(\lambda)\psi(\lambda)\in L^2(a,\infty)$ $(\lambda\in C_\pm)$; (iv) $(m_\pm(\lambda))^*=m_\mp(\lambda^*)$ $(\lambda\in C_\pm)$; (v) either (a) there is a limit circle (LC) at ∞, $\varphi(\lambda)\in L^2(a,\infty)$ and $\psi(\lambda)\in L^2(a,\infty)$ $(\lambda\in C)$, in which case m is not unique or (b) there is a limit point (LP) at ∞, $\varphi(\lambda)\notin L^2(a,\infty)$ and $\psi(\lambda)\notin L^2(a,\infty)$ $(\lambda\in C_\pm)$, in which case $m_\pm$ is unique.

In the m theorem above a is a regular point of the interval. If both endpoints, however, are singular, the associated m function has to be determined via appropriate matchings of logarithmic derivatives of the left- and right-hand regular solutions at some nonsingular a within the interval. This situation is incorporated in our Titchmarsh–Weyl form[8] for m via the phase angle α determined so that $\tan\alpha$ becomes the logarithmic derivative of the solution regular at the left boundary. A careful study of Weyl's theory for the singular case in both endpoints can be found in Coddington and Levinson.[10] The development follows closely the formulation deduced for one singular boundary.

At the matching point a we define two solutions φ_1 and φ_2 with the dependence on the complex parameter $\lambda=E+i\epsilon$, $\epsilon>0$,

$$\begin{aligned}&\varphi_1(a)=1, \quad \varphi_2(a)=0,\\ &p(a)\varphi_1'(a)=0, \quad p(a)\varphi_2'(a)=1.\end{aligned} \tag{2.2}$$

By assigning the real boundary conditions

$$\chi_{IB_I}(B_I)\cos\beta_I + p(B_I)\chi'_{IB_I}(B_I)\sin\beta_I = 0, \tag{2.3}$$

where $I=1$ (2) corresponds to the left (right) part of the interval $(-\infty,+\infty)$ partitioned by the point a, we will study the limiting procedures $B_1\to-\infty$ and $B_2\to+\infty$.

Since the problem $[a,+\infty)$ and $(-\infty,a]$ yields, point by point,

$$\begin{aligned}\chi_{-\infty} &= \varphi_1+\varphi_2 m_{-\infty} = \lim_{B_1\to-\infty}\chi_{1B_1}\\ &= \varphi_1+\varphi_2\Big(\lim_{B_1\to-\infty} m_{1B_1}\Big),\end{aligned} \tag{2.4a}$$

$$\begin{aligned}\chi_{+\infty} &= \varphi_1+\varphi_2 m_{+\infty} = \lim_{B_2\to+\infty}\chi_{2B_2}\\ &= \varphi_1+\varphi_2\Big(\lim_{B_2\to+\infty} m_{2B_2}\Big),\end{aligned} \tag{2.4b}$$

it is a simple matter to match $\chi_{-\infty}$ and $\chi_{+\infty}$ in order to obtain eigensolutions of (2.1) that are regular at both endpoints. The spectral matrix associated with (2.1) and (2.3) is directly obtained from the continued use of the general form of Parseval's relation as

$$\begin{aligned}&\int_\delta f_1^*(t)f_2(t)dt\\ &\quad=\sum_{n=1}^{\infty}\Big(\int_\delta f_1^*(t)h_{\delta n}(t)dt\Big)\Big(\int_\delta h_{\delta n}^*(t)f_2(t)dt\Big),\end{aligned} \tag{2.5}$$

where f_i, $i=1,2$, are square integrable functions, δ is the interval (B_1,B_2), and $h_{\delta n}$ are Sturm–Liouville normalized eigenfunctions of (2.1) [with b.c. (2.3), $I=1,2$] on the interval δ with eigenvalue $E_{\delta n}$. One also gets

$$h_{\delta n}=r_{\delta n1}\varphi_1(E_{\delta n})+r_{\delta n2}\varphi_2(E_{\delta n}), \tag{2.6}$$

where $r_{\delta n}$ are complex constants. For $f_1=f_2=f$, we obtain

$$\int_\delta |f(t)|^2\,dt=\int_{-\infty}^{+\infty}\sum_{j,k=1}^{2} g_{\delta j}(\omega)g_{\delta k}^*(\omega)d\rho_{\delta jk}(\omega), \tag{2.7a}$$

with

$$g_{\delta j}(\omega)=\int_\delta \varphi_j^*(\omega,t)f(t)dt \tag{2.7b}$$

and

$$\rho_{\delta jk}(E_{\delta n}+0)-\rho_{\delta jk}(E_{\delta n}-0)=\sum_m r_{\delta mj}r_{\delta mk}, \tag{2.8}$$

where the sum is taken over all m such that $E_{\delta m}=E_{\delta n}$. In the limits $B_1\to-\infty$ and $B_2\to+\infty$ one finds

$$\int_{-\infty}^{+\infty}|f(t)|^2\,dt=\int_{-\infty}^{+\infty}\sum_{j,k=1}^{2} g_j(\omega)g_k^*(\omega)d\rho_{jk}(\omega), \tag{2.9a}$$

with

$$g_j(\omega) = \int_{-\infty}^{+\infty} \varphi_j^*(\omega,t) f(t) dt. \tag{2.9b}$$

A closer analysis based on the appropriate Green's functions of the left and right boundary value problem shows that the matrix ρ can be obtained from the Coddington–Levinson M matrix, via

$$M_{11}(\lambda) = (m_{-\infty}(\lambda) - m_{+\infty}(\lambda))^{-1}, \tag{2.10a}$$

$$M_{22}(\lambda) = m_{-\infty}(\lambda) m_{+\infty}(\lambda)(m_{-\infty}(\lambda) - m_{+\infty}(\lambda))^{-1}, \tag{2.10b}$$

$$M_{12}(\lambda) = M_{21}(\lambda) = \tfrac{1}{2}(m_{-\infty}(\lambda) + m_{+\infty}(\lambda)) \times (m_{-\infty}(\lambda) - m_{+\infty}(\lambda))^{-1}, \tag{2.10c}$$

and the Stieltjes inversion formula ($\omega_2 > \omega_1$)

$$\rho_{jk}(\omega_2) - \rho_{jk}(\omega_1) = \lim_{\epsilon \to 0^+} \frac{1}{\pi} \int_{\omega_1}^{\omega_2} \mathrm{Im}(M_{jk}(E + i\epsilon)) dE. \tag{2.11}$$

The formulation above applies as well to any interval with a singularity in both the left and the right boundary point, whether it is finite or not.

Defining the Wronskian as

$$W(u,v) = (uv' - u'v)p, \tag{2.12}$$

we note that the "Titchmarsh–Weyl" m expression on the interval $(0,\infty)$, with α energy dependent, can be written as

$$m_{(\mathrm{TW})}(\lambda) = \frac{W(\varphi,\chi_{+\infty})}{W(\chi_{+\infty},\psi)} = \frac{Z\sin\alpha + \cos\alpha}{\sin\alpha - Z\cos\alpha} = \frac{Z\tan\alpha + 1}{\tan\alpha - Z}, \tag{2.13}$$

where $Z(a) = p(a)\chi'_{+\infty}(a)/\chi_{+\infty}(a)$ and $\tan\alpha = p(a)\psi'(a)/\psi(a)$ (ψ regular at the origin). From Eq. (2.13), it follows that $m_{(\mathrm{TW})}(\lambda)$ is related to the Coddington–Levinson M matrix with the left singular boundary being the point at zero, via

$$\mathrm{Tr}(M) = m_{(\mathrm{TW})}. \tag{2.14}$$

Relation (2.14) follows simply from the fact that m with the initial conditions (2.2) is directly given by the logarithmic derivatives of respective regular solution at the endpoints. For details, see Refs. 8 and 19.

It is important to point out, however, that the simple expression derived by Titchmarsh should, strictly speaking, be considered only for energy-independent boundary conditions, i.e., α independent of λ. In this context one may, for $\alpha \neq \pi/2$, deduce the Hille–Wray[20,21] function theoretical form

$$m(\lambda) = -\tan\alpha + \int_{-\infty}^{+\infty} \frac{d\rho(\omega,\alpha)}{\omega - \lambda}, \quad \alpha \neq \pi/2, \tag{2.15}$$

or in its most general Nevanlinna form[11] (including also $\alpha = \pi/2$)

$$m(\lambda) = \beta + \gamma\lambda + \int_{-\infty}^{+\infty} \left(\frac{1}{(E-\lambda)} - \frac{E}{(E^2+1)}\right) d\rho(E), \tag{2.16}$$

with β and γ real and λ independent with $\gamma \geqslant 0$. Since m given by (2.13) necessarily implies $\alpha = \alpha(\lambda)$, the forms (2.15) and (2.16) are not strictly valid, although they can be simply generalized to the λ-dependent case. In the latter case one must interpret the Titchmarsh–Weyl expression (for a being an interior point of the interval I) in terms of the expansion theorem (2.9)–(2.13) and $m(\lambda) = m_{(\mathrm{TW})}$ as the appropriate trace associated with the singular endpoint theory.

B. Connections with scattering theory, completeness relations

In the applications treated below, we will put $p = 1$, and let $q(r) = l(l+1)/r^2 + V(r)$, where $V(r)$ is less singular than r^{-2} at the origin and decreases faster than the centrifugal term at infinity. Since the origin for $l \neq 0$ is also a singular point, we will develop the explicit m expression for each partial wave corresponding to the actual rotational quantum number. For simplicity we begin with the case $l = 0$. We can here choose $a = 0$ and $\alpha = \pi/2$. The regular solution $\psi(r)$ is defined by the conditions $\psi(0) = 0$ and $\psi'(0) = 1$. Although $l = 0$ corresponds to a limit circle situation, with all solutions square integrable at origin, the condition $\psi(0) = 0$ is necessary to assure self-adjointness of the three-dimensional Laplacian. This means that, with $k = (E - q(\infty))^{1/2}$ and E real, we have

$$\lim_{r\to 0}(1/r)\psi(r) = 1, \tag{2.17a}$$

$$\lim_{r\to\infty}(\psi(r) - (|f(k)|/k)\sin(kr + \delta(k))) = 0, \tag{2.17b}$$

where $f(k,r)$ is the exponential solution (Jost solution), i.e.,

$$\lim_{r\to\infty} \exp(\mp ikr) f^{\pm}(k,r) = 1 \tag{2.18}$$

and

$$f(k) = f(k,0) = W(f,\psi) = |f(k)|\exp(-i\delta) \tag{2.19}$$

is the Jost function. Using the relation between m and $f(k)$, one obtains [$E \in \sigma_c(L)$] the Kodaira form for the spectral density,[9,22]

$$\pi\left(\frac{d\rho}{d\omega}\right) = \mathrm{Im}(m(E)) = \frac{iW(f^+,f^-)}{2|W(f^+,\psi)|^2} = \frac{k}{|f(k)|^2}. \tag{2.20}$$

Although (2.13) is valid for all l (and even for more general potentials like the Coulomb or Stark potentials), the limit $a \to 0$ cannot be taken directly. To treat the case $l > 0$ we instead proceed by defining the regular solution by

$$\lim_{r\to 0}(2l+1)!!\, r^{-l-1}\psi_l(r) = 1 \tag{2.21}$$

and

$$\lim_{r\to\infty}\left(\psi_l(r) - \frac{|f_l(k)|}{k^{l+1}}\sin\left(kr - \frac{l\pi}{2} + \delta_l\right)\right) = 0. \tag{2.22}$$

In (2.22) the partial wave Jost function $f_l(k)$ is given by

$$f_l(k) = (-k)^l W(f_l(k,r),\psi_l) = |f_l(k)|\exp(-i\delta_l), \tag{2.23a}$$

$$\lim_{r\to\infty}(i^l \exp(ikr) - f_l(k,r)) = 0. \tag{2.23b}$$

Note that the two branches $f_l^{\pm}(k,r)$ are defined as

$$\lim_{r\to\infty}((\pm i)^l \exp(\pm ikr) - f_l^{\pm}(k,r)) = 0, \tag{2.24}$$

with (k real)

$$(f_l^+(k,r))^* = f_l^-(k,r) = (-1)^l f_l(-k,r), \tag{2.25a}$$

$$(f_l^+(k))^* = f_l^-(k) = f_l(-k). \tag{2.25b}$$

Generalizing (2.20) to $l>0$, one obtains

$$\pi\left(\frac{d\rho_l}{d\omega}\right) = \frac{iW(f_l^+, f_l^-)}{2|W(f_l^+, \psi_l)|^2} = \frac{k^{2l+1}}{|f_l(k)|^2}. \tag{2.26}$$

As mentioned for $l=0$ and $E>0$, it is possible to take the limit $a\to 0$ in (2.13). At a bound state, $E<0$, the spectral function has the jump

$$|a_j^l|^2 = \left(\int_0^\infty |\psi(E_j^l, r)|^2\, dr\right)^{-2}, \tag{2.27}$$

valid for all l.

We note that for $l=0$, $q\equiv 0$ ($a=0$ and $\alpha=\pi/2$), $m=i\sqrt{\lambda}$ is its Nevanlinna form. However, for $l\geqslant 0$ the growth of ρ with E is too rapid to allow for a direct representation of Nevanlinna character (see ρ theorem in Ref. 12). We then construct the completeness relation as

$$\delta(r-r') = \int_{-\infty}^{+\infty} \psi_l(\omega,r)\psi_l(\omega,r')d\rho_l(\omega). \tag{2.28}$$

The spectral density thus has the form

$$\left(\frac{d\rho_l}{d\omega}\right)_{\omega=E} = \begin{cases} \dfrac{1}{\pi}\dfrac{E^{l+1/2}}{|f_l(E^{1/2})|^2}, & E\geqslant 0, \\ \displaystyle\sum_j |a_j^l|^2 \delta(E-E_j^l), & E<0. \end{cases} \tag{2.29}$$

The spectral density derived here can be found, for instance, in Newton's book.[23] It is interesting to note, however, that (2.13) for $a\neq 0$ is of general validity, covering all values of l, and, as we will see, it will be of practical value in searching for the poles.

Summarizing this section, we have derived two types of spectral densities. The first one, given by formula (2.13), associated with expansion theorems (2.9)–(2.11), is the Titchmarsh–Weyl spectral density. The other one, defined by (2.26), associated with the completeness relations (2.28) and (2.29), will be called the Kodaira form. For $l=0$ one can make the two types coalesce by taking the limit $a\to 0$, whereas for $l>0$ this is not directly possible.

III. SPECTRAL DENSITIES

The spectral densities defined in the previous section, i.e., for $E\in\sigma_c(L)$,

$$\pi\left(\frac{d\rho}{d\omega}\right)_{\omega=E} = \mathrm{Im}(m_{(\mathrm{TW})}(E)), \tag{3.1}$$

$$\pi\left(\frac{d\rho}{d\omega}\right)_{\omega=E} = \frac{E^{l+1/2}}{|f_l(E^{1/2})|^2}, \tag{3.2}$$

are straightforward to evaluate numerically. In particular (3.1) is simple to obtain, since $m_{(\mathrm{TW})}$, see Eq. (2.13), only depends on logarithmic derivatives and can be determined from the Riccati equation.[19] The Kodaira density (3.2) on the other hand depends on the Jost function, see (2.26), and hence the Wronskian between the regular solutions at both endpoints has to be calculated. Unlike (3.1), the Kodaira form ($l\neq 0$) is not derivable from a Nevanlinna function and hence the information from the Riccati equation is not sufficient to obtain (3.2).

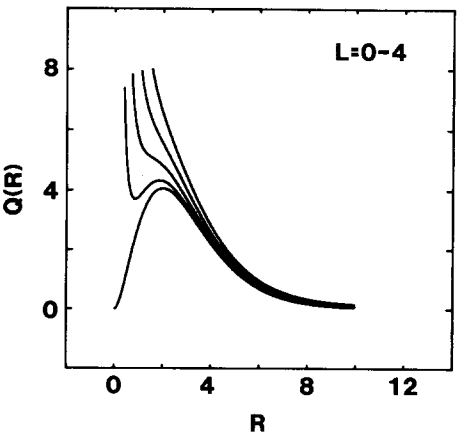


FIG. 1. The potential $V(r) = V_0 r^2 \exp(-r)$ was chosen in our numerical illustrations. The five different curves show, from bottom to top, the potential including the centrifugal term for $l=0, 1, 2, 3, 4$.

In order to compare the different spectral densities derived above, we made a numerical investigation of their properties using the model potential of Bain *et al.*[24]

$$V(r) = V_0 r^2 \exp(-r) \tag{3.3}$$

and a Hamiltonian which in our units is (here $p=\frac{1}{2}$)

$$H = -\frac{1}{2}\frac{d^2}{dr^2} + V(r) + \frac{l(l+1)}{2r^2}. \tag{3.4}$$

The effective potential, $q(r) = V(r) + l(l+1)/2r^2$, is shown in Fig. 1.

To obtain the Titchmarsh–Weyl form for the spectral density, the required logarithmic derivatives [see Eq. (2.13)] were calculated by numerical integration of the Riccati equation using the method by Johnson.[25] To evaluate the Wronskian in the Kodaira expression the regular solutions were integrated using the Numerov technique. The resulting spectral densities are displayed in Figs. 2–4 for $l=0$–2 (Kodaira) and Figs. 5–7 for $l=0$–2 (Titchmarsh–Weyl) with varying matching point a.

Note the difference between the Titchmarsh–Weyl den-

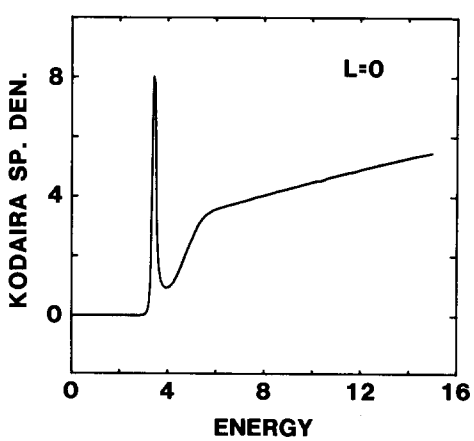


FIG. 2. The Kodaira spectral density for the illustration potential with $l=0$. Note the sharp peak indicating the resonant state.

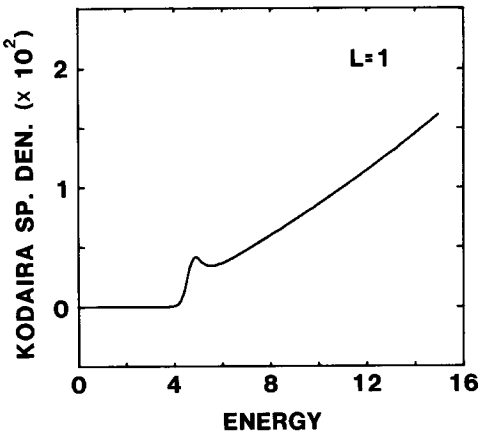


FIG. 3. Same as in Fig. 2 but for $l = 1$. Note that the peak here is less pronounced.

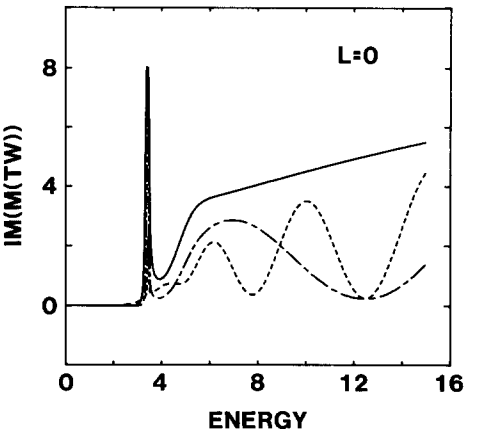


FIG. 5. Spectral densities for $l = 0$. The full line (—) shows the Kodaira form. The semidashed line (- - -) and the dashed line (- - - -) show the Titchmarsh–Weyl densities for $a = 1.0$ and $a = 2.5$, respectively.

sity and the Kodaira density. For $l = 0$, (2.26) or (2.20) corresponds to (2.13) if the matching point a is zero. In Fig. 5 the difference is displayed explicitly. The spectral density given by formula (2.20) approaches k for large energies and $l = 0$, and k^{2l+1} for $l > 0$. As will be shown in the next section, an oscillatory behavior of $m_{(\mathrm{TW})}$ for $a > 0$ appears in the region of large energy. This is clearly brought out in Figs. 5–7, where different matching points a have been used. Note that if a increases so do the oscillations. Only for $a = 0$ or by using (2.26) does one get the free-particle density k^{2l+1}. Since peaks in m, corresponding to resonances, have different magnitudes for different matching points, earlier numerical treatments were somewhat dependent on a. In the complex scaling approach (see below), this is no longer the case.

For an isolated resonance close to the real axis, one finds that the real part of m rapidly changes sign around the maximum of the imaginary part. This feature is indeed very important in locating sharp resonances. It is illustrated for $l = 0$ and $a = 0$ in Fig. 8. However, as resonances move down in the complex plane, the sign change in the real part will not occur on the maximum of the imaginary part. The latter will show non-Lorentzian behavior, and the associated width is much broader than the one given by the actual pole.

When poles approach and go beyond the complex threshold (26), the corresponding real parts of the pole may be of the same magnitude. Due to such pole interactions, one will find that the real part of m may no longer change sign. The onset of this trend is illustrated in Figs. 9 and 10. In such cases we have found that ansatz-dependent analytical procedures, as well as phase shift methods, become very unreliable. In obtaining the appropriate background the properties of the potential become crucially important. To deal with this situation we will show that the combination of the Titchmarsh–Weyl theory with complex scaling techniques supplies us with a rigorous formulation with attractive consequences for numerical applications.

IV. ASYMPTOTIC ANALYSIS OF THE TITCHMARSH–WEYL THEORY AND ITS RELATION WITH THE SCATTERING CROSS SECTION

The spectral densities derived in the previous section are intimately connected with the scattering matrix and thus with the total and differential cross sections. Previously this has been demonstrated in terms of Wronskians involving regular, irregular, and Jost solutions. As commented upon in Sec. II, the direct evaluation of resonance poles via the Titchmarsh–Weyl m function (2.13) becomes particularly attractive when angular momentum quantum numbers become

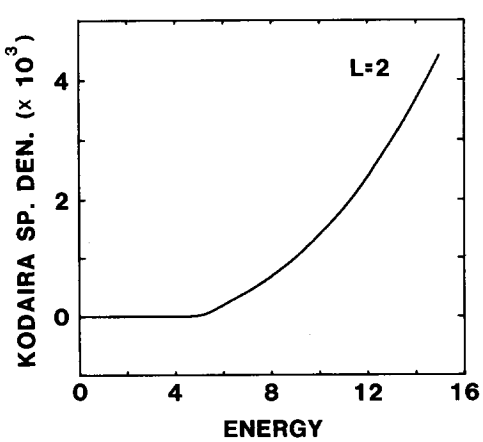


FIG. 4. Same as in Fig. 2 but $l = 2$. Here only the most optimistic reader will find a hint of a peak structure around 7 to 8 energy units.

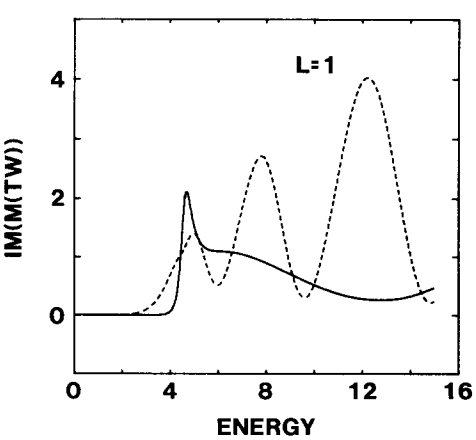


FIG. 6. $\mathrm{Im}(m_{\mathrm{TW}}(E))$ for $l = 1$, $a = 0.5$ (full line) and $a = 2.5$ (dashed line).

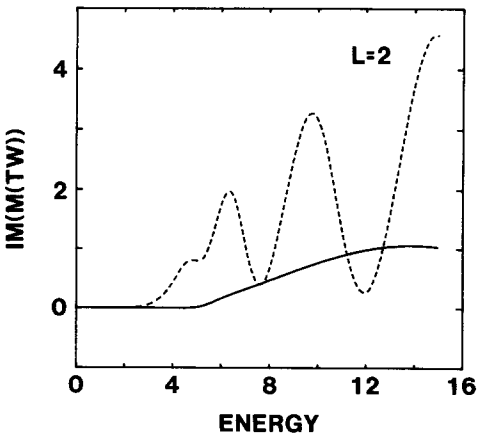


FIG. 7. Same as Fig. 6 but for $l = 2$, $a = 0.5$ (full line) and $a = 2.5$ (dashed line). The maximum that usually is associated with the resonance is here only an asymmetry in a shoulder of an oscillation.

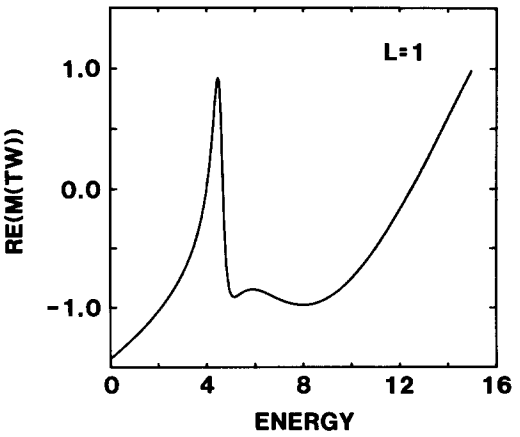


FIG. 9. Same as Fig. 8 but for $l = 1$ and $a = 0.5$. The behavior for resonances close to the real axis, i.e., the sharp down–up variation, has almost disappeared here.

large as well as in more general cases like the Coulomb interaction or the Stark effect. As we will be interested in the asymptotic behavior of the m function (2.13), the trace of the Coddington–Levinson spectral matrix, a simple evaluation of the appropriate Wronskian forms (see previous section) at an asymptotic region (kr large) will be carried out and tested.

In principle, this asymptotic formulation will yield complete spectral information, and hence scattering data, without explicitly having to determine Riccati–Bessel and Riccati–Neumann functions for all l values of interest. Although the latter conclusion is of practical significance in the actual resonance pole determination, the scattering information will be much more difficult to obtain, due to slow convergences. Nevertheless, our asymptotic formulation explains the oscillatory behavior of the spectral density previously noted as well as is entirely expressed in resonance parameters.

To proceed we notice that the asymptotic form of the regular solution is

$$\psi_l(kr) \underset{kr\to\infty}{\to} \exp(i\delta_l(k))\sin(kr - l\pi/2 + \delta_l(k)). \tag{4.1}$$

The Weyl solution is proportional to the Jost solution, $E \in \sigma_c(L)$,

$$\chi_l(kr) \propto f_l(kr) \underset{kr\to\infty}{\to} \exp(ikr). \tag{4.2}$$

We may now evaluate $m_{(\mathrm{TW})}$ as $kr\to\infty$ by using the asymptotic expressions above in (2.13):

$$\lim_{a\to\infty}\left(m^l_{(\mathrm{TW})}(k,a) - \frac{ik^2\cot(ka - l\pi/2 + \delta_l(k)) + 1}{k\cot(ka - l\pi/2 + \delta_l(k)) - ik}\right) = 0$$

or (4.3)

$$\lim_{a\to\infty}(m^l_{(\mathrm{TW})}(k,a) - (1/2ik)((1-k^2)\exp(2i(ka - l\pi/2 + \delta_l(k)) - (1+k^2)))) = 0$$

and

$$\lim_{a\to\infty}(\mathrm{Im}\, m^l_{(\mathrm{TW})}(k,a) - (1/k)((k^2-1)\cos^2(ka - l\pi/2 + \delta_l(k)) + 1)) = 0. \tag{4.4}$$

For a free particle, where $V(r)\equiv 0$ for all r, we have that the solution regular at the origin is the Riccati–Bessel function

$$\hat{j}_l(kr) = krj_l(kr), \tag{4.5}$$

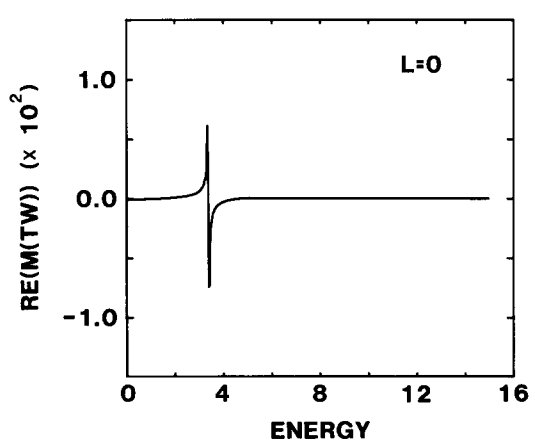


FIG. 8. For isolated resonances close to the real axis $\mathrm{Re}(m_{\mathrm{TW}}(E))$ rapidly changes sign from plus to minus around the resonance position. This is here illustrated for $l = 0$ and $a = 0.5$.

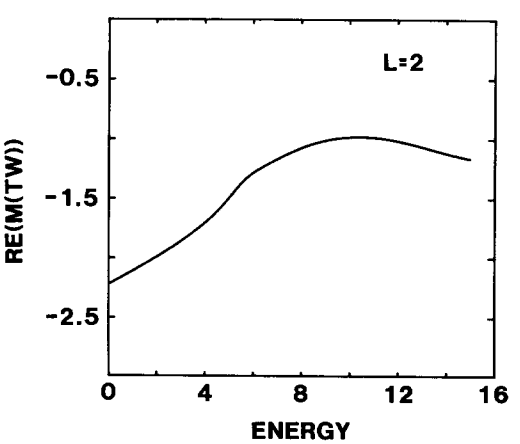


FIG. 10. Same as in Fig. 8 but for $l = 2$ and $a = 0.5$. Here the sharp variation close to the resonant level has disappeared completely.

where $j_l(kr)$ is the spherical Bessel function. For large ka the asymptotic form of the Riccati–Bessel function,

$$\hat{j}_l(kr) \underset{kr\to\infty}{\longrightarrow} \sin(kr - l\pi/2), \tag{4.6}$$

together with the asymptotic form of the Jost solution now gives

$$\lim_{a\to\infty}\left(m^l_{(\mathrm{TW})0}(k,a) - \frac{ik^2\cot(ka - l\pi/2) + 1}{k\cot(ka - l\pi/2) - ik}\right) = 0$$

or

$$\lim_{a\to\infty}(m^l_{(\mathrm{TW})0}(k,a) - (1/2ik)((1-k^2)\exp(2i(ka - l\pi/2) - (1+k^2)))) = 0 \tag{4.7}$$

and

$$\lim_{a\to\infty}(\mathrm{Im}\, m^l_{(\mathrm{TW})0}(k,a) - (1/k)((k^2-1)\cos^2(ka - l\pi/2) + 1)) = 0. \tag{4.8}$$

We may now express the scattering amplitude

$$F_l(k) = (1/k)\exp(i\delta_l(k))\sin\delta_l(k) \tag{4.9}$$

by

$$F_l(k) = \lim_{a\to\infty}\left(\frac{(m^l_{(\mathrm{TW})}(k,a) - m^l_{(\mathrm{TW})0}(k,a))}{(1-k^2)} \times\exp\left(-2i\left(ka - \frac{l\pi}{2}\right)\right)\right). \tag{4.10}$$

The total elastic cross section may be written as

$$\sigma = \sum_l \sigma_l, \tag{4.11}$$

in which the contribution to the total cross section by each partial wave is

$$\sigma_l(k) = 4\pi(2l+1)|F_l(k)|^2. \tag{4.12}$$

Using (4.10) we get

$$\sigma_l(k) = 4\pi(2l+1)\lim_{a\to\infty}\left(\frac{|m^l_{(\mathrm{TW})}(k,a) - m^l_{(\mathrm{TW})0}(k,a)|^2}{|1-k^2|^2}\right). \tag{4.13}$$

From the asymptotic formulas above, it is obvious that the Titchmarsh–Weyl m function will display oscillatory behavior for large k. This is clear from (4.7) and (4.8) with $q(r)\equiv 0$, and from (4.3) and (4.4). In the latter case the phase shift appears explicitly. Consequently the oscillatory behavior of the Titchmarsh–Weyl spectral density as compared to the free particle spectral density will yield information on the partial wave cross section. We will come back to this problem in the following sections.

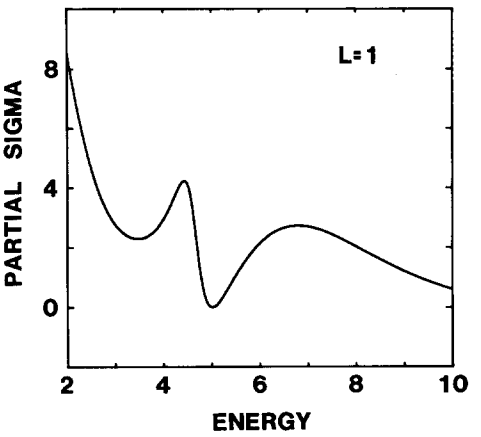


FIG. 12. Same as in Fig. 11 but for $l = 1$.

V. CROSS-SECTION CALCULATIONS

To give an idea of the usefulness of the asymptotic formulas of the previous section, the partial cross section for $l = 0$ was calculated (see Fig. 11). The Titchmarsh–Weyl form for the m function was again obtained by the numerical integration of the appropriate logarithmic derivatives. For $l = 0$ the convergence was satisfactory but for higher l values the convergence properties got worse, as expected, and the evaluation had to be made at very large r values. For comparison, the partial cross section was obtained from the K matrix using the method of Johnson[25] via

$$S = (1 + iK)(1 - iK)^{-1}. \tag{5.1}$$

In this case explicit evaluation of Riccati–Bessel and Riccati–Neumann functions is required. Although the convergence using asymptotic formulas of the last section was bad for $l > 0$ it is interesting to note that even if background scattering data evidently becomes very difficult to obtain from

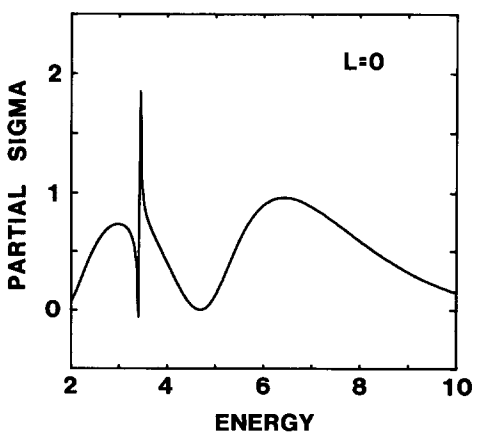


FIG. 11. The partial cross section σ_l for $l = 0$ calculated with the method of Johnson.[25]

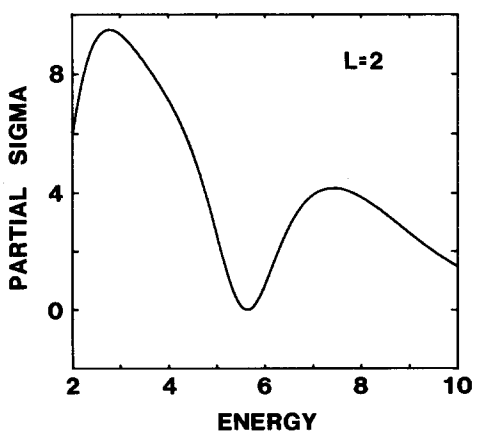


FIG. 13. Same as in Fig. 11 but for $l = 2$.

(4.13), the associated pole search was very efficiently determined through $m_{(\mathrm{TW})}$ for all values of l.

In order to see how the actual poles of the potential shown in Fig. 1 show up in the various spectral densities, we will in addition to the Kodaira densities (Figs. 2–4) and the Titchmarsh–Weyl densities (Figs. 5–7) also display the partial cross section (Figs. 11–13). For $l=0$ and $l=1$, the resonance closest to the real axis can be easily detected. The case $l=2$ is more difficult to analyze. In the Titchmarsh–Weyl case (Fig. 7) an outgrowth on the harmonic feature of the density tells us that something is to be found further down in the complex plane. In the Kodaira case ($l=2$) (Fig. 4) this feature is hardly noticeable. In the cross section the resonance shows up as a shoulderlike behavior (Fig. 13). Note that the partial cross section dips to zero as it should, and that the total cross section (Fig. 14) is quite insensitive in the corresponding region, except the small bump coming from the zero partial wave. In the next two sections we will return to the actual determination of the whole string of complex resonance poles.

VI. COMPLEX ROTATION

In earlier applications $m_{(\mathrm{TW})}$ was calculated on the real axis and continued out in the complex plane by an ansatz of rational fractions.[13,14] This method makes it possible to determine positions and widths of narrow resonances rather accurately. It is, however, dependent on the particular ansatz chosen. With the introduction of complex scaling[4,5,26] one is freed from this dependence and can, in principle, calculate pole strings in the whole section of the complex plane opened up by the rotation.[27–31] The exterior scaling[30,32,33] allows, furthermore, the accurate determination of resonances associated with sets of numerical potentials.[30]

Complex scaling is generated by the densely defined operator ($\rho = -i\hbar\, d/dx$)

$$A = 1/2(px + xp). \tag{6.1}$$

The extension problem here is by no means trivial; for instance, the operator

$$A' = 1/2(px^3 + x^3p) \tag{6.2}$$

has no such extension. For discussions on self-adjointness of A and related problems see Thirring.[17]

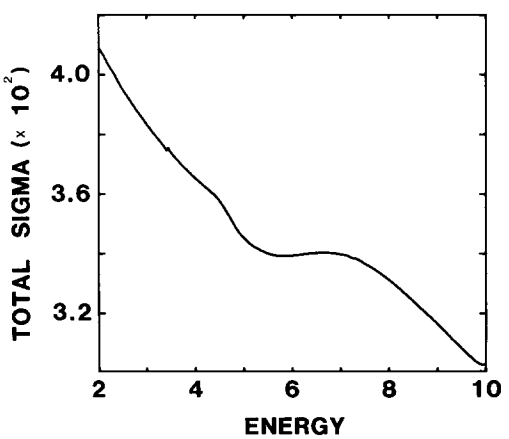


FIG. 14. Total cross section associated with the illustration potential evaluated with the method of Johnson.[25]

In this section we will assume that the potential $q(r)$ [or in the couple equation variant, the matrix $q(r)$] satisfies the conditions of Weyl's limit point. We will further require the analytic extension $q(\eta r)$, $\eta = e^{i\theta}$, to be such that there exists one unique square-integrable solution of the complex deformed differential operator along the complex path. Obviously this includes dilation analyticity as a special case. In addition to the spectral properties derived by Balslev and Combes[5] we may encounter new complex structures such as limiting rotation angles, different pole strings for different angle intervals, etc. (see below). We will here proceed to give some examples and demonstrate how the actual computation can be carried out.

By scaling $r \to \eta r$, $\eta = e^{i\theta}$ in (2.1) one obtains the differential equation ($p = \frac{1}{2}$)

$$u''_\eta(r) + 2\eta^2(\lambda - q(\eta r))u_\eta(r) = 0, \tag{6.3}$$

in which η appears as a complex parameter. To (4.3) we have the associated scaled Riccati equation

$$Z'_\eta = 2\eta^2(q(\eta r) - \lambda) - Z_\eta^2, \tag{6.4}$$

$$Z_\eta = u'_\eta(r)u_\eta^{-1}(r). \tag{6.5}$$

If needed, numeric potentials with analytic tails can also be handled by the exterior scaling method,[30,32,33] i.e.,

$$r \to \begin{cases} r, & r < R_0, \\ R_0 + e^{i\theta}(r - R_0), & r \geqslant R_0. \end{cases} \tag{6.6}$$

Before analyzing the effect of the scale transformations $r\eta$, or (6.6) we see immediately that (2.13) can be analytically extended as

$$m_{(\mathrm{TW}),\eta}(\lambda) = (Z_\eta^J Z_\eta^\psi + 1)/(Z_\eta^\psi - Z_\eta^J), \tag{6.7}$$

with Z_η^J being the logarithmic derivative of the square-integrable solution at infinity [that is, square integrable after rotation, provided $\arg(\eta)$ is sufficiently large, see below], i.e.,

$$Z_\eta^J \underset{r\to\infty}{\to} ik\eta \tag{6.8}$$

and

$$Z_\eta^\psi = \psi'_\eta \psi_\eta^{-1}, \tag{6.9}$$

with $\psi_\eta(r) \underset{r\to 0}{\to} 0$ (all values of l).

At the complex energies ϵ_0 for which (6.7) blows up, direct numerical integration of (6.3) yields the scaled outgoing wave which asymptotically behaves as

$$f_\eta \underset{r\to\infty}{\to} e^{ik\eta r}. \tag{6.10}$$

By putting $\eta = 1$ and evaluating (6.3) with $\lambda = \epsilon_0$, we get the asymptotically growing solution traditionally associated with a Gamow wave. In Fig. 15, the scaled and nonscaled Gamow waves are shown for a resonance energy of the potential in Fig. 1 for $l=0$. Note how the scaling forces the otherwise exponentially growing solution to zero for large r values.

It is interesting to investigate why a Gamow wave cannot be directly obtained from (2.1) or (2.3) with $\eta = 1$, using a complex λ and allowing for a change of the boundary condition at infinity. The situation has been analyzed from the

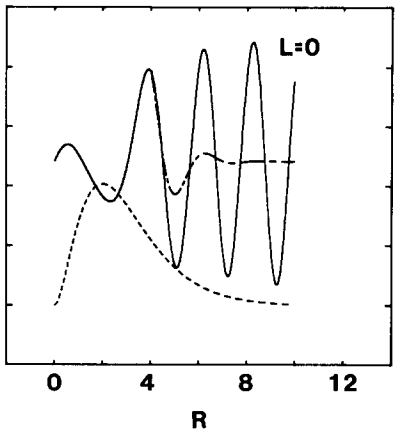


FIG. 15. Our illustration potential with $l = 0$ (dashed line). The full line shows the real part of the associated Gamow wave for the unscaled problem whereas the half-dashed line shows the corresponding exterior scaled Gamow wave. The difference between the unscaled and the scaled Gamow waves starts with the onset of the exterior scaling at $R_0 = 4.0$.

point of view of Titchmarsh–Weyl theory.[34] We will here give a more direct argument.

Since the actual stability problem of the nonlinear Riccati equation focuses on the logarithmic derivative, we will analyze the logarithmic derivative ($l = 0$ for simplicity)

$$Z(\eta,k,r) = \chi'(\eta,k,r)/\chi(\eta,k,r), \tag{6.11}$$

with

$$\chi(\eta,k,r) \underset{r\to\infty}{\cong} (e^{-ik\eta r} - S(\eta,k)e^{ik\eta r}). \tag{6.12}$$

First put $\eta = 1$. Then in the upper half-plane of k we have the following:

$$\lim_{r\to\infty} Z(k,r) \to -ik, \quad k \in C^+, \quad |S(k)| < \infty, \tag{6.13}$$

$$\lim_{r\to\infty} Z(\tilde{k},r) \to i\tilde{k} = -\kappa, \quad \tilde{k} \in C^+, \quad \lim_{k\to\tilde{k}} |S(k)| = \infty \tag{6.14}$$

($\kappa > 0$ for a bound state).

However, for k being continued into the lower half of the complex plane we have

$$\lim_{r\to\infty} Z(k,r) \to ik, \quad k \in C^-, \quad |S(k)| < \infty, \tag{6.15}$$

$$\lim_{r\to\infty} Z(\tilde{k},r) \to i\tilde{k}, \quad \tilde{k} \in C^-, \quad \lim_{k\to\tilde{k}} |S(k)| = \infty. \tag{6.16}$$

Hence, unlike when k belongs to the upper half-plane (physical sheet), the logarithmic derivative for the solution corresponding to a pole of $S(k)$ in the lower half-plane is the same as the logarithmic derivative for any other solution [except the one corresponding to $S(k) = 0$]. When one starts a numerical procedure for the Riccati equation with $k \in C^-$ and $Z(k,r)$ equal to ik in the asymptotic region of large r values, all solutions, except the one corresponding to $S(k) = 0$, will thus have the same asymptotic value of $Z(k,r)$ for r large. The solution corresponding to a pole of $S(k)$ mixes with other solutions in the Riccati procedure, and the latter becomes unstable.

However, if we spectrally deform the differential operator, i.e., let $r \to \eta r$, $\eta = \exp(i\theta)$, $\theta > 0$ ($k \in C^-$) and furthermore let the pole $\tilde{k}$ of $S(k)$ be exposed by the rotation η, i.e., $\arg \eta > |\arg \tilde{k}|$, it then follows immediately that

$$\lim_{r\to\infty} Z(\eta,k,r) = -ik\eta, \quad k \in C^-, \quad k \neq \tilde{k}, \tag{6.17}$$

$$\lim_{r\to\infty} Z(\eta,\tilde{k},r) = i\tilde{k}\eta, \quad \tilde{k} \in C^-. \tag{6.18}$$

Hence the resonance solution, for $k = \tilde{k}$ leads to a logarithmic derivative numerically different from other solutions in the asymptotic range. The spectral deformation makes the branch of the Riccati equation associated with the resonance solution stable, and the corresponding starting asymptotic value of $Z = i\tilde{k}\eta$ does not "mix" with other solutions.

We also note that S, being the quotient of the Jost functions,

$$S(k) = f^-(k)/f^+(k), \tag{6.19}$$

with

$$f^\pm(k) = W(f^\pm, \psi), \tag{6.20}$$

where $f^\pm(k,r)$ are the Jost solutions[35] and ψ the regular solution, seems to indicate that scaling changes $f^\pm(k)$ into $f^\pm(\eta,k)$ and $S(k)$ into $S(\eta,k)$. However, in a numerical computation of the analytic continuation of $S(k)$ it is, in general, necessary to calculate f^+ and f^- using different η factors. From the Jost–Pais theorem[36] one finds, for instance, that $f(k)$ can be identified with the determinant of $(1 - G_0V)$, with G_0 being the resolvent $(E + i0 - H_0)^{-1}$ and H_0 defined from (2.1) without the potential q. Since G_0V, spectrally deformed via the associated scaling operator, will not alter the determinant of $(1 - G_0V)$, one finds that f is independent of scaling for fixed k. This yields the necessary invariance condition for the Jost solutions and the S matrix. The present argument is somewhat sketchy. A more detailed account will be given elsewhere.

In the next section, we will comment on some examples of complex resonance pole determination and their role in building up the spectral density or scattering background.

VII. CALCULATION OF COMPLEX POLES

The present theory has been tested on several model potentials, both analytically as well as numerically defined ones.[24–28] We will here apply our formulas to the model potential used in the previous sections, see Eq. (3.3). For such model potentials (see below) a rigorous analytic theory that does not employ spectral deformation has been formulated by Meyer.[37,38] Unfortunately, this theory does not apply directly to the Coulomb problem or to the case of $l > 0$ or even for the Gaussian potential of Moiseyev *et al.*[39] As we have seen, a particular stability problem appears in the complex energy plane, which necessitates the employment of a complex deformation technique.

To find the location of resonant states we can either focus on the evaluation of the Wronskian corresponding to $f_l^+(k)$ or, equivalently, examine the function appearing in the denominator of Eq. (6.7). We have here chosen the latter procedure. This is a convenient choice, especially in the case of several coupled channels where instabilities frequently occur in the numerical integration of the wave function matrices. The log-derivative method does not suffer from such problems. If we, however, are interested in obtaining an ana-

lytic continuation of the whole S-matrix element we must return to the evaluation of the wave function.

In Fig. 16 we show the resulting pole strings for several l values for our model problem. The obtained pole strings have certain features in common with the ones associated with the Moiseyev-type potential studied by Rittby *et al.*[26,29] For both types of potentials there exist a real energy well above the potential maximum such that for all resonance energies ϵ_{res}, we will have

$$\text{Re}(\epsilon_{\text{res}}) \leqslant E_{\text{thresh}} < \infty . \tag{7.1}$$

The resonance that defines the complex threshold—the threshold resonance fulfilling the equality sign in (7.1)—defines further a natural partitioning point between the two classes of primary and secondary resonances.

The primary class contains all resonances and bound states with

$$\text{Im}(\epsilon_{\text{res}}) < \text{Im}(\epsilon_{\text{thresh}}), \tag{7.2}$$

and we assume that these resonances correspond to the peaks found in the velocity dependence of the scattering cross section. The secondary class of resonances would, with this assumption, only contribute to the background effects in potential scattering. We will return to this division of the pole distribution in connection with a paper discussing the Breit–Wigner ansatz and the analytic continuation of the whole S-matrix element.

In connection with the Moiseyev model potential two separate pole strings appeared: Rittby *et al.*[26] and Korsch *et al.*[40] An explanation of this phenomenon was given by Rittby *et al.*[29] in terms of the existence of a critical rotation angle. In rotations larger than the critical angle one needs to account for a Stokes-like behavior which, if erroneous boundary conditions are supplied, leads to false resonances. An interesting supporting analysis has been given by Atabek and Lefebvre.[41] A recent study of confinement potentials of Moiseyev and Katriel[42] shows that different pole strings for different rotation angles are possible.

At first glance it is remarkable that the pole string associated with the Bain potential has resonances for $l > 1$. This can, however, be understood if it is recognized that only the real part of the potential is affected by the addition of the centrifugal term. One can also notice that the pole spectrum

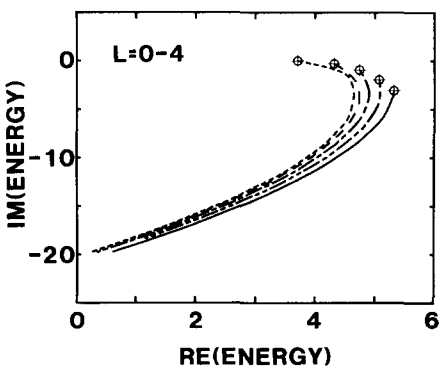


FIG. 16. The pole strings associated with different l values of the illustration potential. The l values range from $l = 0$ to $l = 4$. The first pole for $l = 0$ has the lowest real energy of all the first poles in the respective strings. Note that all strings, although somewhat displaced, follow a similar trajectory.

associated with the potential for $l = 3$ and 4 only has a secondary class. One sees that when poles reach the threshold, i.e., become secondary, the effect on the spectral density is a shoulderlike behavior, see, e.g., Rittby *et al.*[31] The secondary poles or "background poles," if located, can be used to model the S matrix via a pole ansatz. Recent progress along those lines has been obtained by Atabek *et al.*[43] and Korsch *et al.*[44]

VIII. CONCLUDING REMARKS

In this work we have emphasized the use of the Titchmarsh–Weyl approach in connection with studies of analytic properties of scattering theory. In particular we have utilized the completeness of the TW theory to combine the description of the continuum with that of the bound states. In addition to yielding information of the relevant point spectra it also supplies a rigorous inclusion of the appropriate scattering background.

In the development of the theory, various types of spectral densities related to exact completeness relations were derived, analyzed, and computed. It was also pointed out that the extension of Hermitian operators is far from a trivial problem but necessary for further analysis. The analytic continuation of the unique square-integrable solution associated with the (limit-point) differential operator, defined as a two singular endpoint problem, was further carried out, calculated, and interpreted.

Since our main objective is to find the underlying analytic structure associated with the behavior of the actual spectral density as a function of k (or E) on the real axis, a theory that elucidates the full analytic content of the differential operator is needed. Titchmarsh–Weyl theory is precisely the answer to this demand, which furthermore is a formulation that is computationally very attractive. As has been shown here, the combination of complex scaling with TW theory allows a numerical description of pole strings valid far out into the complex plane.

The actual need for doing complex scaling has also been analyzed in connection with general numerical integration techniques applicable to potentials of a very general class. Note that equations are written so that a generalization to coupled equations is straightforward.[45]

In comparison with WKB approaches, one may also bypass the need for a convergence analysis as well as the investigation of problems related to Stokes phenomena. Complex scaling and Titchmarsh–Weyl theory focus directly on the square-integrable solution, i.e., a change of asymptotic branch is automatically picked up. Furthermore one can also apply the TW theory numerically to the case when the asymptotic form of the solution is not known, providing that the potential belongs to Weyl's limit-point classification.[46]

Finally we would like to stress that the present formulations provides a first step towards an alternative solution of the inverse problem. Using pure resonance information as well as the remaining background, this approach can in principle be applied to, for example, barrier penetration phenomena. Work along these lines is in progress.

ACKNOWLEDGMENTS

This work was carried out (in part) when one of us (E. B.) was an Uppsala Visiting Professor at the Quantum Theory Project of the University of Florida. E. B. would like to thank the University and in particular the Office of Academic Affairs for the warm hospitality that was shown to him during that period.

[1] *Electronic and Atomic Collisions Abstracts of Contributed Papers*, edited by J. Eickler, W. Fritsch, I. V. Hertel, N. Stolterfoht, and U. Wollie (ICPEAC, Berlin, 1983), p. 701.

[2] *Production and Physics of Highly Charged Ions*, edited by L. Liljeby (Alden, Oxford, 1983).

[3] R. Lavine, in *Atomic Scattering Theory*, edited by J. Nuttall (Univ. Western Ontario Press, London, Ontario, 1978).

[4] J. Aguilar and J. M. Combes, Commun. Math. Phys. **22**, 269 (1971).

[5] E. Balslev and J. M. Combes, Commun. Math. Phys. **22**, 280 (1971).

[6] H. Weyl, Ann. Math. **68**, 220 (1910).

[7] E. C. Titchmarsh, *Eigenfunction Expansions Associated with Second-Order Differential Equations* (Oxford U. P., London, 1946).

[8] E. Brändas, M. Hehenberger, and H. V. McIntosh, Int. J. Quantum Chem. **9**, 103 (1975).

[9] M. Hehenberger, B. Laskowski, and E. Brändas, J. Chem. Phys. **65**, 4559 (1976).

[10] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (Mc Graw–Hill, New York, 1955).

[11] N. I. Akhiezer, *The Classical Moment Problem* (Oliver and Boyd, Edinburgh, 1965).

[12] W. N. Everitt and C. Bennewitz, in "Tribute to Åke Pleijel," *Proceedings of the Conference*, September 1979 (Department of Mathematics, University of Uppsala, Uppsala, Sweden, 1980), p. 49.

[13] M. Hehenberger, H. V. McIntosh, and E. Brändas, Phys. Rev. A **10**, 1494 (1974).

[14] M. Hehenberger, P. Froelich, and E. Brändas, J. Chem. Phys. **65**, 4571 (1976).

[15] N. Elander, M. Hehenberger, and P. R. Bunker, Phys. Scripta **20**, 631 (1979).

[16] M. Reed and B. Simon, *Methods of Modern Mathematical Physics* (Academic, New York, 1972–79), Vols. I–IV.

[17] W. Thirring, *Quantum Mechanics of Atoms and Molecules* (Springer, New York, 1981).

[18] R. D. Richtmyer, *Principles of Advanced Mathematical Physics* (Springer, New York, 1978).

[19] E. Brändas and M. Hehenberger, *Lecture Notes in Mathematics Vol. 415* (Am. Math. Soc., Providence, RI, 1974), p. 316.

[20] E. Hille, *Lectures on Ordinary Differential Equations* (Addison–Wesley, London, 1969).

[21] S. D. Wray, Proc. R. Soc. Edinburgh Ser. A **14**, 14 (1974/75).

[22] K. Kodaira, Am. J. Math. **71**, 921 (1949).

[23] R. G. Newton, *Scattering Theory of Waves and Particles* (Mc Graw–Hill, New York, 1966).

[24] R. A. Bain, J. N. Bardsley, B. R. Junker, and C. V. Sukumar, J. Phys. B **7**, 2189 (1974).

[25] B. R. Johnson, J. Comput. Phys. **13**, 445 (1973).

[26] B. Simon, Ann. Math. **97**, 247 (1973).

[27] M. Rittby, N. Elander, and E. Brändas, Phys. Rev. A **24**, 1636 (1981).

[28] M. Rittby, N. Elander, and E. Brändas, Mol. Phys. **45**, 553 (1982).

[29] M. Rittby, N. Elander, and E. Brändas, Phys. Rev. A **26**, 1804 (1982).

[30] M. Rittby, N. Elander, and E. Brändas, Int. J. Quantum Chem. **23**, 865 (1983).

[31] M. Rittby, N. Elander, and E. Brändas, Chem. Phys. **87**, 55 (1984).

[32] B. Simon, Phys. Lett. A **71**, 211 (1979).

[33] J. D. Morgan and B. Simon, J. Phys. B **14**, L167 (1981).

[34] E. Brändas, P. Froelich, C. Obcemea, N. Elander, and M. Rittby, Phys. Rev. A **26**, 3656 (1982).

[35] R. Jost, Physica **12**, 509 (1946).

[36] R. Jost and A. Pais, Phys. Rev. **82**, 840 (1951).

[37] H-D. Meyer, J. Phys. B **16**, 2265 (1983).

[38] H-D. Meyer, J. Phys. B **16**, 2785 (1983).

[39] N. Moiseyev, P. R. Certain, and F. Weinhold, Mol. Phys. **36**, 1613 (1978).

[40] H. J. Korsch, H. Laurent, and R. Möhlenkamp, Phys. Rev. A **26**, 1802 (1982).

[41] O. Atabek and R. Lefebvre, Nuovo Cimento B **76**, 176 (1983).

[42] N. Moiseyev and J. Katriel, Chem. Phys. Lett. **105**, 194 (1984).

[43] O. Atabek, R. Lefebvre, and M. Jacon, J. Phys. B **15**, 2689 (1982).

[44] H. J. Korsch, R. Möhlenkamp, and H-D. Meyer, J. Phys. B **17**, 2955 (1984).

[45] N. Elander, M. Rittby, and E. Brändas, Int. J. Quantum Chem. **22**, 445 (1982).

[46] E. Engdahl, E. Brändas, M. Rittby, and N. Elander (to be published).

On extensions of supersymmetry algebras in $d = 4$ and 11

E. Ahmed,[a] B. L. Rands, and J.G. Taylor
Department of Mathematics, King's College, University of London, Strand, London, WC2R 2LS, England

(Received 11 October 1984; accepted for publication 21 March 1985)

SUSY algebras, in $d = 4$ and 11, are analyzed, extended by the addition of both spinorial and bosonic central charge generators. It is found that only limited extensions are possible, on use of Jacobi identities. Casimirs, irreps, and spin reduction are discussed, and the problems of constructing actions for SUSY and supergravity theories as full superspace actions are considered.

I. INTRODUCTION

There has recently been a great deal of interest in theories of gravity in higher dimensions[1] and especially in those involving N-extended supergravity (N-SGR). In particular, the process of spontaneous compactification has led to various possible models in which the "standard model" gauge group $SU(3)\times SU(2)\times U(1)$ is the symmetry of the compactified solution.[2–4]

Even though these models still have problems associated with the possible fermion representations there remains the more fundamental question as to their quantum properties. In particular, it is not known if any version of supergravity, in higher dimensions or not, has sufficient cancellation of the usual ultraviolet divergences to be finite on-shell.

The most appropriate way to tackle this problem is to construct a superspace version of the corresponding theory so as to "automate" the Fermi–Bose ultraviolet divergence cancellation mechanism. Such an approach has been used with great effect to prove the finiteness of $d = 4, N = 4$ super Yang–Mills theory ($N = 4$ SYM) to all orders of perturbation theory.[5,6] A similar proof for $d = 4, N = 8$ SGR is lacking due to the existence of a barrier at $N = 3$ for the construction of extended supergravities in superspace.[7,8] Since the $N = 3$ barrier in $d = 4$ becomes even more restrictive with a barrier at $N = 1$ for the case of $d = 10$ and even $N = 0$ for $d = 11$, it is necessary to take seriously the question of evading these various barriers for constructing N-SGR's in $d \geqslant 4$.

The methods used in the proof of finiteness of $N = 4$ SYM were either by use of light-cone variables and gauge,[5] or by use of an $N = 2$ superfield formulation of the theory.[6] A light-cone or $N/2$ superspace formulation of extended supergravities in $d \geqslant 4$ seems possible, especially following the recent construction of a component version of $N = 1$ supergravity in light-cone coordinates in closed form.[9] Even if a light-cone superspace version of maximal N-SGR in various dimensions could be constructed it is not clear that this would allow a proof of finiteness.[10] The discovery of the $N = 4$ relaxed hypermultiplet[11,12] should allow a complete version of $N = 8$ SGR in terms of $N = 4$ superfields. However, this theory would suffer from the existence of a putative three-loop counterterm.[13] Higher-dimensional SYM and SGR's are expected to be even more ultraviolet divergent.

The explicit preservation of maximal symmetry may be of importance in achieving finiteness of $N = 8$ SGR. There is an alternative approach to the construction of superspace versions of N-SYM and N-SGR beyond the N barrier which uses the notion of "spin reduction" and which avoids the loss of explicit Lorentz invariance (as in the light-cone gauge) or full internal symmetry (as in the $N/2$ version of N-SYM or N-SGR).

Spin reduction arises when only massless representations are allowed, through the imposition of a suitable constraint. In order to achieve this, it is essential that off-shell central charges are introduced. The appropriate constraints can then be obtained by means of consideration of representations of supersymmetry algebras in the presence of central charges[14–16] or by means of the Casimirs of the supersymmetry algebra.[17] It has proved possible to construct a full superspace version of $N = 4$ SYM by these means,[18] using the analysis of linearized actions for $N = 2$ (Ref. 19) and $N = 4$ and 8 (Ref. 20).

A similar approach to $d = 11, N = 1$ SGR will require the introduction of central charges in the supersymmetry algebra. So far we have assumed that the central charges are always bosonic. However, a geometric formulation of $d = 11, N = 1$ SGR required the introduction of spinorial central charges.[21] We must therefore also consider such a possibility seriously in the construction of full superspace versions of extended supergravities.

The need to consider additional fermionic central charges also stems from the problem of obtaining a measure in full superspace which has the correct dimensions. Thus for $d = 4$, N-SGR requires $2(N\text{-}1)$ extra central charge dimensions,[22] whilst in $d = 11, N = 1$ SGR requires seven extra such dimensions. Both of these cases are not natural, in the sense that the $d = 4$ case has $N(N\text{-}1)$ such bosonic (real) central charges Z^{ij} in the algebra for the spinorial SUSY generators S^i_α (we use chiral-projected notation in $d = 4$ with $S_{\alpha\pm} = \frac{1}{2}[(1 \pm i\gamma_5)S]_\alpha$)

$$[S^i_{\alpha+}, S^j_{\beta+}]_+ = C^{-1}_{\alpha+\beta+} Z^{ij}. \tag{1.1}$$

Thus for $N = 8$ there are 56 such central charges, whereas only 14 are required to give a measure in full superspace $d^4x \prod^{14}_{l=1} dz^l \prod^8_{i=1} d^4\theta^i$ with suitable length dimension of 2. In $d = 11$ there are 55 two-tensor central charges $Z_{\mu\nu}$ and ${}^{11}C_5$ five-tensors $Z_{\mu\nu\lambda\sigma\rho}$ arising in the anticommutator

$$[S_\alpha, S_\beta]_+ = \sum_{k=1,2,5} (Z_k/k!)(\Gamma^k C^{-1})\alpha\beta, \tag{1.2}$$

where $Z_1 = \{P_\mu\}$, $Z_2 = \{Z_{\mu\nu}\}$, $Z_5 = \{Z_{\mu\nu\lambda\sigma\rho}\}$, with associated antisymmetrized products $\Gamma^1, \Gamma^2, \Gamma^5$ defined in the usual way from the 11-dimensional Dirac algebra. Again it appears unnatural to attempt to choose a subset of seven of

[a] On leave of absence from Department of Mathematics, Faculty of Science, Mansoura University, Mansoura, Egypt.

 0022-2488/85/102659-07$02.50

these central charges Z_2 and Z_5. It is possible that additional spinorial central charges will redress this imbalance and lead to a satisfactory full superspace measure. This paper is an attempt to analyze this possibility.

An immediate question which arises from the introduction of spinorial central charges is the existence of any limit to the number of these charges which may be introduced. Since there are further bosonic charges which may also arise, this question does not have an obvious answer. We follow the analysis of $N=1$ SUSY algebras in $d=2,3,4 \bmod 8$ (Ref. 23) and use Jacobi identities to impose constraints on possible additional central charges. We do this only for the Poincaré case, unlike the above-mentioned discussion which is mainly concerned with the conformal and de Sitter cases; our analysis could be extended to these latter situations by similar techniques.

We consider the case of N-SUSY in $d=4$ in the next section, and show that provided there are no tensorial central charges then the SUSY algebra may only be extended to a spinorial central charge $S'_{\alpha i}$ and a new bosonic central charge Z^i_j which is a Hermitian $N \times N$ matrix. Our spinor S' is defined by the commutator

$$[P_\mu, S'^i_{\alpha +}]_- = \tfrac{1}{2}\gamma^{\beta -}_{\mu\alpha +} S'^i_{\beta -} \tag{1.3}$$

and its complex conjugate. Other additional spinorial central charges may be introduced which commute with P_μ, but do not obviously relate to those arising from the reduction of those in $d=11$ (Ref. 21). We hope to discuss aspects of the former elsewhere. In the following section we extend our analysis to $d=11$ and show that a similar limitation arises. Construction of the Casimirs for the new algebras is then considered in Sec. IV, and the N barrier against the construction of N-SYM or N-SGR in superspace shown to exist for these extended algebras in $d=4$ and 11. Spin reduction is then discussed in Sec. V, and a suitable constraint obtained which achieves off-shell character for all the central charges. We then consider in the following section the dimension-matching problem in the light of our results. A final section on extensions of our work concludes the paper.

II. THE EXTENDED ALGREBRA IN $d=4$ DIMENSIONS

Our aim in this section is to obtain a maximal extension of the usual supersymmetry algebra with central charges in four dimensions via the introduction of a new set of spinor generators which arise from the commutators of the momentum generator P_μ with the chiral supersymmetry generators $S^i_{\alpha +}$ and $S_{\alpha -i}$; this approach has already been discussed for $N=1$ supersymmetry algebras in various dimensions[23] after such a spinor generator was first introduced in an 11-dimensional $N=1$ context.[21]

Our discussion here is thus not completely general, since we sidestep the introduction of independent spinor generators (i.e., those not arising from $[P_\mu, S]$ commutators), which would be distinguished from extra supersymmetry generators by their transformation properties under the internal symmetry group, e.g., $S^{ij}_{\alpha +}, S^{ijk}_{\alpha +} S_{\alpha +i}$, etc. We also avoid here the introduction of tensorial central charges, restricting such generators to be Lorentz scalars. Both these omissions deserve, we feel, further investigation.

We begin with the Poincaré algebra, for the momentum generator P_μ, and the angular momentum $J_{\mu\nu}$. We have

$$[P_\mu, P_\nu] = 0, \tag{2.1}$$

$$[J_{\mu\nu}, P_\rho] = i(P_\mu g_{\nu\rho} - P_\nu g_{\mu\rho}), \tag{2.2}$$

$$[J_{\mu\nu}, J_{\lambda\sigma}] = -i(g_{\mu\lambda} J_{\nu\sigma} + g_{\nu\sigma} J_{\mu\lambda} - g_{\mu\sigma} J_{\mu\lambda} - g_{\nu\lambda} J_{\mu\sigma}), \tag{2.3}$$

where the metric $g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$.

Next we introduce the supersymmetry generators $S^i_{\alpha +}, S_{\alpha -i}$ [where $(S^i_{\alpha +})^* = S_{\alpha -i}$] and the central charge generators $Z^{ij}, (Z^{ij})^*$ with anticommutation relations

$$\{S^i_{\alpha +}, S_{\beta -j}\} = (\not{p} C^{-1})_{\alpha + \beta -} \delta^i_j, \tag{2.4}$$

$$\{S^i_{\alpha +}, S^j_{\beta +}\} = Z^{ij} C^{-1}_{\alpha +\beta +}, \tag{2.5}$$

$$\{S_{\alpha -i}, S_{\beta -j}\} = (Z^{ij})^* C^{-1}_{\alpha -\beta -}, \tag{2.6}$$

where C is the charge conjugation matrix.

The supersymmetry generators are spinors, with commutation relations

$$[J_{\mu\nu}, S^i_{\alpha +}] = -i(\sigma_{\mu\nu})^{\beta +}_{\alpha +} S^i_{\beta +}, \tag{2.7}$$

$$[J_{\mu\nu}, S_{\alpha -i}] = -i(\sigma_{\mu\nu})^{\beta -}_{\alpha -} S_{\beta -i}, \tag{2.8}$$

where $\sigma_{\mu\nu} = \frac{1}{4}[\gamma_\mu, \gamma_\nu]$. Then, $Z^{ij}, (Z^{ij})^*$ will be scalars under Lorentz transformations, i.e.,

$$[J_{\mu\nu}, Z^{ij}] = 0. \tag{2.9}$$

Normally, the next step in the derivation of the supersymmetry algebra is to write down a general form for the $[P_\mu, S^i_{\alpha +}]$ commutator. Using Lorentz covariance, unless new spinor generators are introduced this could only be of the form $[P_\mu, S^i_{\alpha +}] = C^{ij} \gamma^{\beta -}_{\mu\alpha +} S_{\beta -j}$ for some constant C^{ij} of dimension -1. But, no such constant is available to us, and we are thus required to put $[P_\mu, S^i_{\alpha +}]$ equal to zero.

Instead, we introduce new spinors $S'_{\alpha +i}$, and their complex conjugates $S'^i_{\alpha -}$, and write

$$[P_\mu, S^i_{\alpha +}] = \tfrac{1}{2}\gamma^{\beta -}_{\mu\alpha +} S'^i_{\beta -}, \tag{2.10}$$

$$[P_\mu, S_{\alpha -i}] = \tfrac{1}{2}\gamma^{\beta +}_{\mu\alpha -} S'_{\beta +i}. \tag{2.11}$$

We now postulate anticommutation relations for S'_+ and S'_- both among themselves, and with the S_+ and S_- supersymmetry generators. In doing this, we restrict ourselves to the introduction only of new Lorentz scalar generators, in order to avoid problems of working with tensorial objects and subsequent breaking of the Lorentz invariance on reduction to a single central charge. We take

$$\{S'_{\alpha +i}, S^j_{\beta +}\} = C^{-1}_{\alpha +\beta +} Z^j_i, \tag{2.12}$$

$$\{S'^i_{\alpha -}, S_{\beta -j}\} = C^{-1}_{\alpha +\beta -} (Z^j_i)^*, \tag{2.13}$$

$$\{S'_{\alpha +i}, S_{\beta -j}\} = 0, \tag{2.14}$$

$$\{S'^i_{\alpha -}, S^j_{\beta +}\} = 0, \tag{2.15}$$

$$\{S'_{\alpha +i}, S'_{\beta +j}\} = C^{-1}_{\alpha +\beta +} Z'_{ij} \quad (Z'_{ij} = -Z'_{ji}), \tag{2.16}$$

$$\{S'^i_{\alpha -}, S'^j_{\beta -}\} = C^{-1}_{\alpha -\beta +} (Z'_{ij})^*, \tag{2.17}$$

$$\{S'_{\alpha +i}, S'^j_{\beta -}\} = 0. \tag{2.18}$$

We may now analyze the remainder of the algebra, together with any restrictions on our assumed anticommutation relations by means of the graded Jacobi identities; if B denotes a Bose, and F a Fermi operator, these are

(B_1,B_2,B_3): $[B_1,[B_2,B_3]]+[B_2,[B_3,B_1]]$
$+[B_3,[B_1,B_2]]=0,$ (2.19)

(B_1,B_2,F_1): $[B_1,[B_2,F_1]]+[B_2,[F_1,B_1]]$
$+[F_1,[B_1,B_2]]=0,$ (2.20)

(B_1,F_1,F_2): $[B_1,\{F_1,F_2\}]+\{F_1,[F_2,B_1]\}$
$-\{F_2,[B_1,F_2]\}=0,$ (2.21)

(F_1,F_2,F_3): $[F_1,\{F_2,F_3\}]+[F_2,\{F_3,F_1\}]$
$+[F_3,\{F_1,F_2\}]=0.$ (2.22)

For example, considering the identity $(P_\mu,S^i_{\alpha+},S_{\beta-j})$, we have

$$[P_\mu,\{S^i_{\alpha+},S_{\beta-j}\}]+\{S^i_{\alpha+},[S_{\beta-j},P_\mu]\}$$
$$-\{S_{\beta-j},[P_\mu,S^i_{\alpha+}]\}=0.$$

This may be rewritten as

$$[P_\mu,P_\nu](\gamma^\nu C^{-1})_{\alpha+\beta-}\,\delta^i_j-\tfrac{1}{2}(\gamma_\mu)^{\gamma+}_{\beta-}\,\{S^i_{\alpha+},S'_{\gamma+j}\}$$
$$-\tfrac{1}{2}(\gamma_\mu)^{\gamma-}_{\alpha+}\,\{S_{\beta-j},S''^i_{\gamma-}\}=0$$

or

$$(\gamma_\mu)^{\gamma+}_{\beta-}C^{-1}_{\gamma+\alpha+}Z^i_j+(\gamma_\mu)^{\gamma-}_{\alpha+}C^{-1}_{\gamma-\beta-}(Z^j_i)^*=0,$$

or finally as

$$(\gamma_\mu C^{-1})_{\alpha+\beta-}(Z^i_j+(Z^j_i)^*)=0.$$

This requires

$$(Z^j_i)^*=-Z^i_j,\tag{2.23}$$

so that Z^j_i is an anti-Hermitian matrix.

We list the important identities, and the properties they impose upon our algebra:

(P,P,S_+) leads to $[P_\mu,S''^i_{\alpha-}]=0,$
(P,S_+,S_+) leads to $[P_\mu,Z^{ij}]=0,$
(P,S_+,S'_+) leads to $[P_\mu,Z^j_i]=0,$
(P,S_+,S'_-) leads to $\{S''^i_{\alpha-},S''^j_{\beta-}\}=0,$
so $Z''^{ij}=(Z''^{ij})^*=0,$
(S_+,S_-,S'_+) leads to $[S_{\alpha-i},Z^k_j]=0,$
(S'_+,S_+,S_+) leads to $[S'_{\alpha+i},Z^{jk}]=0,$
(S'_+,S_-,S_-) leads to $[S'_{\alpha+i},(Z^{jk})^*]=0,$
(S'_+,S'_-,S_+) leads to $[S''^i_{\alpha-},Z^k_j]=0,$
(S_-,S_+,S_+) leads to $[S_{\alpha-i},Z^{jk}]$
$=S''^{ij}_{\alpha-}\,\delta^k_i-S''^{ik}_{\alpha-}\,\delta^j_i.$ (2.24)

The identities (Z,S_+,S_-) and (P,S_+,Z) imply that $[S^i_{\alpha+},Z^{jk}]$ is an object which commutes with P_μ and anticommutes with $S_{\beta-l}$. Without introduction of more spinor generators, we must write this as $[S^i_{\alpha+},Z^{jk}]=M^{ijk}_l S''^l_{\alpha+}$, where M^{ijk}_l is a Lorentz scalar of dimension zero. In the absence of such an object, we set it to zero to obtain $[S^i_{\alpha+},Z^{jk}]=0.$

Then

(Z,S_+,S_+) leads to $[Z^{ij},Z^{jk}]=0,$ (2.25)
(Z,S_+,S'_+) leads to $[Z^{ij},Z^l_k]=0,$ (2.26)
(Z,S_-,S_-) leads to $[(Z^{lm})^*,Z^{ij}]$
$=Z^i_l\,\delta^j_m-Z^i_m\,\delta^j_l-Z^j_l\,\delta^i_m+Z^j_m\,\delta^i_l.$ (2.27)

All other Jacobi identities between $P,S_+,S_-,Z^{ij},(Z^{ij})^*$, S'_+,S'_-, and Z^j_i are then satisfied trivially.

We mentioned briefly how the algebra may be enlarged still further if we allow spinor generators with more complicated internal symmetry structure to occur. We may introduce

$$V^{ijk}_{\alpha+}=[S^i_{\alpha+},Z^{jk}],$$

which will satisfy the Jacobi identity (S_+,S_+,S_+) if $V^{ijk}_{\alpha+}$ is totally antisymmetric on ijk. Further analysis of the Jacobi identities sets the graded commutator brackets of all other generators except $S^i{}_+$ and Z^{ij} (and V_+, itself) with V_+ to be zero. Other commutation relations are preserved.

The brackets of V_+ with $S^i{}_+,Z^{ij}$ may be nonzero, but further generators with "exotic" internal symmetry representations would be required if these are to be incorporated. Setting these to zero then restricts the anticommutator $\{V^{ijk}_{\alpha+},V^{lmn}_{\beta+}\}$ to be zero, and thus $V^{ijk}_{\alpha+}$ (anti) commutes with everything.

The implicit use of Lorentz covariance throughout ensures that S'_+,S'_-, and Z^j_i transform as spinors and scalars, respectively,

$$[J_{\mu\nu},S'_{\alpha+i}]=-i(\sigma_{\mu\nu})^{\beta+}_{\alpha+}S'_{\beta+i},$$
$$[J_{\mu\nu},S''^i_{\beta-}]=-i(\sigma_{\mu\nu})^{\beta-}_{\alpha-}S''^i_{\beta-},$$
$$[J_{\mu\nu},Z^j_i]=0,$$

as can be checked explicitly from the relevant Jacobi identities. We have thus determined, under the above provisos, the maximum extension of the supersymmetry algebra when spinorial and further bosonic central charge generators are allowed.

III. THE EXTENDED ALGEBRA IN 11 DIMENSIONS

In this section we will try to extend the $d=11$ algebra in an analogous way to that used in $d=4$. First we list the $d=11$ algebra mentioned in[23]:

$$\{S_\alpha,S_\beta\}=\sum_{k=1,2,5}\frac{1}{k!}Z_k(\Gamma^kC^{-1})_{\alpha\beta},$$
$$[Z_k,S_\alpha]=(-1)^ku^k(\Gamma_kS')_\alpha\quad\text{(no sum over } k),$$
$$[Z_k,Z_l]=0,\quad\forall k,l=1,2,5,$$
$$[J_{\mu\nu},S']=-\tfrac{1}{2}\Gamma_{\mu\nu}S',\tag{3.1}$$

where S' commutes with the remaining generators and u_k satisfies $u_1+5u_2-6u_5=0$. This implies that if $Z_2=Z_5=0$ then also $S'=0$.

We start the extension by assuming the commutation relations, obtained from (3.1) by setting $u_k=-1(k=1,2,5)$

$$\{S_\alpha,S_\beta\}=\sum_k\frac{1}{k!}Z_k(\Gamma^kC^{-1})_{\alpha\beta},$$
$$[Z_k,S_\alpha]=(-1)^{k-1}(\Gamma_kS')_\alpha,\tag{3.2}$$
$$[P_\mu,P_\nu]=0.$$

We then investigate the case that S' is not a central charge so that we may introduce further bosonic central charges Z'_k,Z''_k by the definitions

$$\{S_\alpha,S'_\beta\}=\sum_{n=0}^{5}\frac{1}{n!}Z''_n(\Gamma^n C^{-1})_{\alpha\beta},$$

$$\{S'_\alpha,S'_\beta\}=\sum_{k=1,2,5}\frac{1}{k!}Z'_k\,(\Gamma^k C^{-1})_{\alpha\beta}. \tag{3.3}$$

An analysis of Jacobi identities leads to the result that

$$Z''_0=Z''_{\mu\nu}=0,\quad\text{and thence}\quad Z'_\mu=0,$$
$$(6\eta_{\mu[\nu}Z'_{\rho\sigma]}-4\eta_{\nu[\rho}Z'_{\sigma]\mu})-\mu\leftrightarrow\nu=0,$$
$$(10\eta_{\sigma[\mu_1}Z'_{\mu_2\cdots\mu_5]\rho}-[2/(5!)^2]\epsilon_{\sigma MN}\epsilon_\rho^{NN_1}Z'_{N_1})$$
$$-\rho\leftrightarrow\sigma=0,$$

which gives—by contracting the indices $\mu\,\rho$ in the first, and $\sigma\,\mu_1$ in the second equation

$$Z'_k=0,\quad\forall k=1,2,5. \tag{3.4}$$

Thus, finally, we get the closed algebra

$$[Z_k,S']=0,\quad\forall k=1,2,5,$$
$$[Z''_n,S]=[Z''_n,Z_k]=[Z''_n,Z''_m]=0,$$
$$\forall m,n=1,3,4,5,$$
$$[P_\mu,Z_{\rho\sigma}]=4\eta_{\mu[\rho}Z''_{\sigma]}+2Z''_{\rho\sigma\mu},$$
$$[P_\mu,Z_{\nu_1\cdots\nu_5}]$$
$$=10\eta_{\mu[\nu_1}Z''_{\nu_2\cdots\nu_5]}+(2i/5!)\epsilon_{\mu\nu_1\cdots\nu_5}MZ'^M,$$
$$[Z_{\mu\nu},Z_{\rho\sigma}]=0,$$
$$[Z^{\rho\sigma},Z_{\nu_1\cdots\nu_5}]=-\{40\,\delta^\rho_{[\nu_1}\delta^\sigma_{\nu_2}Z''_{\nu_3\nu_4\nu_5]}$$
$$+(2i/4!)\epsilon^{\rho\sigma}_{\lambda_1\cdots\lambda_4\nu_1\cdots\nu_5}Z'^{\lambda_1\cdots\lambda_4}$$
$$+20\,\delta^\rho_{[\nu_1}Z''^\sigma_{\nu_2\cdots\nu_5]}\}, \tag{3.5}$$
$$Z_{\mu_1\cdots\mu_5},Z_{\lambda_1\cdots\lambda_5}$$
$$=\frac{5!}{2}\left\{\frac{2i}{2!}\epsilon_{M\nu M'}Z'^\nu+\frac{5i}{5!}\epsilon_{\mu_1\cdots\mu_4\nu_2\nu_3M'}Z''^{\nu_2\nu_3}_{\mu_2}\right.$$
$$\left.+\frac{10i}{3(5i)}\epsilon_{\mu_1\mu_2\mu_3\nu_3\nu_4\nu_5M'}Z''^{\nu_3\nu_4\nu_5}_{\mu_5\mu_4}\right\}-M\leftrightarrow M',$$

where $M=\mu_1,\ldots,\mu_5, M'=\lambda_1,\ldots,\lambda_5$.

We can therefore state that starting from an algebra satisfying (3.2) the maximum extension is to introduce another type of central charge Z''_n and the maximal algebra is then given by the commutation relations (3.5). We point out here that there is no possible introduction of further spinorial central charges as there was in the four-dimensional case for the bracket $[S^i_{\alpha+},Z^{jk}]_-$. This is because we assumed the stronger commutator in (3.1) defining $S'_\alpha,[Z_k,S_\alpha]_-=(-1)^k u^k(\Gamma_k S')\alpha$. This could have been relaxed, so introducing more spinorial central charges.

IV. ON THE CASIMIRS OF HIGHER-DIMENSION SUPERSYMMETRY ALGEBRAS

We first investigate the Casimirs of the Poincaré algebra in d dimensions starting from the algebra (2.1)–(2.3), where $\mu=1,2,\ldots,d$. An obvious generalization of the $d=4$ Casimirs P^2_μ and W^2_μ is P^2_μ and $(W^{\mu_1\mu_2\cdots\mu_{d-3}})^2:=W^{\mu_1\cdots\mu_{d-3}}W_{\mu_1\cdots\mu_{d-3}}$, where

$$W^{\mu_1\cdots\mu_{d-3}}:=\epsilon^{\mu_1\cdots\mu_{d-3}\rho\sigma\lambda}J_{\rho\sigma}P_\lambda. \tag{4.1}$$

This can alternatively be written as $J_{[\rho\sigma}P_{\lambda]}$ and called $W^{(3)}_{\rho\sigma\lambda}$. Other Casimirs can be obtained by including further J operators; for example, $(W^{(5)})^2$ is a Casimir where $W^{(5)}_{\mu\nu\rho\sigma\lambda}:=J_{[\mu\nu}J_{\rho\sigma}P_{\lambda]}$ (or equivalently, it can be called W^{d-5}). In general there are $(d/2-1)$ such Casimirs for even d and $(d-1)/2$ for d odd. The Casimirs $(W^{(r)})^2$ $(r=3,5,\ldots,[d/2]+1)$ are just those expected for the group I $O(d-1,1)$.

Next we study the Casimirs of the graded Poincaré algebra without central charges, i.e., the algebra given by

$$\{S_\alpha,S_\beta\}=P_\mu(\Gamma^\mu C^{-1})_{\alpha\beta},\quad[P_\mu,S]=0,$$
$$[J_{\mu\nu},S]=-\tfrac{1}{2}\Gamma_{\mu\nu}S.$$

For such an algebra P^2 is still a Casimir but the other Casimirs have to be modified. This follows the standard procedure for introducing compensator terms in the Pauli–Lubanski vector in $d=4$. For example, $(W^{\mu_1\cdots\mu_{d-3}})^2$ is modified to be $(W^{\mu_1\cdots\mu_{d-3}}+\frac{1}{12}K^{\mu_1\cdots\mu_{d-3}1})^2$, where

$$K^{\mu_1\cdots\mu_{d-3}1}=K^{\mu_1\cdots\mu_{d-3}}-p^{-2}P^{[\mu_1}(P\cdot K)^{\mu_2\cdots\mu_{d-3}]} \tag{4.2}$$

and $K^{\mu_1\cdots\mu_{d-3}}=\bar{S}\Gamma_\rho\Gamma_\sigma\Gamma_\lambda S\epsilon^{\mu_1\cdots\mu_{d-3}\rho\sigma\lambda}$. An alternative form for this Casimir is $(W_{\rho\sigma}+K_{\rho\sigma})^2$, where

$$W_{\rho\sigma}=J_{[\rho\sigma}P_{\lambda]}P^\lambda\quad\text{and}\quad K_{\rho\sigma}=\tfrac{1}{12}\bar{S}\Gamma_{\rho\sigma\lambda}SP^\lambda. \tag{4.2'}$$

It is straightforward to modify the other Casimirs. Thus the notion of a single superspin developed for irreps of $d=4$ unextended SUSY[24] must now be extended to a set of $([d/2]+1)$ such quantities.

When we include the bosonic central charges $Z_{\mu\nu},Z_M$ without the spinorial central charge [cf. (3.1) with $u_k=0$, $\forall k$] we do not seem able to obtain a suitable generalization of the extended Pauli–Lubanski vectors $W^{(r)}$ defined above. Obviously, P^2, $Z_{\mu\nu}Z^{\mu\nu}$, and Z_MZ^M are Casimirs for this algebra.

As will be discussed later, the algebra (3.1) seems more suitable for the study of $N=1$, $d=11$ supergravity than the extended algebra (3.5). Therefore we will discuss the Casimirs of the first one. We find that the inclusion of the spinorial central charges S'_α changes things drastically. First, P^2 is no longer a Casimir, if $u_1\neq0$, and instead we obtain the Casimir

$$\mathscr{P}^2=P^2-\frac{u_1}{2!u_2}Z_2^2+\frac{u_1}{5!u_5}Z_5^2+2u_1C^{\alpha\beta}S_\alpha S'_\beta. \tag{4.3}$$

Second, since S' has nontrivial commutation relations only with $J_{\mu\nu}$, it is impossible, we think, that J can enter into any Casimir of this algebra. The argument for this is the following. If $JA+B$ is an operator that commutes with S' such that A does not contain any J, then A will commute with S'. Then $[JA,S']=-\frac{1}{2}\Gamma_{\mu\nu}S'A\cdots[B,S']=\frac{1}{2}\Gamma_{\mu\nu}S'A$. Since B does not commute with S' then B has to contain J. Therefore $B=JC$ for some C. Here, A does not contain J, so that C does not contain J, otherwise one gets $[B,S']=J\Gamma S'C_1\neq[JA,S']$. Thus C will commute with S' and hence $C=A$. The case A contains J can be treated similarly by setting $A=JA_1$, A_1 does not contain J, and so on. Thus for this algebra there appears to be no simple generalization of the superspin Casimir. We note that besides (4.3) another Casimir for this algebra is $\bar{S}'S'$.

Finally, this discussion does not guarantee to give the whole Casimirs of supersymmetry algebras. We may appeal to arguments from the contraction of OSP($n|1$) to our algebra that indeed we have obtained all the Casimirs for the SUSY algebra (3.1) in $d = 11$ when $Z = 0$. The situation is clearly unresolved in the alternative case.

V. THE N BARRIER FOR THE EXTENDED ALGEBRAS

The original proof[7,8] of the existence of a barrier at various values of N for different dimensions to the construction of auxiliary fields for N-extended super Yang–Mills and supergravities required knowledge of the dimensions of the fundamental irrep of the N-SUSY algebra in the various dimensions being considered. As N increases the dimensions of these irreps increase exponentially, so making it ever more difficult to redefine the component fields so as to be either the appropriate physical or possible auxiliary fields for the N-extended supergravity. We now attempt to answer the question as to whether or not the extended algebras we have constructed in Secs. II and III can help bypass the N barrier in either $d = 4$ (when the barrier is at $N = 3$) or in $d = 11$ (when the barrier is at $N = 0$).

Let us first consider the case $d = 4$. We may regard the anticommutation relations (2.4)–(2.6) as being those for the elements of a Clifford algebra, with quadratic form having values in the algebra $\mathscr{A}$ generated by the purely bosonic generators $\{Z^{ij},(Z^{ij})^*,Z^j_i,P_\mu\}$ with commutators being given by (2.24)–(2.27). If this algebra were Abelian, as it would be if $Z^j_i = 0$, and the quadratic form nondegenerate then the standard representation theory of Clifford algebras would show that there exists a unique irrep of dimension 2^{2N}. In the more general case $\mathscr{A}$ can be written as the semidirect product $A_0 \oplus_s A_1$, where $A_0 = \{Z^j_i,P_\mu\}$, $A_1 = \{Z^{ij},(Z^{ij})^*\}$, with $[A_i,A_j]_- \subset A_{i+j}(\text{mod } 2)$. We may then represent Z^{ij} and Z^j_i as functions of the variables z^i_j,z_{ij}

$$Z^{ij} = \frac{\partial}{\partial z_{ij}} + \frac{1}{2}\left(Z^*_{il}\frac{\partial}{\partial z^l_j} - Z^{*jk}\frac{\partial}{\partial z^k_i}\right), \quad Z^i_j = \frac{\partial}{\partial z^i_j}. \tag{5.1}$$

Again we expect $\mathscr{A}$ to have the features of a field, so that there are no divisors of zero, and there is a unique irrep of $\{S^i_{\alpha+},S_{\beta-j}\}$ of dimension 2^{2N}. Since the dimension 2^{2N} is the same as if there were no central charges, the $N = 3$ barrier still exists. The introduction of the spinorial central charges S' have not helped the situation. This argument is not, however, an exact proof though we feel it can be made so by suitable analysis.

For $d = 11$ a similar situation persists. Since the central charges Z_k defined in (3.1) are Abelian, then provided the quadratic form

$$Q^{(11)}_{\alpha\beta} = \sum_{k=1,2,5}\frac{1}{k!}Z_k(\Gamma^k C^{-1})_{\alpha\beta} \tag{5.2}$$

is nondegenerate the dimension of the unique irrep of the Clifford algebra generated by the set $\{S_\alpha\}$ will be 2^{16}, as it is without the presence of the extra spinorial central charges S'.

We note here that if we take representations of the total algebra in either $d = 4$ or 11 for which S' is nontrivial then ghost states will be present. This may be seen following the arguments used in analyzing all supersymmetries of the S matrix.[25] For expectation values in a given state of the anticommutators (2.17), (2.18) (for $d = 4$), or (3.3) with $Z'_k = 0$ (for $d = 11$) can only be zero if there are intermediate state contributions on the left-hand side (lhs) which have negative norm.

To avoid this unphysical feature of the representations, but yet not require the new spinorial central generators S' to vanish identically, we must require the S' to vanish on-shell. They are thus "off-shell" central charges. To achieve this off-shell character, and at the same time avoid the above N barrier, we must require the degeneracy of the quadratic forms discussed above. In the case of $d = 11$ this degeneracy is given simply by the requirement that

$$\det Q^{(11)}_{\alpha\beta} = 0, \tag{5.3}$$

whereas the corresponding matrix $Q^{(4)}_{\underline{\alpha}\underline{\beta}}$ (the index $\underline{\alpha}$ has both a chiral spinor index $\alpha\pm$ together with either an upper or a lower internal index) in the $d = 4$ case is, from Sec. II,

$$Q^{(4)} = \begin{pmatrix} Z^{ij}C^{-1}_{\alpha+\beta+} & (\not{p}C^{-1})_{\alpha+\beta-}\,\delta^i_j & -Z^i_jC^{-1}_{\alpha+\beta+} & 0 \\ (\not{p}C^{-1})_{\alpha-\beta+}\,\delta^j_i & (Z^{ij})^*C^{-1}_{\alpha-\beta-} & 0 & Z^j_iC^{-1}_{\alpha-\beta-} \\ Z^j_iC^{-1}_{\alpha+\beta+} & 0 & 0 & 0 \\ 0 & -Z^i_jC^{-1}_{\alpha-\beta-} & 0 & 0 \end{pmatrix}. \tag{5.4}$$

In either $d = 4$ or 11 the condition that $Q^{(d)}_{\underline{\alpha}\underline{\beta}}$ is singular is equivalent to the existence of a spinor $\chi^{(d)}_{\underline{\beta}}$ for which

$$Q^{(d)\underline{\beta}}_{\underline{\alpha}}\chi^{(d)}_{\underline{\beta}} = 0. \tag{5.5}$$

The only spinors available for $d = 11$ are S_α and S'_α. If we take $\chi^{(11)}_\alpha = S_\alpha$ in (5.3), then the commutators (3.1) set $S' = 0$. We thus take $\chi^{(n)}_\alpha = S'_\alpha$, and have consistency, from (3.1), provided

$$\text{tr}[Q^{(11)}]^2 \propto \sum\frac{1}{k!}(Z_kZ^k) = 0. \tag{5.6}$$

For the case $d = 4$, we have four possible choices for χ [these are dictated by the need for the matching of both spinorial and internal indices in the summation over $\underline{\beta}$ in (5.5), and also by the dimensions of the various entries in the matrix $Q^{(4)}_{\underline{\alpha}\underline{\beta}}$]. These are

$$(S'_+,S'_-,S_+,S_-), \quad (\not{p}S_-,\not{p}S_+,\not{p}^{-1}S'_-,\not{p}^{-1}S'_+),$$
$$(\not{p}S_-,\not{p}S_+,S_+,S_-), \quad (S'_+,S'_-,\not{p}^{-1},S'_-,\not{p}^{-1}S'_+).$$

For each of these choices the imposition of (5.5) together with the commutation relations of Sec. II seem to restrict

both $Z^i{}_j$ and S' to be zero, so that the algebra reduces to the original one without the extra central charges.

This reduction will still occur even if the three-index spinorial central charge, mentioned in Sec. II, is included, since this will also be set to zero by the spin reduction condition in a similar manner to S' discussed above.

We conclude that degeneracy may be achieved in $d = 4$ only by loss of the additional central charges, whilst in $d = 11$ it is possible for it to occur provided the massless constraint (5.4) is valid. In both cases the mass-shell constraint $p^2 = 0$ will set all central charge variables to zero.

VI. DIMENSION MATCHING

The problem of constructing a full superspace version of $N \geqslant 2, d = 4$ or $N = 1, d = 11$ supergravities is still unsolved. One of the first questions related to this for $N = 8$ is the set of extra fermionic and bosonic variables, beyond the set of $d\,x^\mu$'s ($d = 4$ or 11) and 32 θ^i_α's, over which integration is to be performed. We will consider first the case $d = 4$, where we showed in Sec. II that the maximal set of off-shell central charges is given by the set $\{Z^{ij}, Z^j_i, S'_{\alpha+r}, S''_{\alpha-}\}$. The associated variables lengths dimensions are $\{Z_{ij}(1),\ Z^i{}_j(2),\ \theta''_{\alpha+}(\tfrac{3}{2}),\ \theta'_{\alpha-r}(\tfrac{3}{2})\}$. The maximal measure for a full superspace integral will therefore be

$$d\mu_4 = d^4x\, d^{32}\theta\, d^{28}Z_{ij}\, d^{64}Z^i{}_j\, d^{32}\theta'. \tag{6.1}$$

Taking into account the complex (Hermitian) nature of $Z_{ij}(Z^i{}_j)$, respectively, the (length) dimension of (6.1) is found to be 60, which is very unsatisfactory. Nor can removal of integration over one or more of $Z_{ij}, Z^i{}_j$, or θ' make this value better. It would appear that only the reduced case of seven complex Z_{ij}'s is possible, with no $Z^i{}_j$ or θ' integration. However, there are measures which can be constructed from extra spinorial generators S' of length dimension $\frac{1}{2}$ (not $\frac{3}{2}$) and associated bosonic generators of dimension one which are satisfactory but had been excluded in Sec. II. We propose to discuss these elsewhere.

The case of $d = 11$ is more hopeful. If we take $Z_5 = 0$, the measure

$$d\mu_{11} = d^{11}x\, d^{32}\theta\, d^{55}z_{ij}\, d^{32}\theta' \tag{6.2}$$

has net dimension (for real Z_{ij}) of 2. Thus a possible candidate action for supergravity in $d = 11$ would be

$$\int d\mu_{11} \det E^A_M\, \mathscr{L}, \tag{6.3}$$

where $\mathscr{L}$ is a scalar with dimension L^{-2} and E^M_A is the usual multibein with A denoting the indices for the various coordinates in (6.2). Suitable torsion constraints would have to be determined in order that (6.3) has the correct field equations and also satisfies the degeneracy condition (5.3) and (5.4) of the previous section. The only nonzero dimension-zero torsions would be expected to be those of the flat algebra of Sec. III. However, a detailed analysis would appear necessary in order to construct such a theory, and again that is outside the scope of this analysis. In any case the simple and most natural case

$$\mathscr{L} = 1/\kappa_{11}^2 \tag{6.4}$$

is excluded by the above dimensional reasons.

VII. DISCUSSION

We have shown that it is possible to extend the maximal supersymmetry algebras in $d = 4$ or 11 by addition of a limited set of bosonic and fermionic central charges. This restricts the allowed set of superspace measures, all of which appear unsatisfactory if they are to give a simple action based on the volume of the associated superspace. In all the cases considered it is necessary to perform integration over a subspace of the full superspace obtained by reducing the number of bosonic variables. This has serious consequences for $d = 11$, as we shall now discuss.

We have already seen how the enlarged algebra in $d = 11$ has a Clifford algebra for the SUSY generators S_α. The dimension 2^{16} of the unique irrep of this Clifford algebra was that used in the original proof of the No-Go theorem in $d = 11$ (Ref. 7).

To avoid the resulting $N = 0$ barrier a spin-reducing constraint was required, arising from the degeneracy of the quadratic form of the Clifford algebra. It is necessary to interpret an action in the enlarged superspace with the spin-reducing constraint.

This question has been analyzed in $d = 4$ (Ref. 22), where the further requirement that the difference between the (mass) dimensions of the bosonic (d_B) and fermionic (d_F) central charges in the measure is $d_B - d_F = 2$. This insures that integration over all of these central charge variables produces only two bosonic central charge derivatives, as has been shown explicitly for linearized actions with $N = 2$ (Ref. 19) and $N = 4$ and 8 (Ref. 20). The detailed extension of these results to $d = 11$ should be straightforward, with the same criterion

$$d_B - d_F = 2. \tag{7.1}$$

In order to satisfy (7.1) we must break the full SO(1,10) invariance by keeping at most seven extra central charges if we wish also to have (6.3) and (6.4) in $d = 11$. Such breaking is clearly unappealing, though we expect the full invariance to be restored on-shell. However, if we wish to have maximal uv-divergence cancellation we must have maximal invariance. Thus the construction of a superspace version of $d = 11$ with maximal local SO(1,10) symmetry is not possible. A similar problem does not arise for $d = 4$, where the local SO(1,3) can be preserved off-shell in the construction of a full superspace form of $d = 4, N = 8$ SGR.[26] However, the use of the extended algebra discussed in Sec. II does not seem possible in that case, since we showed in Sec. V that the degeneracy constraint required to bypass the $N = 3$ barrier annihilates the extra central charges of Sec. II.

We conclude that there are serious difficulties in constructing a fully symmetric full-superspace version of $d = 11$ SGR. If such a construction is essential in showing the possible finiteness of the theory we see that such finiteness will therefore be impossible to determine. The dimensionally reduced version $d = 4$ of $N = 8$ SGR is a much more interesting situation; its finiteness is still allowed to be analyzed using full superspace techniques. Such a possibility may be an indication of potential finiteness as distinct from that for $d = 11, N = 1$ SGR. This may well be the reason why we live in four dimensions; all higher-dimensional maximal super-

gravities are ultraviolet infinite on-shell.

ACKNOWLEDGMENTS

Financial support by the Science and Engineering Research Council of Great Britain is gratefully acknowledged by B. L. Rands. E. Ahmed would like to acknowledge financial support by the Government of Egypt.

[1]Th. Kaluza, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Kl. 966 (1921); O. Klein, Z. Phys. **37**, 895 (1926).
[2]E. Witten, Nucl. Phys. B **186**, 412 (1981).
[3]L. Castellani, R. d'Auria, and P. Fré, Turin University preprint, 1983.
[4]S. Watamura, Phys. Lett. B **129**, 188 (1983).
[5]S. Mandelstam, in *Proceedings of the International High Energy Physics Conference, Paris 1982;* p. 331; J. Phys. **43**, Colloq. C-3, Supplement to No. 12, Nucl. Phys. B **213**, 419 (1983).
[6]P. Howe, K. Stelle, and P. Townsend, *Proceedings of the International High Energy Physics Conference, Paris 1982,* p. 326; Nucl. Phys. B **214**, 519 (1983).
[7]V. Rivelles and J. G. Taylor, Phys. Lett. B **104**, 131 (1981); Phys. Lett. B **121**, 37 (1983).
[8]J. G. Taylor, J. Phys. A. Math. Gen. **15**, 867 (1982).
[9]T. Hori and K. Kamimura, Phys. Lett. B **132**, 65 (1983).
[10]J. G. Taylor, "Extended superspace supergravity in the light-cone gauge," *Proceedings of the Karpacz School on SUSY and Supergravity*, 1983, edited by B. Milewski (World Scientific, Singapore, 1983).
[11]G. Bufton, Ph.D. thesis, University of London, 1983 (unpublished).
[12]G. Bufton and J. G. Taylor (in preparation).
[13]M. Grisaru and W. Siegel, Nucl. Phys. B **201**, 292 (1982).
[14]M. Sohnius, Nucl. Phys. B **138**, 109 (1978).
[15]S. Ferrara and C. Savoy, in *Supergravity '81*, edited by S. Ferrara and J. G. Taylor (Cambridge U. P., Cambridge, 1982).
[16]J. G. Taylor, in *Superspace, Supergravity*, edited by S. Hawking, and M. Rocek, (Cambridge U. P., Cambridge, 1981), p. 363 *et seq.*
[17]B. Rands and J. G. Taylor, J. Phys. A. Math. Gen. **16**, 1005 (1983).
[18]Y. Hassoun, A. Restuccia, and J. G. Taylor, An unconstrained superspace form of $N = 4$ super Yang–Mills, King's College, London preprint, 1983.
[19]Y. Hassoun, A. Restuccia, and J. G. Taylor, J. Math. Phys. **26**, 345 (1985).
[20]G. Amerighi, Y. Hassoun, L. Hørnfeldt, A. Restuccia, and J. G. Taylor, "Superfield actions for $N = 4$ and 8 degenerate central charges," King's College, London preprint, 1983.
[21]R. d'Auria and P. Fré, Nucl. Phys. B **201**, 101 (1982).
[22]A. Restuccia and J. G. Taylor, J. Phys. A. Math. Gen. **16**, 4097 (1983).
[23]J. Van Holten and A. Van Proyen, J. Phys. A. Math. Gen. **15**, 3763 (1982).
[24]E. Sokatchev, Nucl. Phys. B **99**, 96 (1975).
[25]R. Haaag, M. Sohnius, and J. Lopuszanski, Nucl. Phys. B **88**, 257 (1975).
[26]P. Davis, A. Restuccia, and J. G. Taylor, Phys. Lett. B **144**, 46 (1984).

Quantum theory of particles of indefinite mass: Spin-½

Levere Hostler
Physics Department, Wilkes College, Wilkes Barre, Pennsylvania 18766

(Received 30 January 1985; accepted for publication 26 April 1985)

The quantum theory of particles of indefinite mass investigated earlier by Hostler [J. Math. Phys. **21**, 2461 (1980); **22**, 2307 (1981)] is adapted to apply to spin-½ particles. The proposed wave equation is $\{\frac{1}{2}\Pi\cdot(1+i\sigma)\cdot\Pi+(1/i)(\partial/\partial\lambda)\}\phi=0$, $\Pi_\mu=p_\mu-eA_\mu$, in which ϕ is a 2×1 Pauli spinor and σ is a spin tensor whose Lorentz components are 2×2 Pauli spin matrices. Particle scattering by an external c-number static field is computed within the framework of the new formalism, and agreement with the known physics of a charged spin-½ particle is obtained for this special case.

I. INTRODUCTION

The "second-order Dirac equation"[1]

$$\{\Pi\cdot(1+i\sigma)\cdot\Pi+m^2\}\phi=0, \tag{1.1}$$

$\Pi=p-eA$, has been investigated earlier.[2,3] In Eq. (1.1) ϕ is a 2×1 Pauli spinor, and σ is a self-dual spin tensor, whose Lorentz components are

$$\sigma_{\mu\nu}=\begin{array}{c}\\ \mu\downarrow\end{array}\begin{array}{c}\nu\longrightarrow\\ \left(\begin{array}{c|c|c|c} 0 & \sigma_3 & -\sigma_2 & \sigma_1 \\ \hline -\sigma_3 & 0 & \sigma_1 & \sigma_2 \\ \hline \sigma_2 & -\sigma_1 & 0 & \sigma_3 \\ \hline -\sigma_1 & -\sigma_2 & -\sigma_3 & 0 \end{array}\right)\end{array}, \tag{1.2}$$

where $\sigma_{1,2,3}$ are the ordinary 2×2 Pauli spin matrices. It has been shown[2] that Eq. (1.1), equipped with the inner product

$$(\phi_b;\phi_a)=\frac{1}{i}\int d^3r\,\bar{\phi}_b\,\overrightarrow{\Pi}_4\,\phi_a, \tag{1.3}$$

provides an equivalent alternative description of a Dirac particle. Indeed, a full quantum electrodynamics (QED) based on Eq. (1.1) has been investigated,[4] and has been proven entirely equivalent to conventional QED. The dual vector $\bar{\phi}$ in Eq. (1.3) is defined through the equation

$$\bar{\phi}=\phi^\dagger(-\overleftarrow{\Pi}_4-\boldsymbol{\sigma}\cdot\overleftarrow{\boldsymbol{\Pi}}), \tag{1.4}$$

and plays a role in the second-order Dirac equation, which corresponds to the role of $\bar{\psi}=\psi^\dagger\gamma_4$ in the conventional Dirac equation (herein called the "first-order Dirac equation").

Our present interest in Eq. (1.1) stems from the fact that it contains the mass only quadratically. This feature enables us to convert Eq. (1.1) into a Fock-type wave equation. The resulting wave equation

$$\left\{\frac{1}{2}\Pi\cdot(1+i\sigma)\cdot\Pi+\frac{1}{i}\frac{\partial}{\partial\lambda}\right\}\phi=0 \tag{1.5}$$

differs from the original Fock equation[5] primarily in that it involves a 2×1 Pauli spinor, rather than a 4×1 Dirac spinor. Accordingly, we avoid the doubling of the number of states associated with the original Fock equation; according to which there are two wave functions, rather than only one, belonging to each physical particle state.

We here propose to adopt Eq. (1.5) as the wave equation for a charged spin-½ particle, rather than the second-order Dirac equation from which we started. A wave packet in our theory will be a superposition of states of different mass. The author has investigated such "indefinite mass quantum theories" earlier[6,7] in the spin-0 and spin-1 cases. The present work merges the two earlier lines of research involving the author's indefinite mass quantum theory and the theory of the second-order Dirac equation.

Analogous theories have been proposed by a number of authors including Stueckelberg,[8] Davidon,[9] Horwitz and coworkers,[10] and Fanchi and Collins.[11] These theories are characterized by the use of Fock-type wave equations involving a new evolution parameter corresponding to our λ in Eq. (1.5). The new evolution parameters all are related to proper time. The "mass Hamiltonian" methods of Enatsu and co-workers[12] and of Feynman, Kislinger, and Ravndal[13] are quantum theories of this same general type as well. An alternative treatment of spin-½ using Fock-type methods has been given earlier by Piron, Reuse, and Grelland.[14] Also, the concept of mass wave packets has been explored before by Greenberger.[15]

We will here continue the program of Refs. 6 and 7 but using the Fock equation (1.5) instead of the scalar form described earlier. In Sec. II a five-current conservation law for Eq. (1.5) is derived. It follows from this conservation law that the not-positive-definite inner product

$$(\phi_b;\phi_a)=\int d^4x\,\bar{\phi}_b\,\phi_a \tag{1.6}$$

is independent of "evolution time," λ. In analogy to the scalar case; we shall give $(\phi_b;\phi_a)$ the physical interpretation of the probability amplitude, and $|(\phi_b;\phi_a)|^2$ the probability, to see the state ϕ_b if a suitable measurement is made on the system when its state is originally ϕ_a. In Sec. III we apply this rule to calculate the scattering cross section for scattering of a charged spin-½ particle by a static (in x^0 and in λ) external potential A_μ. We find agreement with the known physics of a Dirac particle in this case. Indeed, our spin-½ wave equation (1.5) has been carefully designed to preserve this known body of knowledge in a suitable limiting case, in this example the static case.

II. FIVE-CURRENT CONSERVATION LAW

Our calculation is greatly simplified by recognizing that the operator $-\Pi\cdot(1+i\sigma)\cdot\Pi$ factorizes in the form

 0022-2488/85/102666-04$02.50

$-\Pi\cdot(1+i\sigma)\cdot\Pi = A\,B$, where

$$A = A^\dagger = -i\Pi_4 + \boldsymbol{\sigma}\cdot\boldsymbol{\Pi} \tag{2.1}$$

and

$$B = B^\dagger = -i\Pi_4 - \boldsymbol{\sigma}\cdot\boldsymbol{\Pi}. \tag{2.2}$$

Then the new Fock equation (1.5) can be written

$$\left\{\frac{1}{2}A\,B + \frac{1}{i}\frac{\partial}{\partial\lambda}\right\}\phi = 0. \tag{2.3}$$

Taking the adjoint and then acting with $\overleftarrow{B}$ on the right gives the differential equation of the dual state in the form

$$\bar{\phi}\left\{\frac{1}{2}\overleftarrow{A}\,\overleftarrow{B} + \frac{1}{i}\frac{\overleftarrow{\partial}}{\partial\lambda}\right\} = 0,$$

equivalently

$$\bar{\phi}\left\{\frac{1}{2}\overleftarrow{\Pi}\cdot(1+i\sigma)\cdot\overleftarrow{\Pi} + \frac{1}{i}\frac{\overleftarrow{\partial}}{\partial\lambda}\right\} = 0. \tag{2.4}$$

This calculation assumes the four-potential A_μ to be independent of evolution time λ.

Now let ϕ_a obey Eq. (1.5) and $\bar{\phi}_b$ Eq. (2.4). By multiplying Eq. (1.5) on the left by $\bar{\phi}_b$, and Eq. (2.4) on the right by ϕ_a, and subtracting, we can proceed in a standard way to obtain a conservation law. This conservation law is

$$\frac{\partial\rho^{ba}}{\partial\lambda} + \frac{\partial j_\mu^{ba}}{\partial x_\mu} = 0, \tag{2.5}$$

where

$$\rho^{ba} = \bar{\phi}_b\,\phi_a \tag{2.6}$$

and

$$j^{ba} = \tfrac{1}{2}\bar{\phi}_b(\overleftarrow{\Pi}\cdot(1+i\sigma) + (1+i\sigma)\cdot\overrightarrow{\Pi})\,\phi_a. \tag{2.7}$$

It is characteristic of Fock-type wave equations that the four-current (2.7) is not a conserved four-current. On the other hand, a conserved four-current does emerge after integrating Eq. (2.5) over all evolution time:

$$\frac{\partial J_\mu^{ba}}{\partial x_\mu} = 0, \tag{2.8}$$

$$J^{ba} = \int d\lambda\, j^{ba}. \tag{2.9}$$

It is this J^{ba} that is the conserved transition current that we are accustomed to. This point will be critical for the cross-section calculation taken up in Sec. III.

Instead of integrating over evolution time, we may integrate Eq. (2.5) over all Minkowski space. Dropping surface terms gives the identity

$$\frac{\partial}{\partial\lambda}\int d^4x\,\bar{\phi}_b\,\phi_a = 0.$$

This justifies the statement made in the Introduction that the dot product (1.6) is independent of the evolution time at which it is calculated. Now let ϕ be normalized at one evolution time in such a way that the probability to see the state ϕ, if a suitable measurement is made on the system when it is already known to be in the state ϕ, is $|\int d^4x\,\bar{\phi}\phi\,|^2 = 1$. Then ϕ will remain so normalized for all evolution time, thereby justifying our choice of inner product (1.6) for the new Fock equation.

III. POTENTIAL SCATTERING

We will here apply Eq. (1.5) and its inner product (1.6) in a situation for which known physics is expected to prevail—scattering by a four-potential independent of λ. For simplicity we assume the four-potential independent of x^0 as well. The standard lore of scattering theory will be assumed as needed, but will be rephrased in terms of the evolution parameter λ instead of the usual x^0. The calculation follows closely a similar calculation for spin-0 carried out in Ref. 7, Appendix B. The initial state at $\lambda = -\infty$ is taken to be

$$\phi_i(x,\lambda) = (V\,Tm)^{-1/2}\,u^{\rho_i}(p_i)\,e^{i(p_i\cdot x + (\lambda/2)(m_i)^2)}, \tag{3.1}$$

with dual state

$$\bar{\phi}_i(x,\lambda) = (V\,Tm)^{-1/2}\,\bar{u}^{\rho_i}(p_i)\,e^{-i(p_i\cdot x + (\lambda/2)(m_i)^2)}; \tag{3.2}$$

in which the plane wave spinors $u^\rho(p)$, $\bar{u}^\rho(p)$ are defined by

$$u^\rho(p) = [(E+m+\boldsymbol{\sigma}\cdot\mathbf{p})/(2m(E+m))^{1/2}]\,u^\rho(0), \tag{3.3}$$

and

$$\bar{u}^\rho(p) = u^\rho(0)^\dagger m(E+m-\boldsymbol{\sigma}\cdot\mathbf{p})/(2m(E+m))^{1/2},$$
$$u^1(0) = \begin{pmatrix}1\\0\end{pmatrix},\quad u^2(0) = \begin{pmatrix}0\\1\end{pmatrix}. \tag{3.4}$$

The state $\phi_i(x,\lambda)$ is normalized to unity in a space-time box with three-volume V and duration $T = \int dx^0$ in "observers time," x^0. The plane wave spinor $u^\rho(p)$ corresponds to a particle which has spin up ($\rho = 1$) or spin down ($\rho = 2$) in the comoving frame, and is obtained by Lorentz transforming from the comoving frame back to the frame in which the three-momentum is $\mathbf{p}_i$.

First we convert Eq. (1.5) into an equivalent integral equation that incorporates the boundary condition at $\lambda = -\infty$:

$$\phi_\epsilon^+(x_2,\lambda_2) = \phi_i(x_2,\lambda_2) + \frac{1}{2}\int d\lambda_1\int d^4x_1 \times G_\epsilon^0(x_2,\lambda_2;x_1,\lambda_1)\,V(x_1)\,\phi_\epsilon^+(x_1,\lambda_1). \tag{3.5}$$

In Eq. (3.5) strictly retarded boundary conditions in evolution time are assumed: $G_\epsilon^0(x_2,\lambda_2;x_1,\lambda_1)$ is the retarded Green's function of the free Fock equation (1.5), with Fourier representation

$$G_\epsilon^0(x_2,\lambda_2;x_1,\lambda_1) = \int\frac{dN}{4\pi}\int\frac{d^4p}{(2\pi)^4} \times\exp[ip\cdot(x_2-x_1) + (i/2)N(\lambda_2-\lambda_1)]\,G_\epsilon^0(p,N), \tag{3.6}$$

in which

$$G_\epsilon^0(p,N) = 2/(p^2+N-i\epsilon),\quad \epsilon>0. \tag{3.7}$$

As has been noted before,[7,16] for Fock-type wave equations retarded boundary conditions in evolution time correspond to Feynman propagation of the timelike states, characterized by $N>0$: for the timelike states, positive frequencies propagate forward in observer's time, x^0, and negative frequencies propagate backward. This is evidenced by the $-i\epsilon$ in the denominator of Eq. (3.7).

By expanding in a Born series, the integral equation (3.5) can be formally solved for $\phi_\epsilon^+(x_2,\lambda_2)$:

$$\phi_\epsilon^+(x_2,\lambda_2) = \int d\lambda_1 \int d^4x_1\, G_\epsilon^A(x_2,\lambda_2;x_1,\lambda_1) \times\left[\frac{1}{2}(\overleftarrow{\partial}_1)^2 + \frac{1}{i}\frac{\overleftarrow{\partial}}{\partial\lambda_1} - i\epsilon\right]\cdot\phi_i(x_1,\lambda_1), \tag{3.8}$$

in which $G_\epsilon^A(x_2,\lambda_2;x_1,\lambda_1)$ is the interacting Green's function of Eq. (1.5), obeying the differential equation

$$\left[\frac{1}{2}\Pi_2\cdot(1+i\sigma)\cdot\Pi_2 + \frac{1}{i}\frac{\partial}{\partial\lambda_2} - i\epsilon\right] G_\epsilon^A(x_2,\lambda_2;x_1,\lambda_1) = \delta(\lambda_2-\lambda_1)\delta^4(x_2-x_1). \tag{3.9}$$

Now we employ the dot product (1.6) to calculate the probability amplitude to see the particle in a state $\phi_f(x_2,\lambda_2)$ in the remote future in evolution time, $\lambda_2\to\infty$. This probability amplitude is

$$S_{fi} = \lim_{\lambda_2\to\infty}\int d^4x_2\, \bar{\phi}_f(x_2,\lambda_2)\,\phi_i^+(x_2,\lambda_2),$$

where $\phi_i^+(x_2,\lambda_2)$ is the interacting wave function (3.8). By use of standard reduction techniques, and assuming $f\neq i$, this can be rewritten in the form

$$S_{fi} = i\int d^4x_2\int d\lambda_2\int d^4x_1\int d\lambda_1\,\bar{\phi}_f(x_2,\lambda_2) \times\left[\frac{1}{2}(\vec{\partial}_2)^2 + \frac{1}{i}\frac{\vec{\partial}}{\partial\lambda_2} - i\epsilon\right]\cdot G_\epsilon^A(x_2,\lambda_2;x_1,\lambda_1) \times\left[\frac{1}{2}(\overleftarrow{\partial}_1)^2 + \frac{1}{i}\frac{\overleftarrow{\partial}}{\partial\lambda_1} - i\epsilon\right]\phi_i(x_1,\lambda_1). \tag{3.10}$$

If A_μ is independent of evolution time λ; then the propagator $G_\epsilon^A(x_2,\lambda_2;x_1,\lambda_1)$ has the form

$$G_\epsilon^A(x_2,\lambda_2;x_1,\lambda_1) = \int\frac{dN}{4\pi}\,e^{(1/2)iN(\lambda_2-\lambda_1)}\,G_\epsilon^A(x_2,x_1,N), \tag{3.11}$$

in which $G_\epsilon^A(x_2,x_1,N)$ is proportional to the familiar interacting Green's function of the second-order Dirac equation, Eq. (1.1), with "mass squared" $= N$:

$$\tfrac{1}{2}\left[\Pi_2\cdot(1+i\sigma)\cdot\Pi_2 + N - 2i\epsilon\right] G_\epsilon^A(x_2,x_1,N) = \delta^4(x_2-x_1). \tag{3.12}$$

When the representation (3.11) is substituted in Eq. (3.10) and explicit expressions are substituted for the initial and final plane wave states; it becomes possible to perform the λ_1 and λ_2 integrals explicitly, with the result

$$S_{fi} = [4\pi\delta((m_f)^2-(m_i)^2)/VT]\,M_{fi}, \tag{3.13}$$

$$M_{fi} = \frac{i}{m_i}\int d^4x_2\int d^4x_1\,\bar{u}^{\rho_f}(p_f)\,e^{-ip_f\cdot x_2} \times\frac{1}{2}\left[(\vec{\partial}_2)^2 + (m_i)^2\right] G_\epsilon^A(x_2,x_1,(m_i)^2) \times\frac{1}{2}\left[(\overleftarrow{\partial}_1)^2 + (m_i)^2\right] u^{\rho_i}(p_i)\,e^{ip_fx_1}. \tag{3.14}$$

Next we use the relation

$$4\pi\delta(N=0) = \int d\lambda\, e^{-i(1/2)N\lambda}\,|_{N=0} = \Lambda,$$

where Λ is the duration of the scattering in evolution time, to obtain the probability $W_{fi} = |S_{fi}|^2$ to see the state $\phi_f(x_2,\lambda_2)$ as $\lambda_2\to\infty$ in the form $W_{fi} = \Lambda\,4\pi\delta((m_f)^2-(m_i)^2)\times|M_{fi}|^2/(V^2T^2)$. This gives a probability per unit of observer's time, T, of seeing the final state $\phi_f(x_2,\lambda_2)$ in the form

$$W_{fi}/T = (\Lambda/V^2T^3)4\pi\delta((m_f)^2-(m_i)^2)|M_{fi}|^2. \tag{3.15}$$

We divide this probability per unit of observer's time by the total incident current, in order to obtain a differential cross section. Note that by the total current here we mean the empirical conserved four-current $J_\mu = \int d\lambda\, j_\mu$, where j_μ is given by Eq. (2.7) with $\phi_b = \phi_a = \phi_i$. The magnitude of the space part of J_μ for the state (3.1) is

$$J_{\rm inc} = (\Lambda/VT)\,p_i. \tag{3.16}$$

The duration Λ in evolution time cancels out when the cross section $d\sigma' = (W_{fi}/T)/J_{\rm inc}$ is formed, and we find

$$d\sigma' = 4\pi\delta((m_f)^2-(m_i)^2)(|M_{fi}|^2/VT^2p_i). \tag{3.17}$$

The cross section $d\sigma$ for scattering into a final state with three-momentum in the range d^3p_f irrespective of the final mass squared is obtained by multiplication of Eq. (3.17) on both sides by $d^4n_f = VT\,d(m_f)^2\,d^3p_f/((2\pi)^32E_f) =$ the number of final states for which the final mass squared and final linear momentum lie within specified limits, followed by integration over all $(m_f)^2$. Actually, the mass squared conserving delta function in Eq. (3.17) has the effect of eliminating all contributions to the $(m_f)^2$ integration except contributions for which $(m_f)^2$ is in an infinitesimal neighborhood of $(m_i)^2$. We have here a guarantee that the timelike initial state $\phi_i(x_1,\lambda_1)$ shall scatter into a timelike final state of the same mass squared. For $d\sigma = \int d^4n_f\,d\sigma'$, we get

$$d\sigma = \frac{d^3p_f}{(2\pi)^3E_f}\frac{|M_{fi}|^2}{p_iT}. \tag{3.18}$$

Note that all reference to the evolution time Λ has here disappeared: the right-hand side of Eq. (3.18) involves the familiar space-time degrees of freedom only. The expression (3.18) is precisely the expression for $d\sigma$ that one would obtain in a standard treatment of the second-order Dirac equation, Eq. (1.1), with inner product (1.3), and is in agreement with the known physics of a charged Dirac particle.

ACKNOWLEDGMENTS

Part of this work was completed during my sabbatical leave at Yale University. I am grateful to the Physics Department at Yale University for their hospitality.

This research was supported in part by the National Science Foundation, Grant No. PHY-8403322.

[1] If Lorentz indices are exhibited explicitly, Eq. (1.1) reads $\{\Pi_\mu(\delta_{\mu\nu}+i\sigma_{\mu\nu})\times\Pi_\nu + m^2\}\,\phi = 0$. Our four-vectors have an imaginary fourth component, and our Lorentz metric is simply $\delta_{\mu\nu}$. Also, a derivative acting to the left signifies minus differentiation of the object on the left, as in $\bar{\phi}\overleftarrow{\partial}_\mu = -\partial\bar{\phi}/\partial x_\mu$.

[2] L. C. Hostler, J. Math. Phys. **23**, 1179 (1982).
[3] L. C. Hostler, J. Math. Phys. **24**, 2366 (1983).
[4] L. C. Hostler, J. Math. Phys. **26**, 1348 (1985).
[5] V. A. Fock, Phys. Z. Sowjetunion **12**, 404 (1937).
[6] L. C. Hostler, J. Math. Phys. **21**, 2461 (1980).
[7] L. C. Hostler, J. Math. Phys. **22**, 2307 (1981).
[8] E. C. G. Stueckelberg, Helv. Phys. Acta **14**, 322 (1941); **14**, 588 (1941); **15**, 23 (1942).

[9]W. C. Davidon, Phys. Rev. **97**, 1131 (1955); **97**, 1139 (1955).
[10]L. P. Horwitz and C. Piron, Helv. Phys. Acta **46**, 316 (1973); L. P. Horwitz and F. C. Rotbart, Phys. Rev. D **24**, 2127 (1981); L. P. Horwitz and Y. Lavie, Phys. Rev. D **26**, 819 (1982).
[11]J. R. Fanchi and R. E. Collins, Found. Phys. **8**, 851 (1978); J. R. Fanchi, Phys. Rev. D **20**, 3108 (1979).
[12]H. Enatsu, Prog. Theor. Phys. **30**, 236 (1963); H. Enatsu and S. Kawaguchi, Nuovo Cimento A **27**, 458 (1975); H. Enatsu, A. Takenaka, and M. Okazaki, Nuovo Cimento A **43**, 575 (1978).
[13]R. P. Feynman, M. Kislinger, and F. Ravndal, Phys. Rev. D **3**, 2706 (1971).
[14]C. Piron and F. Reuse, Helv. Phys. Acta **51**, 146 (1978); F. Reuse, Helv. Phys. Acta **51**, 157 (1978); H. H. Grelland, Int. J. Quantum Chem. **XIX**, 873 (1981).
[15]D. M. Greenberger, J. Math. Phys. **15**, 395, 406 (1974).
[16]Y. Nambu, Prog. Theor. Phys. (Kyoto) **5**, 82 (1950).

Analytic strong-potential Born approximation cross sections for electron capture from the K shell by fast ions

N. C. Deb and N. C. Sil
Department of Theoretical Physics, Indian Association for the Cultivation of Science, Jadavpur, Calcutta 700032, India

J. H. McGuire
Department of Physics, Kansas State University, Manhattan, Kansas 66506

(Received 18 March 1985; accepted for publication 23 May 1985)

The strong potential Born approximation has been applied to evaluate the transition amplitude for the collision process in which a fast ion of charge Z_P captures an electron from the K shell of an atomic target with nuclear charge Z_T, where $Z_T \geqslant Z_P$. The transition amplitude has been evaluated analytically in terms of Gauss hypergeometric functions. No peaking approximation has been made. We have considered the K-shell capture of electrons by high-energy protons from Li, C, Ne, Na, and Ar. Our results for C, Ne, and Ar are in good agreement with the available experimental data.

I. INTRODUCTION

In the past several years a variety of methods[1-6] have been developed to deal with the relatively difficult problem of electron capture. The development of the strong potential Born (SPB) approximation, among them, has led to significant progress in understanding the nature of electron capture. In this approximation the weaker of the two interactions—electron-target nucleus and electron-projectile nucleus interactions—is taken to first order and the stronger interaction is kept to all orders via Coulomb Green's function, whereas the nuclear–nuclear interaction is completely neglected.

Following the work of Briggs,[1] Macek and coworkers[2-4] first systematically derived the SPB amplitude as a three-dimensional integral using an off-energy-shell Coulomb wave function. Employing a further peaking approximation they[4] also reduced their three-dimensional integral into a closed form. In this peaking approximation the main assumption was to neglect angular dependence in the integrand and consequently their calculations were restricted to electron capture in the final s states for systems asymmetric in the projectile and target charges, Z_P and Z_T.

In a later development Alston[5] employed a less restrictive transverse peaking approximation to reduce the $1s$–$1s$ capture amplitude to a one-dimensional integral. In this approximation the author considered only the component of the intermediate momentum parallel to $\mathbf{v}$, the relative impact velocity, while the perpendicular component was neglected. Dube[6] has proposed explicit expressions for capture of an electron to and from arbitrary hydrogenic states for various multiple scattering approaches with peaking approximations.

Sil and McGuire[7] reported a treatment for the evaluation of the SPB amplitude in principle accurate to the error of the SPB approximation itself. They treated the term responsible for the Thomas peak exactly and expanded other terms in powers of Z_P/v, of which terms through first order were considered. They obtained the amplitude as a linear combination of one-dimensional integrals, each of which can be derived from a basic integral by suitable parametric differentiation. This method is applicable to electron capture into final higher angular momentum states as well. They also discussed briefly the possibility of evaluating this basic integral analytically in terms of Gauss hypergeometric functions. In another work McGuire and Sil[8] computed some $1s$–$1s$ differential capture cross sections by evaluting the one-dimensional integrals numerically.

In this work we calculate in the framework of the SPB approximation the $1s$–$1s$ capture cross sections at high incident energies for some asymmetric systems where the projectile charge is smaller than the target charge, viz., protons on Li, C, Ne, Na, and Ar. We present the detailed analytic evaluation of a basic integral in terms of a single Gauss hypergeometric function from which the SPB amplitude can be derived easily.

II. EVALUATION OF THE SPB AMPLITUDE

Using an integral representation[9] for the off-shell Coulomb wave function, Macek and Alston derived the expression[10] for the capture amplitude from the ground state to a final state ϕ_f (hydrogenic), for the case $Z_T \geqslant Z_P$ as (in their notation)

$$T_{\mathrm{SPB}} = -4\pi Z_P \int d\mathbf{p}\, \tilde{\phi}_f^*(\mathbf{p}) |\mathbf{p}-\mathbf{K}|^{-2} \left(\frac{4\pi^2 N_i}{(2\pi)^{3/2}} \tau \frac{e^{-i\pi\tau}}{\sin \pi\tau}\right) \frac{\partial}{\partial \mu} \left[(\mu^2+J^2)^{-1} \times \left\{\frac{[(\mu-iv)^2+K^2+2\mathbf{p}\cdot(\mathbf{J}-i\mu\hat{v})](p^2+Z_P^2/n^2)}{4(\mu^2+J^2)(v^2+2\mathbf{v}\cdot\mathbf{p})}\right\}^{-\tau}\right], \tag{1}$$

 0022-2488/85/102670-06$02.50

which is correct to the order $(Z_P/v)^2$. Sil and McGuire[7] gave an elaborate treatment for the reduction of the SPB amplitude to a linear combination of one-dimensional integrals by expanding the terms $|\mathbf{p}-\mathbf{K}|^{-2}$, τ, and $(v^2+2\mathbf{v}\cdot\mathbf{p})^{-\tau}$ in powers of $\mathbf{p}\cdot\mathbf{K}/K^2$ and $\mathbf{p}\cdot\mathbf{v}/v^2$, which is equivalent to an expansion in Z_P/v since $K\geqslant v/2$ and p may be taken to be of the order of Z_P. Results[8] for the $1s$–$1s$ is differential capture cross sections at several energies were obtained by them for the p–H and p–He systems by evaluating the terms through first order in Z_P/v. The zeroth-order term in their expansion is given by

$$T^{(0)}_{\mathrm{SPB}}=\frac{ch_0}{K^2}\frac{\partial}{\partial\mu}\left[\kappa^{2i\nu-2}\int d\mathbf{p}\,\tilde{\phi}^{*}_{f}(\mathbf{p})(TS)^{-i\nu}\right], \tag{2}$$

where

$$c=16\left(\frac{Z_T}{2}\right)^{3/2}Z_P,\quad h_0=\frac{2\nu}{1-e^{2\pi\nu}},\quad \nu=\frac{Z_T}{v},$$
$$\kappa^2=\mu^2+J^2,\quad T=(\mu-iv)^2+K^2+2\mathbf{p}\cdot(\mathbf{J}-i\mu\hat{v}), \tag{3}$$
$$S=(p^2+\lambda^2)/4v^2,\quad \lambda=Z_P/n,$$

and $\tilde{\phi}_f(\mathbf{p})$ is the Fourier transform of the final bound state wave function:

$$\tilde{\phi}_f(\mathbf{p})=(2\pi)^{-3/2}\int e^{i\mathbf{p}\cdot\mathbf{r}}\phi_f(\mathbf{r})d\mathbf{r}. \tag{4}$$

We consider the type integral

$$I_{MN}=\int d\mathbf{p}(p^2+\lambda^2)^{-M-i\nu}T^{-N-i\nu}, \tag{5a}$$

where M and N are integers such that $M\geqslant 2$ and $N\geqslant 0$.

We note that for an arbitrary $\phi_f(\mathbf{r})$ all the integrals in Eq. (2) can be derived from I_{MN}.

Separating the angular dependent part of the integral I_{MN} in Eq. (5a) we write

$$I_{MN}=\int_0^\infty p^2(p^2+\lambda^2)^{-M-i\nu}f(p)dp, \tag{5b}$$

where

$$f(p)=\int(A+2\mathbf{p}\cdot\mathbf{L})^{\tau}\,d\Omega \tag{6a}$$

with

$$A=(\mu-iv)^2+K^2,\quad \mathbf{L}=\mathbf{J}-i\mu\hat{v}, \tag{6b}$$

and

$$\tau=-N-i\nu.$$

For a real vector $\mathbf{L}$ we could choose the Z axis along $\mathbf{L}$ and carry out the θ,ϕ integrations to get

$$f(p)=[2\pi/(\tau+1)2Lp]\,[(A+2Lp)^{\tau+1}-(A-2Lp)^{\tau+1}]. \tag{7}$$

For a complex vector $\mathbf{L}$ as in our case the validity of the result in Eq. (7) has been justified (see the Appendix).

Substituting the value of the angular integral $f(p)$ from Eq. (7) in Eq. (5b) and putting $p=\lambda x$ we can write

$$I_{MN}=\frac{D_{MN}}{2L}\frac{\partial}{\partial L}\{(2L)^{\tau_2}H\}, \tag{8}$$

where

$$H=\int_{-\infty}^{\infty}(x^2+1)^{\tau_1}(x+\alpha)^{\tau_2}\,dx, \tag{9a}$$
$$\tau_1=-M-i\nu,\quad \tau_2=\tau+2,$$
$$\alpha=A/2L\lambda,\quad L=(\mathbf{L}\cdot\mathbf{L})^{1/2},$$

and

$$D_{MN}=\frac{\pi(Z_P)^{-2M-N+3-3i\nu}}{(-i\nu-N+1)(-i\nu-N+2)}. \tag{9b}$$

On writing $(x+\alpha)=(\alpha-i)\omega$, the basic integral H can be recast as

$$H=[-(\alpha+i)]^{\tau_1}[-(\alpha-i)]^{\tau_1}(\alpha+i)^{\tau_2+1}$$
$$\times\int_\Gamma\omega^{\tau_2}(1-\omega)^{\tau_1}(1-\omega z)^{\tau_1}\,d\omega, \tag{10}$$

where $z=(\alpha+i)/(\alpha-i)$ and Γ is a suitable contour in the complex ω plane.

The contour integral in Eq. (10) is a solution of the hypergeometric equation

$$\left\{z(1-z)\frac{d^2}{dz^2}+[c-(a+b+1)z]\frac{d}{dz}-ab\right\}$$
$$\times u(a,b;c;z)=0, \tag{11}$$

with $a=-\tau_1$, $b=\tau_2+1$, and $c=\tau_1+\tau_2+2$, since the function $\omega^b(1-\omega)^{c-b}(1-\omega z)^{-a-1}$ vanishes at the end points of the contour Γ (see Ref. 11). So H can be expressed as a linear combination of an independent pair of two Gauss hypergeometric functions satisfying Eq. (11). Thus we may write

$$H=(\alpha+i)^{\tau_1+\tau_2+1}(\alpha-i)^{\tau_1}$$
$$\times[E\,{}_2F_1(-\tau_1,\tau_2+1;\tau_1+\tau_2+2;z)+Dz^{-\tau_1-\tau_2-1}$$
$$\times{}_2F_1(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)]. \tag{12}$$

It is to be noted that the physical values of α always lie in the lower half of the complex α plane where the function $(\alpha+i)^{\tau_1+\tau_2+1}$ has a branch point at $\alpha=-i$, but the function $(\alpha-i)^q$, q being any complex number, is single valued and analytic. So we rewrite

$$H=(\alpha-i)^{2\tau_1+\tau_2+1}$$
$$\times[Ez^{\tau_1+\tau_2+1}\,{}_2F_1(-\tau_1,\tau_2+1;\tau_1+\tau_2+2;z)$$
$$+D\,{}_2F_1(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)]. \tag{13}$$

The transformation $\alpha\to z$ maps the entire lower half of the α plane into the unit circle in the z plane. Since H must be single valued and analytic around the point $z=0$ (corresponding to $\alpha=-i$), the constant E is identically zero. Thus we have

$$H=D(\alpha-i)^{2\tau_1+\tau_2+1}$$
$$\times{}_2F_1(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z). \tag{14}$$

We note that by using relation (3.251 (2) on page 295 of Ref. 12 the value of the integral H in Eq. (9a) as $\alpha\to\infty$ can be obtained in closed form as

$$H_{\alpha\to\infty}=\alpha^{\tau_2}B(-\tau_1-\tfrac{1}{2},\tfrac{1}{2}), \tag{15}$$

where B stands for the beta function. So we can obtain the constant D by comparing the asymptotic value (as $\alpha\to\infty$) of H from Eq. (14) with that from Eq. (15).

With the help of the linear transformation, formula 15.3.6 on page 559 of Ref. 13 for the Gauss hypergeometric functions, Eq. (14) can be recast as

$$H = D(\alpha - i)^{2\tau_1+\tau_2+1}\left[\frac{\Gamma(-\tau_1-\tau_2)\Gamma(2\tau_1+1)}{\Gamma(\tau_1+1)\Gamma(-\tau_2)} \times {}_2F_1\left(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;\frac{-2i}{\alpha-i}\right) + \left(\frac{-2i}{\alpha-i}\right)^{2\tau_1+1}\frac{\Gamma(-\tau_1-\tau_2)\Gamma(-2\tau_1-1)}{\Gamma(-2\tau_1-\tau_2-1)\Gamma(-\tau_1)} \times {}_2F_1\left(\tau_1+1,-\tau_2;2\tau_1+2;\frac{-2i}{\alpha-i}\right)\right]. \quad (16)$$

As $\alpha\to\infty$ the second term within the large bracket in Eq. (16) is the dominating one since $\mathrm{Re}(2\tau_1+1)$ is negative. So we can write

$$H_{\alpha\to\infty} = D\alpha^{\tau_2}2^{2\tau_1+1}e^{-i(\pi/2)(2\tau_1+1)} \times\frac{\Gamma(-\tau_1-\tau_2)\Gamma(-2\tau_1-1)}{\Gamma(-2\tau_1-\tau_2-1)\Gamma(-\tau_1)}. \quad (17)$$

From Eqs. (15) and (17) we have

$$D = \frac{e^{i(\pi/2)(2\tau_1+1)}\Gamma(\frac{1}{2})\Gamma(-2\tau_1-\tau_2-1)\Gamma(-\tau_1-\frac{1}{2})}{2^{2\tau_1+1}\Gamma(-\tau_1-\tau_2)\Gamma(-2\tau_1-1)}. \quad (18)$$

Substituting the value of D from Eq. (18) in Eq. (14) we finally get

$$H = \frac{e^{i(\pi/2)(2\tau_1+1)}\Gamma(\frac{1}{2})\Gamma(-2\tau_1-\tau_2-1)\Gamma(-\tau_1-\frac{1}{2})(\alpha-i)^{2\tau_1+\tau_2+1}}{2^{2\tau_1+1}\Gamma(-\tau_1-\tau_2)\Gamma(-2\tau_1-1)}\,{}_2F_1(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z). \quad (19)$$

This expression for H can be checked by directly evaluating the basic integral H in Eq. (9a) for $\alpha = 0$ and comparing this result with the corresponding result obtained from Eq. (19).

For the direct evaluation of H in Eq. (9a) we split the range of integration into two parts and write

$$H = \int_0^\infty (x^2+1)^{\tau_1}(x+\alpha)^{\tau_2}\,dx + \int_{-\infty}^0 (x^2+1)^{\tau_1}(x+\alpha)^{\tau_2}\,dx. \quad (20)$$

Changing x to $-x$ in the second integral and taking the phase term consistent with the imaginary part of x, which is negative in our case, we obtain

$$H = \int_0^\infty (x^2+1)^{\tau_1}(x+\alpha)^{\tau_2}\,dx + e^{-i\pi\tau_2}\int_0^\infty (x^2+1)^{\tau_1}(x-\alpha)^{\tau_2}\,dx. \quad (21)$$

Setting $\alpha = 0$ and using the relation 3.251 (2) on page 295 of Ref. 12, we arrive at

$$H_{\alpha\to0} = \frac{1+e^{-i\pi\tau_2}}{2}B\left(\frac{\tau_2+1}{2},\frac{-2\tau_1-\tau_2-1}{2}\right). \quad (22)$$

We now evaluate H for $\alpha = 0$ from Eq. (19). Putting $\alpha = 0$ we get

$$H_{\alpha\to0} = D(-i)^{2\tau_1+\tau_2+1}\,{}_2F_1(a,b;c;-1), \quad (23a)$$

where $a = 2\tau_1-\tau_2-1$, $b = -\tau_1$, and $c = -\tau_1-\tau_2$. Since $a-b+1 = c$, we can use the relation 15.1.21 on page 557 of Ref. 13 to obtain

$$H_{\alpha\to0} = D(-2i)^{2\tau_1+\tau_2+1}\frac{\Gamma(\frac{1}{2})\Gamma(-\tau_1-\tau_2)}{\Gamma(-\tau_1/2+\frac{1}{2})\Gamma(-\tau_1-\tau_2/2)}. \quad (23b)$$

Substituting the value of D in Eq. (23b) and using the results[14]

$$\Gamma(-2\tau_1-1) = (2\pi)^{-1/2}2^{-2\tau_1-3/2}\Gamma(-\tau_1-\tfrac{1}{2})\Gamma(-\tau_1),$$
$$\Gamma(-2\tau_1-\tau_2-1) = (2\pi)^{-1/2}2^{-2\tau_1-\tau_2-3/2}\Gamma(-\tau_1-\tau_2/2-\tfrac{1}{2}) \times\Gamma(-\tau_1-\tau_2/2),$$

and

$$\Gamma(-\tau_2/2+\tfrac{1}{2})\Gamma(\tau_2/2+\tfrac{1}{2}) = \pi\csc\pi(\tau_1/2+\tfrac{1}{2}),$$

we finally get exactly the same result for H as in Eq. (22).

III. 1s–1s CAPTURE

For 1s–1s transition,

$$\tilde\phi_f(\mathbf{p}) = (2\sqrt{2}\lambda^{5/2}/\pi)(p^2+\lambda^2)^{-2},$$

so that Eq. (2) becomes

$$T^{(0)}_{\mathrm{SPB}} = \frac{ch_0 2\sqrt{2}\lambda^{5/2}}{\pi K^2(4v^2)^{-iv}}\frac{\partial}{\partial\mu}(\kappa^{2iv-2}I_{20}). \quad (24)$$

Differentiating with respect to μ, we obtain

$$T^{(0)}_{\mathrm{SPB}} = c_1(T_1+T_2), \quad (25)$$

where

$$c_1 = \frac{ch_0 2\sqrt{2}\lambda^{5/2}k^{2iv-2}}{\pi K^2(4v^2)^{-iv}},$$
$$T_1 = 2\mu(iv-1)\kappa^{-2}I_{20},$$
$$I_{20} = D_{20}\left[2(2-iv)(2L)^{-iv}H + (2L)^{1-iv}\frac{\partial H}{\partial L}\right],$$
$$D_{20} = \pi(Z_P)^{-1-3iv}/(1-iv)(2-iv),$$
$$T_2 = D_{20}\left[\left\{4(2-iv)(-iv)(2L)^{-1-iv}H + 2(1-iv)(2L)^{-iv}\frac{\partial H}{\partial L}\right\}\frac{\partial L}{\partial\mu} + 2(2-iv)(2L)^{-iv}\frac{\partial H}{\partial\mu} + (2L)^{1-iv}\frac{\partial}{\partial\mu}\left(\frac{\partial H}{\partial L}\right)\right],$$

$$\frac{\partial H}{\partial L}=\frac{\partial H}{\partial \alpha}\frac{\partial \alpha}{\partial L},\quad \frac{\partial H}{\partial \mu}=\frac{\partial H}{\partial \alpha}\frac{\partial \alpha}{\partial \mu},$$

$$\frac{\partial}{\partial \mu}\left(\frac{\partial H}{\partial L}\right)=\frac{\partial \alpha}{\partial L}\frac{\partial^2 H}{\partial \alpha^2}\frac{\partial \alpha}{\partial \mu}+\frac{\partial H}{\partial \alpha}\frac{\partial}{\partial \mu}\left(\frac{\partial \alpha}{\partial L}\right),$$

$$\frac{\partial H}{\partial \alpha}=D\Big[(2\tau_1+\tau_2+1)(\alpha-i)^{2\tau_1+\tau_2}\times {}_2F_1(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)+(\alpha-i)^{2\tau_1+\tau_2+1}\frac{\partial z}{\partial \alpha}\times {}_2F_1'(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)\Big],$$

$$\frac{\partial^2 H}{\partial \alpha^2}=D\Big[(2\tau_1+\tau_2+1)(2\tau_1+\tau_2)(\alpha-i)^{2\tau_1+\tau_2-1}\times {}_2F_1(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)+2(2\tau_1+\tau_2+1)(\alpha-i)^{2\tau_1+\tau_2}\frac{\partial z}{\partial \alpha}\times {}_2F_1'(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)+(\alpha-i)^{2\tau_1+\tau_2+1}\Big\{\Big(\frac{\partial z}{\partial \alpha}\Big)^2\times {}_2F_1''(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)+\frac{\partial^2 z}{\partial \alpha^2}\,{}_2F_1'(-2\tau_1-\tau_2-1,-\tau_1;-\tau_1-\tau_2;z)\Big\}\Big], \tag{26}$$

$$\frac{\partial}{\partial \mu}\left(\frac{\partial \alpha}{\partial L}\right)=\frac{1}{2\lambda}\left(-\frac{1}{L^2}\frac{\partial A}{\partial \mu}+\frac{2A}{L^3}\frac{\partial L}{\partial \mu}\right),$$

$$\frac{\partial \alpha}{\partial \mu}=\frac{1}{2\lambda}\left(\frac{1}{L}\frac{\partial A}{\partial \mu}-\frac{A}{L^2}\frac{\partial L}{\partial \mu}\right),$$

$$\frac{\partial L}{\partial \mu}=-\frac{\mu+i\mathbf{J}\cdot\hat{v}}{L},\quad \frac{\partial A}{\partial \mu}=2(\mu-iv),$$

$$\frac{\partial \alpha}{\partial L}=-\frac{A}{2\lambda L^2},$$

$$\frac{\partial z}{\partial \alpha}=-\frac{2i}{(\alpha-i)^2},$$

$$\frac{\partial^2 z}{\partial \alpha^2}=\frac{4i}{(\alpha-i)^3},$$

and the primes on the ${}_2F_1$ functions denote the derivatives with respect to z.

The total cross section can now be calculated from the integral over the transverse momentum transfer $K_\perp$:

$$\sigma_{\mathrm{SPB}}^{(0)}=(2\pi v^2)^{-1}\int_0^\infty |T_{\mathrm{SPB}}^{(0)}|^2 K_\perp\, dK_\perp, \tag{27}$$

where $K_\perp$ is defined by the relation

$$K_\perp^2=K^2-K_z^2,$$

with

$$K_z=-\tfrac{1}{2}(v+(Z_T^2-Z_P^2)/v).$$

For $1s$–nlm capture with $l>0$ further dependence on $\hat{p}$ in Eq. (2) occurs through $\tilde{\phi}_f^*(p)$. In that case also the SPB amplitude can be calculated from the integral I_{MN} by suitable parametric differentiations following the method discussed in Sec. III of the work of Sil and McGuire.[7]

IV. RESULTS AND DISCUSSIONS

A generalized computer program has been developed to evaluate the ${}_2F_1$ functions appearing in Eq. (19) for any value

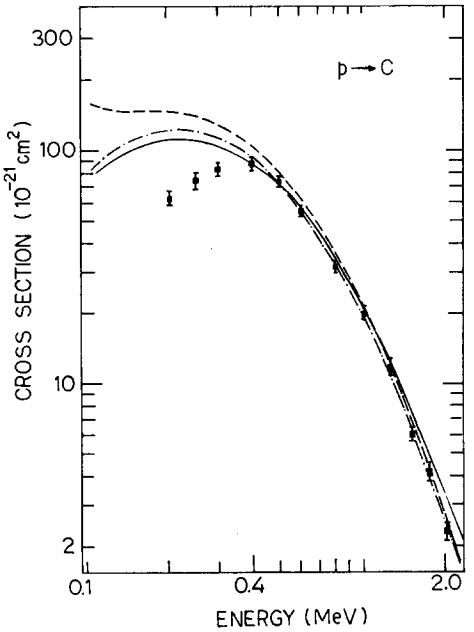


FIG. 1. Cross sections for electron capture by protons from C (in units of 10^{-21} cm^2) for the energy range 0.1 to 2.0 MeV: full curve, present result; dashed–dotted curve, results of SPB full peaking (Ref. 4); dashed curve, results of SPB transverse peaking (Ref. 5); ■, experimental data of Rodbro *et al.*[16]

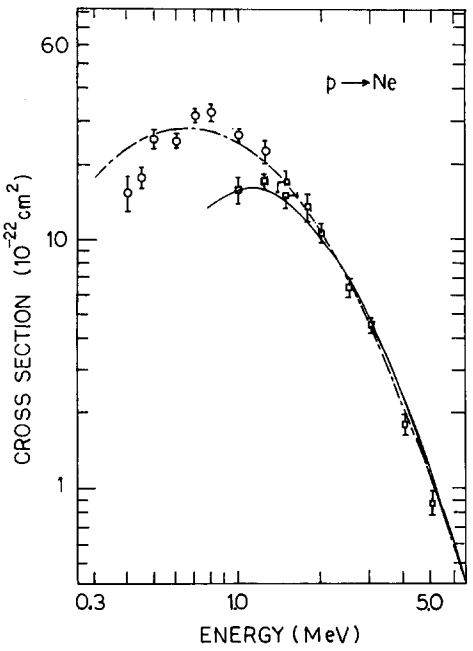


FIG. 2. Cross section for the electron capture by protons from Ne (in units of 10^{-22} cm^2) for the energy range 0.3 to 5 MeV. Full curve, present results; dashed curve, results of SPB transverse peaking (Ref. 5); □, experimental data of Cocke *et al.*[17]; O, experimental data of Rodbro *et al.*[16]

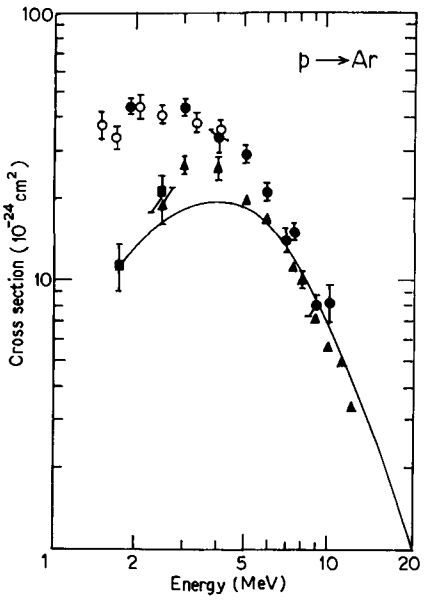


FIG. 3. Cross section for the electron capture by protons from Ar (in units of 10^{-24} cm^2) for the energy range 2 to 20 MeV. Full curve, present result; experimental data, O, Horsdal-Pedersen *et al.*[18] from University of Aarhus; ●, experimental data from Kansas State University; ■, Andriamonje *et al.*[19]; ▲, Macdonald *et al.*[20]

of the complex argument z. Using this program we have calculated the results for the integral H from Eq. (19) for several physical values of the parameters and compared these results with the corresponding values obtained by directly evaluating the integral in Eq. (9a) numerically; the two sets of results are in very good agreement with each other. We have computed the $1s$–$1s$ cross sections for electron capture by protons from the atomic targets Li, C, Ne, Na, and Ar for the energy range 400 keV to 20 MeV (lab). The Hylleraas[15]-type effective target charge has been taken.

Figures 1 and 2 present our $1s$–$1s$ capture cross sections following Eq. (27) for protons on carbon and neon along with the theoretical results in the full peaking and transverse peaking SPB approximations and also the experimental findings. For the carbon target our results are in good agreement with the experimental data of Rodbro *et al.*[16] for the energy range $E \geqslant 400$ keV. Results in full peaking and in transverse peaking SPB approximations also agree well with the experiment for the same energy range. Below this energy range our results are 5%–10% smaller than the results of Macek and Alston[4] but 20%–40% larger than the experimental data. Our results for the Ne target are in close agreement with the

TABLE I. Cross sections (in units of 10^{-20} cm^2) for the electron capture by protons from the K shell of Li.

E (MeV)	Total cross sections (10^{-20} cm^2)
0.4	78.18
0.5	36.04
0.6	18.12
0.7	9.774
0.8	5.584
0.9	3.346
1.0	2.088
2.0	0.0715

TABLE II. Cross sections (in units of 10^{-22} cm^2) for the electron capture by protons from the K shell of Na.

E (MeV)	Total cross section (10^{-22} cm^2)
1	6.48
2	6.45
3	3.59
4	2.00
5	1.12
6	0.65
7	0.39
8	0.24
9	0.15

experimental data of Cocke *et al.*[17] for the energy range $E \geqslant 1$ MeV but below this energy range our results underestimate the experimental data of Rodbro *et al.*[16] The discrepancy between our results and the experimental data in the low-energy region may be due in part to the fact that we have calculated only the zeroth-order term in the expansion of Sil and McGuire.[7]

In Fig. 3 our results for the electron capture from the K shell of Ar by protons are shown together with the experimental data of Horsdal-Pedersen *et al.*,[18] Andriamonje *et al.*,[19] and Macdonald *et al.*[20] Our results agree well with the experimental data for $E \geqslant 5$ MeV.

The calculated cross sections for the electron capture by protons from the K shell of Li and Na are tabulated in Tables I and II, respectively; the energy ranges considered are $0.4 \leqslant E \leqslant 2$ MeV for the Li target and $1 \leqslant E \leqslant 9$ MeV for the Na target.

Since there are two electrons in the K shell the calculated cross sections for a single electron capture have been multiplied by 2 for the presentation of the results in all the figures and tables.

ACKNOWLEDGMENT

J.H.M. was supported by the Division of Chemical Sciences, U. S. Department of energy.

APPENDIX: EVALUATION OF THE INTEGRAL $f(p)$ FOR COMPLEX L

In our problem

$$\mathbf{L} = \mathbf{L}_1 - i\mathbf{L}_2,$$

where $L_1 = \mathbf{J}$ and $L_2 = \mu\hat{v}$.

We choose the Z axis along $\mathbf{L}_1$ and the XZ plane is taken as the plane containing $\mathbf{L}_1$ and $\mathbf{L}_2$, so that in the polar coordinates, $\mathbf{L}_1 = (L_1,0,0)$ and $\mathbf{L}_2 = (L_2, \beta, 0)$.

Then

$$A + 2\mathbf{p} \cdot \mathbf{L} = A + B\cos\theta + C\sin\theta\cos\phi, \qquad \text{(A1)}$$

where

$$B = 2p(L_1 - iL_2\cos\beta),$$
$$C = -2iL_2 p\sin\beta. \tag{A2}$$

Here we note that

$$\sqrt{B^2 + C^2} = 2Lp, \tag{A3}$$

where

$$L = (L_1^2 - L_2^2 - 2iL_1L_2\cos\beta)^{1/2} = (\mathbf{L}\cdot\mathbf{L})^{1/2}. \tag{A4}$$

Using Eq. (A1) in Eq. (6a) we can write, for $|A| > |B| + |C|$,

$$f(p) = \sum_p \sum_s \int_0^{\pi}\int_0^{2\pi} \frac{(-\tau)_{r+s} A^{\tau-r-s}(-B)^r(-C)^s \cos^r\theta \sin^{r+1}\theta \cos^s\phi \, d\theta \, d\phi}{r!s!}, \tag{A5}$$

where $(a)_r$ stands for Pochhammer's symbol, $(a)_r = a(a+1)\cdots(a+r-1)$, for $r = 1,2,3,\ldots$, and $(a)_0 = 1$.

For odd values of s, the ϕ integral in Eq. (A5) vanishes. So putting $s = 2h$ and using the relation 3.621 (1) on page 369 of Ref. 12 we obtain

$$\int_0^{2\pi} \cos^{2h}\phi \, d\phi = \frac{2^{2h+1}[\Gamma(h+\frac{1}{2})]^2}{\Gamma(2h+1)}. \tag{A6}$$

The θ integral will survive only for even values of r. We put $r = 2j$ and use the relation 3.621 (5) on page 369 of Ref. 12 to get

$$\int_0^{\pi} \cos^{2j}\theta \sin^{2h+1}\theta \, d\theta = \frac{\Gamma(h+1)\Gamma(j+\frac{1}{2})}{\Gamma(j+h+\frac{3}{2})}. \tag{A7}$$

On substitution of (A6) and (A7) in Eq. (A5) we arrive at

$$f(p) = 2A^{\tau}\sum_j \sum_h \frac{(-\tau)_{2(j+h)}(B^2/A^2)^j(C^2/A^2)^h 2^{2h}[\Gamma(h+\frac{1}{2})]^2\Gamma(h+1)\Gamma(j+\frac{1}{2})}{(2j)!(2h)!\Gamma(2h+1)\Gamma(j+h+\frac{3}{2})}. \tag{A8}$$

Using the result[14]

$$\Gamma(2h+1) = (2^{2h}/\sqrt{\pi})\Gamma(h+\tfrac{1}{2})\Gamma(h+1),$$

and noting that

$$\Gamma(n+\tfrac{1}{2}) = \pi^{1/2}(\tfrac{1}{2})_n$$

and

$$(2n)! = 2^{2n}n!(\tfrac{1}{2})_n,$$

we can write

$$f(p) = 4\pi A^{\tau}\sum_j \sum_h \frac{(-\tau)_{2(j+h)}(B^2/A^2)^j(C^2/A^2)^h}{2^{2(j+h)}j!h!(\frac{3}{2})_{j+h}}. \tag{A9}$$

Writing $j + h = l$ and performing the summation over h we obtain

$$f(p) = 4\pi A^{\tau}\sum_l \frac{(-\tau)_{2l}}{2^{2l}(\frac{3}{2})_l l!}\left(\frac{B^2+C^2}{A^2}\right)^l. \tag{A10}$$

We now show that the above expansion of $f(p)$ is equivalent to the expression in Eq. (7)

Substituting the value of $2Lp$ from Eq. (A3) in Eq. (7) and using binominal expansion we get

$$f(p) = \frac{2\pi A^{\tau+1}}{(\tau+1)(B^2+C^2)^{1/2}}\sum_m \frac{\{-(\tau+1)\}_m}{m!} \times\left[\left\{-\frac{(B^2+C^2)^{1/2}}{A}\right\}^m - \left\{\frac{(B^2+C^2)^{1/2}}{A}\right\}^m\right]. \tag{A11}$$

The condition for convergence of each series is automatically satisfied when $|A| > |B| + |C|$. It is clear that only odd values of m will survive in the expansion (A11). So we take $m = 2l + 1$ to obtain

$$f(p) = -\frac{4\pi A^{\tau}}{(\tau+1)}\sum_l \frac{\{-(\tau+1)\}_{2l+1}}{(2l+1)!}\left(\frac{B^2+C^2}{A^2}\right)^l. \tag{A12}$$

Using the results

$$\{-(\tau+1)\}_{2l+1} = -(\tau+1)(-\tau)_{2l}$$

and

$$(2l+1)! = 2^{2l}l!(\tfrac{3}{2})_l$$

in Eq. (A12) we finally obtain the same expression for $f(p)$ as in Eq. (A10). This proves that the result in Eq. (7) is justified for a complex vector $\mathbf{L}$ as well under the restriction $|A| > |B| + |C|$. Validity of the result for the general case follows from analytic continuation.

[1] J. S. Briggs, J. Phys. B **10**, 3075 (1977).
[2] J. H. Macek and R. Shakeshaft, Phys. Rev. A **22**, 1441 (1980).
[3] J. Macek and K. Taulbjerg, Phys. Rev. Lett. **46**, 170 (1981).
[4] J. Macek and S. Alston, Phys. Rev. A **26**, 250 (1982).
[5] S. Alston, Phys. Rev. A **27**, 2342 (1983).
[6] L. J. Dube, J. Phys. B **17**, 641 (1984).
[7] N. C. Sil and J. H. McGuire, J. Math. Phys. **26**, 845 (1985).
[8] J. H. McGuire and N. C. Sil, Phys. Rev. A **28**, 3679 (1983).
[9] E. J. Kelsey and J. Macek, J. Math. Phys. **17**, 1182 (1976).
[10] This is Eq. (4.17) of Ref. 4 where the notations are defined in detail.
[11] H. Bateman, *Higher Transcendental Functions* (McGraw-Hill, New York, 1953), Vol. 1, p. 60, Sec. 2.1.3.
[12] I. S. Granshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1980).
[13] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (McGraw-Hill, New York, 1970).
[14] These are formulas 6.1.18 and 6.1.17 on page 256 of Ref. 13.
[15] E. A. Hylleraas, Z. Phys. **54**, 347 (1929).
[16] M. Rodbro, E. Horsdal-Pedersen, C. L. Cocke, and J. R. Macdonald, Phys. Rev. A **19**, 1936 (1979).
[17] C. L. Cocke, R. K. Gradner, B. Curnutte, T. Bratton, and T. K. Saylor, Phys. Rev. A **16**, 2248 (1977).
[18] E. Horsdal-Pedersen, C. L. Cocke, J. L. Rasmussen, S. L. Varghese, and W. Waggoner, J. Phys. B **16**, 1799 (1983).
[19] S. Andriamonje, J. F. Chemin, J. Routurier, J. N. Scheurer, H. Laurent, and J. P. Schapira, in *Abstracts of the Proceedings of the Twelfth International Conference on the Physics of Electronic and Atomic Collisions*, Gatlinburg, 1981, edited by S. Datz (North-Holland, Amsterdam, 1982), p. 657.
[20] J. R. Macdonald, C. L. Cocke, and W. W. Eidson, Phys. Rev. Lett. **32**, 648 (1974).

Hamilton–Jacobi formulation and quantum theory of thermal wave propagation in the solid state

Andreas Mandelis
Photoacoustics Laboratory, Department of Mechanical Engineering, University of Toronto, Toronto, Ontario, Canada M5S 1A4

(Received 27 December 1984; accepted for publication 7 June 1985)

A mathematical formalism has been developed for the description of the propagation of thermal waves in solids. The theory is based on the analogy between thermal waves and mechanics as manifested upon expressing the heat conduction equation in the Hamilton–Jacobi formalism. The transition of the classical formulation to quantum mechanics is accomplished by defining thermal wave operators for the generalized coordinates and the canonically conjugated momenta. The resulting theory shows that propagating, albeit heavily damped, thermal waves can be formally described by a quantum mechanical thermal harmonic oscillator (THO). The expectation values of observables, derived from Ehrenfest's theorems, are found to be of practical importance for the description of the thermal wave field in solids with inhomogeneous thermal and thermodynamic properties.

I. INTRODUCTION

In recent years photoacoustic spectroscopy has been used for the creation and detection of thermal waves in solids in general,[1–8] and in crystalline semiconductors in particular.[9–12] The ability of thermal waves to perform nondestructive depth-profiling studies in materials has been exploited mainly qualitatively due to the lack of appropriate theoretical models in the literature. It has been recognized for some time that the depth profiling of dopant concentrations in semiconductors may be the single most important application of thermal wave physics.[10,11,13] An important theoretical obstacle, however, to the realization of the full potential of thermal wave depth profilometry appears to be the complexity of the mathematical description of thermal wave propagation in continuous solids, especially those which exhibit large local variations of their relevant thermal and thermodynamic properties, i.e., the thermal conductivity, the density, and the specific heat. As a result of the mathematical difficulties, only two theoretical treatments have appeared in the literature, which assume discrete, multilayered solid structures with constant thermal and thermodynamic properties within each thin layer.[13,14] Furthermore, Afromowitz *et al.*[15] have applied discrete Laplace transformations to the heat conduction equation to treat the production of the photoacoustic signal in a solid with continuously variable optical absorption coefficient as a function of depth, however, the thermal parameters of the solid were assumed constant. Thomas *et al.*[3] calculated the Green's function for the three-dimensional heat conduction equation describing thermal wave propagation in a thermally uniform solid with a subsurface discontinuity ("flaw"). In most experimental situations of interest one has to deal with thermal wave propagation in fields where drastic variations of thermal properties occur within a thermal wavelength[16] from the source of oscillation. In this limit the thermal wave behavior can be described formally by using the analogy between classical mechanical plane wave propagation and thermal wave motion. The crucial difference between these two types of waves is that the former kind is a result of the time harmonic solutions to the Helmholtz equation, which is a hyperbolic partial differential equation of second order in the derivative of the wave-function field, whereas the latter kind is a result of time harmonic solutions to the Fourier equation of heat conduction, which is a parabolic partial differential equation of first order in the time derivative of the thermal field wave function. The above difference has a profound effect in the nature of the two kinds of waves: the hyperbolic type is able to propagate freely within matter and suffers attenuation only when the wave vector in the medium has an imaginary component, while the parabolic type always exhibits heavy attenuation as a function of propagation distance in the medium.

In this paper the formal analogy between classical and quantum wave fields and a thermal wave field will be investigated. It will be shown that the thermal wave field Hamiltonian is nondissipative irrespective of the spatial dependence of the relevant thermal/thermodynamic properties of the system. This property, in turn, allows the definition of an eikonal equation and Fermat's principle for thermal wave propagation, as well as the definition of a fundamental spatial eigenfrequency of the thermal oscillations.

The classical mechanical Hamiltonian of the thermal wave field will be further shown to be that of a harmonic oscillator in the temperature potential field. This observation allows the quantization of the thermal wave field, which sets the foundations of the microscopic description of wave phenomena occurring within propagation distances on the order of a thermal wavelength, such as thermal diffraction. The macroscopic thermal wave equations can indeed be recovered in a quantum mechanical expectation function form after introducing expansions of the thermal wave field observables in terms of integrals over complete sets of eigenfunctions of the thermal Schrödinger equations. The formalism allows the explicit evaluation of the macroscopic tempera-

 0022-2488/85/102676-08$02.50

ture field in continuous solids with nonhomogeneous thermal/thermodynamic properties as a result of the propagation of thermal wave fronts in the potential field defined by the temperature-dependent generalized coordinate. A similar approach was used successfully[17,18] in the quantization of light rays propagating in solids, where both Schrödinger-type[17] and Dirac-type[18] theories were advanced for the description of the optical wave field observables.

II. HAMILTON–JACOBI FORMULATION OF THERMAL WAVE PHYSICS

The generation of thermal waves in a medium is the result of the presence of a harmonic heat source at the origin, modulated at some angular frequency ω_0. The most commonly used heat sources are modulated laser beams,[1,3–5,8,12] electron beams,[2,9,11] or ac current sources.[19] The temperature field $\theta(\mathbf{r},t)$ in the medium excited by the heat source is given by the macroscopic Fourier heat conduction equation, which is a statement of energy balance in the medium:

$$\nabla\cdot[k(\mathbf{r})\nabla\theta(\mathbf{r},t)] - \rho(\mathbf{r})c(\mathbf{r})\frac{\partial}{\partial t}\theta(\mathbf{r},t) = 0, \tag{2.1}$$

where $k(\mathbf{r})$, $\rho(\mathbf{r})$, and $c(\mathbf{r})$ are the (generally spatially variant) thermal conductivity (W/m °K), density (kg/m^3), and specific heat (J/kg °K), respectively, of the medium. For most experimental thermal wave solid state geometries $c(\mathbf{r})$ implies the specific heat at constant volume, however, constant pressure specific heats must be used when the experimental conditions require it. In this work the one-dimensional counterpart of (2.1) will be used in order to simplify the formalism, with the three-dimensional case constituting a straightforward extension of the fundamental concepts developed herein.

A harmonic time dependence of the one-dimensional temperature field of the form

$$\theta(x,t) = T(x)\exp(i\omega_0 t) \tag{2.2}$$

yields the Fourier–Helmholtz equation

$$\frac{d}{dx}\left[k(x)\frac{d}{dx}T(x)\right] - i\omega_0\rho(x)c(x)T(x) = 0. \tag{2.3}$$

In passing we note that (2.3) is in the form of the Liouville equation with the single eigenvalue $\lambda = i\omega_0$. The Lagrangian function, which corresponds to (2.3), is[20]

$$\mathscr{L} = \frac{1}{2}k(x)\left[\frac{dT(x)}{dx}\right]^2 + i\omega_0\int_0^T y\,dy \tag{2.4}$$

and satisfies the Euler equation

$$\frac{\partial\mathscr{L}}{\partial T} - \frac{d}{dx}\frac{\partial\mathscr{L}}{\partial(dT/dx)} = 0. \tag{2.5}$$

In thermal wave configurations the following boundary conditions are usually assumed[13,14]:

$$T(x=0) = T_0, \tag{2.6a}$$

$$-k(x)\frac{dT(x)}{dx}\bigg|_{x=0} = Q_0, \tag{2.6b}$$

where the domain of the temperature field encompasses the one-dimensional volume (x). The surface (s) which encloses (x) is defined by the origin $(x=0)$ and some suitable point $L < +\infty$. The Hamiltonian is given by

$$H(x,T,p_T) = p_T\left[\frac{dT(x)}{dx}\right] - \mathscr{L}$$

$$= p_T^2/2k(x) - (i/2)\,\omega_0\rho(x)c(x)T^2, \tag{2.7}$$

with p_T the generalized momentum defined by

$$p_T = \frac{\partial\mathscr{L}}{\partial(dT/dx)} = k(x)\frac{dT(x)}{dx}. \tag{2.8}$$

Equations (2.7) and (2.8) show that for the thermal wave problem the generalized coordinate and momentum are the field temperature and heat flux, respectively. The Hamiltonian form (2.7) is not appropriate, however, for use in the consideration of thermal wave dynamics, because it is an explicit function of the spatial coordinate (x). A canonical transformation is required, such that both coordinate and momentum will be constants of the motion. Consider the following parametric transformations:

$$\zeta \equiv \frac{1}{J}\int_0^x\left[\frac{\rho(y)c(y)}{k(y)}\right]^{1/2}dy, \tag{2.9}$$

$$\tau \equiv [k(x)\rho(x)c(x)]^{1/4}T(x), \tag{2.10}$$

and

$$J \equiv \frac{1}{L}\int_0^L\left[\frac{\rho(y)c(y)}{k(y)}\right]^{1/2}dy. \tag{2.11}$$

The generating function, which produces the desired canonical transformation, is Hamilton's principal function[21] S, in which the time coordinate has been replaced by the spatial coordinate (x). The Hamilton–Jacobi equation

$$\frac{\partial S}{\partial x} = -H\left(x,T,\frac{\partial S}{\partial T}\right)$$

$$= -\frac{1}{2k(x)}\left(\frac{\partial S}{\partial T}\right)^2 + \frac{i}{2}\omega_0\rho(x)c(x)\,T^2 \tag{2.12}$$

transforms to the following equations, after introduction of the new variables ζ and τ:

$$\frac{\partial S}{\partial\zeta} = -\frac{J}{2}\left[\left(\frac{\partial S}{\partial\tau}\right)^2 - i\omega_0\tau^2\right]. \tag{2.13}$$

Using separation of variables in the form

$$S(\tau,\zeta,\alpha) = W(\tau,\alpha) - \alpha\zeta \tag{2.14}$$

yields an equation for Hamilton's characteristic function $W(\tau,\alpha)$:

$$\frac{J}{2}\left[\left(\frac{\partial W}{\partial\tau}\right)^2 - i\omega_0\tau^2\right] = \alpha = \text{const.} \tag{2.15}$$

Equation (2.15) is the canonical transformation in which the new coordinate ζ is cyclic and the transformed Hamiltonian is a constant of the motion and assumes the meaning of the total generalized energy of the thermal wave field:

$$H\left(\tau,\frac{\partial W}{\partial\tau}\right) = \alpha \equiv E. \tag{2.16}$$

In the Hamilton–Jacobi theory of thermal waves the generalized coordinate is $\tau(\zeta)$ and the generalized momentum is $p_\tau = \partial W/\partial\tau$. The functional form of the canonical Hamiltonian

$$H(\tau,p_\tau) = \tfrac{1}{2}Jp_\tau^2 + \tfrac{1}{2}K\tau^2 \tag{2.17}$$

shows that thermal wave field behavior is equivalent to that

of a thermal harmonic oscillator (THO), with the effective mass $m = J^{-1}$ and spring constant K, which is subject to a restoring, conservative force $F = -K\tau$ generated by the effective harmonic potential field $V(\tau) = \frac{1}{2}K\tau^2$. The frequency of the oscillation in the canonical coordinates can be found via the use of the action-angle variable

$$I_\tau = \oint p_\tau d\tau = \oint \left[\frac{\partial W(\tau,\alpha)}{\partial \tau}\right] d\tau. \tag{2.18}$$

Defining the spring constant

$$K \equiv -i\omega_0 J, \tag{2.19}$$

the integral of (2.18) can be evaluated between 0 and 2π to yield the frequency of the motion[21] of the THO,

$$\nu_\tau = \frac{\partial H}{\partial I_\tau} = \frac{1}{2\pi}(KJ)^{1/2}, \tag{2.20}$$

so that an angular frequency can be written as

$$\Omega_\tau \equiv 2\pi\nu_\tau = (KJ)^{1/2}\ [m^{-1}]$$

or

$$\Omega_\tau(\omega_0) = \pm\frac{(1-i)}{L}\int_0^L \left[\frac{\omega_0 \rho(y)\, c(y)}{2k(y)}\right]^{1/2} dy$$

$$\equiv \pm\frac{(1-i)}{L}\int_0^L a_s(\omega_0, y)\, dy, \tag{2.21}$$

where $a_s(\omega_0, y)$ is the local thermal diffusion coefficient of the Rosencwaig–Gersho theory[22] at depth y in the medium. Here, Ω_τ is the spatial angular frequency of oscillation of the THO and is defined in terms of the spatial extent of the medium or, in the case of semi-infinite media, in terms of

$$\lim_{L\to\infty}\left(\frac{1}{L}\int_0^L \alpha_s(\omega_0, y)\, dy\right).$$

The Hamilton–Jacobi formulation of the thermal wave problem leads to an eikonal equation of thermal wave physics, upon combination of (2.15) and (2.16):

$$(\nabla_\tau W)^2 = p_\tau^2. \tag{2.22}$$

Equation (2.22) is the rule for the construction of the surfaces of constant phase, via the equation

$$W(\tau, \alpha) = \text{const.} \tag{2.23}$$

The thermal gradient of W determines the normal to such surfaces. Any surface (s) that satisfies the condition (2.23) is a surface of constant thermal phase and thus defines a thermal wave front. The thermal ray trajectories are determined everywhere in space by (2.22) and are perpendicular to the wave fronts, whose phase velocities are given by

$$(v_\tau)_p = \frac{E}{|\partial W/\partial\tau|} = \left(\frac{J}{2}\right)^{1/2}\frac{E}{\sqrt{E - V(\tau)}}. \tag{2.24}$$

At this stage, a variational Fermat's principle can be formulated for the description of geometrical thermal ray trajectories, analogous to those of classical mechanics and geometrical optics. This principle can be expressed as follows:

$$\delta\int \frac{ds}{(v_\tau)_p} = \delta\int p_\tau\, ds = 0, \tag{2.25}$$

where ds is the incremental length in the configuration space spanned by the generalized coordinate τ and the conjugate momentum p_τ.

Finally, it is interesting to note that there exist three Poisson brackets for the canonical variables of the one-dimensional thermal wave problem,

$$\{\tau, p_\tau\} = 1, \tag{2.26a}$$

$$\{\tau, H\} = Jp_\tau, \tag{2.26b}$$

$$\{p_\tau, H\} = i\omega_0 \tau J, \tag{2.26c}$$

along with the relation

$$\frac{dH}{d\xi} = \frac{\partial H}{\partial \xi}, \tag{2.27}$$

which is a consequence of the fact that the total generalized energy of the medium is conserved over one cycle of spatial oscillation of the THO.

III. QUANTUM THEORY OF THERMAL WAVE PHYSICS

The classical theory of thermal wave fields in one dimension, which was presented in Sec. II above, is capable of describing only macroscopic thermal wave phenomena, such as the trajectories of thermal rays and the eikonal equation (2.22). In this limit of geometrical thermal wave physics the Fermat's principle (2.25) can be used successfully; however, it cannot rederive the Fourier–Helmholtz equation (2.3) through purely algebraic manipulations. Furthermore, the Fourier–Helmholtz equation is itself a reduced form of the canonical Hamilton–Jacobi equation (2.16) and thus it suffers from the mathematical disadvantages of noncanonical differential equations with regard to the degree of difficulty in obtaining the most general solution. In this section it will be shown that the thermal wave equation can be recovered and the most general solution can be obtained relatively easily from geometrical thermal ray physics via quantization, as by-products of thermal wave Ehrenfest's theorems. A complete analogy to the traditional quantum theory can be drawn upon replacing all classical variables of the Hamilton–Jacobi theory with thermal wave quantum mechanical operators:

$$\tau \to \hat{\tau} = \tau, \tag{3.1}$$

$$p_\tau \to \hat{p}_\tau = -i\hbar\frac{\partial}{\partial\tau}, \tag{3.2}$$

$$H \to \hat{H} = i\hbar\frac{\partial}{\partial\xi}. \tag{3.3}$$

The constant $\hbar$ appearing in (3.1)–(3.3) is the thermal wave equivalent of Planck's constant. The Hamiltonian operator $\hat{H}$ is in units of the generalized energy of the THO:

$$[\hat{H}] = [E] = \text{W °K/m}^3$$

so that

$$[\hbar] = \text{W °K/m}^2.$$

Equation (3.3) is now assumed to admit eigenfunction solutions

$$\hat{H}\psi = i\hbar\frac{\partial\psi}{\partial\xi}. \tag{3.4}$$

Use of the definitions (3.1) and (3.2) in (2.17) and insertion in (3.4) yields the canonical coordinate-dependent "Schrödinger equation" of thermal wave quantum mechanics:

$$-\mathcal{b}^2\left(\frac{J}{2}\right)\frac{\partial^2}{\partial\tau^2}\psi(\tau,\zeta)+V(\tau)\psi(\tau,\zeta)=i\mathcal{b}\frac{\partial}{\partial\zeta}\psi(\tau,\zeta), \quad (3.5)$$

where

$$V(\tau)=\tfrac{1}{2}K\tau^2. \quad (3.6)$$

Separation of variables in the form

$$\psi(\tau,\zeta)=\phi(\tau)\exp(-iE\zeta/\mathcal{b}) \quad (3.7)$$

gives the coordinate-independent "Schrödinger equation"

$$\frac{d^2}{d\tau^2}\phi(\tau)+\frac{2}{\mathcal{b}^2J}[E-V(\tau)]\,\phi(\tau)=0. \quad (3.8)$$

The solutions to (3.8) are, in principle, the eigenfunctions of the quantum mechanical harmonic oscillator; however, at this point it must be recognized that the constant $\mathcal{b}$ is consistent with real values of the generalized energies only if it is a complex quantity of the form

$$\mathcal{b}=(1+i)|\mathcal{b}|. \quad (3.9)$$

If (3.9) is satisfied, the anticipated harmonic oscillator eigenvalues[23] $E_n=(n+\frac{1}{2})\mathcal{b}\Omega_\tau$ will be real, as can be verified from (2.21). Using (3.9) and (2.19) in (3.8) results in the following equation:

$$|\mathcal{b}|^2J\frac{d^2}{d\tau^2}\phi(\tau)+(\omega_0J)\tau^2\phi(\tau)=2iE\phi(\tau). \quad (3.10)$$

Defining a new variable

$$z_1\equiv(4\omega_0/|\mathcal{b}|^2)^{1/4}\tau, \quad (3.11)$$

Eq. (3.10) becomes

$$\frac{d^2}{dz_1^2}\phi(z_1)+\left[\frac{z_1^2}{4}-i\left(\frac{E}{\omega_0|\mathcal{b}|J}\right)\right]\phi(z_1)=0. \quad (3.12)$$

Equation (3.12) has well-defined solutions if and only if

$$E/\omega_0^{\frac{1}{2}}|\mathcal{b}|J=p+\tfrac{1}{2} \quad (3.13a)$$

and

$$z_1=ze^{i\pi/4}, \quad (3.13b)$$

so that (3.12) may be transformed to

$$\frac{d^2}{dz^2}\phi_p(z)+\left(p+\frac{1}{2}-\frac{z^2}{4}\right)\phi_p(z)=0. \quad (3.14)$$

Equation (3.14) is a parabolic wave equation or Weber–Hermite equation[24] with eigenvalues

$$E_p=\omega_0^{\frac{1}{2}}|\mathcal{b}|J(p+\tfrac{1}{2}). \quad (3.15)$$

For general values of p the solutions to (3.14) are the Weber functions,[25] $\phi_p(z)=D_p(z)$, where

$$D_p(z)=2^{(p/2)}e^{-z^2/4}\left\{\frac{\sqrt{\pi}}{\Gamma[(1-p)/2]}\,{}_1F_1\left(-\frac{p}{2};\frac{1}{2};\frac{z^2}{2}\right)-\frac{z\sqrt{2\pi}}{\Gamma(-p/2)}\,{}_1F_1\left(\frac{1-p}{2};\frac{3}{2};\frac{z^2}{2}\right)\right\}, \quad (3.16)$$

where ${}_1F_1(a;c;z)$ is a degenerate hypergeometric function given by[26]

$${}_1F_1(a;c;z)=\frac{e^{ia\pi}z^{-a}}{\Gamma(a)}\int_0^\infty e^{-u}u^{a-1}\left(1+\frac{u}{z}\right)^{c-a-1}du, \quad (3.17)$$

with $z=|z|e^{i\theta}$, $-\pi<\theta<\pi$. For integral values of p, i.e., for $p=n$, where n is a positive integer, the Weber functions can be expressed by means of Hermite polynomials[25]:

$$D_n(z)=2^{-(n/2)}e^{-z^2/4}H(z/\sqrt{2}) \quad (3.18a)$$

$$=(-1)^ne^{z^2/4}\frac{d^n}{dz^n}e^{-z^2/2}. \quad (3.18b)$$

In the particular case where $p=n$ the eigenvalues of the Weber–Hermite equation are

$$E_n=(n+\tfrac{1}{2})\omega_0^{1/2}|\mathcal{b}|J. \quad (3.19)$$

Equations (2.11), (2.20), and (2.21) give

$$|\Omega_\tau|=\omega_0^{1/2}J, \quad (3.20)$$

so that (3.19) can be written

$$E_n=(n+\tfrac{1}{2})|\mathcal{b}|\,|\Omega_\tau| \quad (3.21a)$$

$$=(n+\tfrac{1}{2})\mathcal{b}\Omega_\tau. \quad (3.21b)$$

Now Eq. (3.21) can be used to interpret $\mathcal{b}$ of the thermal wave quantum theory in terms of the generalized energy of thermal wave packets (thermions!), as the constant ratio of the energy to the angular frequency of such wave packets:

$$E=\mathcal{b}\Omega_\tau=b\nu_\tau. \quad (3.22)$$

Equation (3.22) in conjunction with (2.22) and (2.24) yields a thermal wave equivalent of the de Broglie relation,

$$\lambda_{\rm th}=\frac{b}{p_\tau}=\frac{1}{2\pi}\left[\frac{\omega_0\rho(\zeta)c(\zeta)}{k(\zeta)}\right]^{1/2}. \quad (3.23)$$

Farther insight into the nature of the constant $\mathcal{b}$ is obtained upon writing (3.22) in the form

$$\mathcal{b}=\frac{E}{(1-i)\,[(1/L)\int_0^L a_s(\omega_0 y)\,dy]}=\frac{E}{\langle k_{\rm th}\rangle}, \quad (3.24)$$

where $k_{\rm th}$ is the one-dimensional component of the complex thermal wave vector[16]

$$\mathbf{k}_{\rm th}=k_{\rm th}\hat{\zeta}, \quad (3.25a)$$

and

$$|k_{\rm th}|=2\pi/\lambda_{\rm th}. \quad (3.25b)$$

The presence of an imaginary component in the thermal wave vector is responsible for the exponential attenuation of thermal waves propagating in a continuous medium. From another point of view, a medium whose thermal wave vector is given by (3.25) can be described as thermally lossy, in analogy to optically lossy media arising in the propagation of electromagnetic radiation.[27] Equations (3.24) and (3.25b) show that $\mathcal{b}$ is proportional to the wavelength $\langle\lambda_{\rm th}\rangle$ of the thermal wave packet, averaged over the entire extent of the propagation medium. The proportionality of $\mathcal{b}$ to $\langle\lambda_{\rm th}\rangle$ is analogous to that observed between the Planck's constant of the quantum theory of light rays and the wavelength of the optical radiation.[17] It is also consistent with the correspondence principle of quantum mechanics: The thermal wave Schrödinger equation (3.5) can be transformed using the substitution

$$\psi(\tau,\zeta)=A\exp[iW(\tau,\zeta)/\mathcal{b}] \quad (3.26)$$

to an equation for W:

$$\frac{\partial}{\partial\zeta}W(\tau,\zeta)+\left(\frac{J}{2}\right)\left[\frac{\partial W(\tau,\zeta)}{\partial\tau}\right]^2+V(\tau)$$
$$-i\hbar\left(\frac{J}{2}\right)\left[\frac{\partial^2 W(\tau,\zeta)}{\partial\tau^2}\right]=0. \tag{3.27}$$

Letting $\hbar\to 0$ in accordance with the requirement of the correspondence principle, Eq. (3.27) becomes identical to (2.15) of the classical mechanical Hamilton–Jacobi theory:

$$H(\tau,p_\tau)+\frac{\partial}{\partial\zeta}W(\tau,\zeta)=0, \tag{3.28}$$

with $W(\tau,\zeta)$ being Hamilton's characteristic function, for which the eikonal equation (2.22) is valid.

In the thermal wave quantum mechanical theory, the eigenvalue equation

$$\hat{H}\psi_n(\tau,\zeta)=E_n\psi_n(\tau,\zeta) \tag{3.29}$$

has the set of eigenfunctions

$$\psi_n(z,\zeta)=N_nD_n(z)\exp(-iE_n\zeta/\hbar), \tag{3.30}$$

where

$$z=(4\omega_0/|\hbar|^2)^{1/4}e^{-i\pi/4}\tau, \tag{3.31}$$

the E_n are given by (3.21), and the N_n are normalization constants, which can be determined using the orthogonality property of the Weber functions[24]:

$$\int_{-\infty}^{\infty}D_m^*(z)D_n(z)dz=(2\pi)^{1/2}n!\delta_{mn}. \tag{3.32}$$

The normalization condition for ψ_n is

$$\int_{-\infty}^{\infty}\psi_n^*(z,\zeta)\psi_n(z,\zeta)dz=1. \tag{3.33}$$

Equations (3.30), (3.32), and (3.33) determine the normalizing constants

$$N_n=[1/(2\pi)^{1/2}n!]^{1/2}. \tag{3.34}$$

IV. EXPECTATION FUNCTIONS AND EHRENFEST'S THEOREMS

The most important application of thermal wave quantum mechanics is its ability to calculate expectation functions for various macroscopic observables, especially those which are difficult or impossible to derive explicitly from the macroscopic Fourier–Helmholtz equation, such as the temperature and the heat flux fields in the medium of thermal wave propagation. In this section we shall derive the expectation values for the potential energy of the THO, expectation functions for the temperature and the heat flux, and a macroscopic heat diffusion equation of the thermal wave center of gravity in the sense and form of Ehrenfest's theorems.

A. Potential energy of THO

We have

$$\langle V(z)\rangle_n=\int_{-\infty}^{\infty}\psi_n^*(z,\zeta)z^2\psi_n(z,\zeta)dz \tag{4.1}$$

$$=N_n^2|G_n(\zeta)|^2\int_{-\infty}^{\infty}z^2D_n^2(z)dz, \tag{4.2}$$

where

$$G_n(\zeta)=\exp(-iE_n\zeta/\hbar). \tag{4.3}$$

Using (3.18a) in (4.2), and the Hermite polynomial identity[28]

$$\int_{-\infty}^{\infty}x^2e^{-x^2}H_n^2(x)dx=\sqrt{x}(2n+1)2^{n-1}n! \tag{4.4}$$

it can be shown that

$$\langle z^2\rangle_n=\langle V(z)\rangle_n=2n+1 \tag{4.5}$$

and, substituting (3.31) in (4.5),

$$\langle V(\tau)\rangle_n=\tfrac{1}{2}(2n+1)(k|\hbar|/2\omega_0^{1/2}e^{-i\pi/2},$$
$$=\tfrac{1}{2}(n+\tfrac{1}{2})\omega_0^{1/2}J|\hbar|$$
$$=\tfrac{1}{2}E_n. \tag{4.6}$$

Equation (4.6) indicates that, for any value of n, the average potential energy is half of the total generalized energy per cycle of oscillation, a result similar to that obtained in the case of the classical mechanical harmonic oscillator.

B. Temperature field

It should be noticed that expectation values are calculated as integrals over the variable z, which is related to the generalized temperature τ via (3.31) and, ultimately, to the temperature T via (2.10). In the present quantum formalism, however, z is considered a dummy variable when under the integral sign, spanning the range of values $(-\infty,+\infty)$. The expectation function for $T(x)$ can be found from

$$\langle z\rangle_{n,m}=\int_{-\infty}^{\infty}\psi_n^*(z,\zeta)z\psi_m(z,\zeta)dz \tag{4.7}$$

$$=(N_nN_m/2^{1/2(n+m)-1})G_n^*(\zeta)G_m(\zeta)$$
$$\times\begin{cases}[\sqrt{\pi}2^n(n+1)]_{m=n+1},\\ [\sqrt{\pi}2^{n-1}n!,]_{m=n-1},\\ 0,\quad m\neq n\pm 1.\end{cases} \tag{4.8}$$

Therefore,

$$\langle z\rangle_{n,m}=\begin{cases}\left(\frac{n+1}{2}\right)^{1/2}G_n^*(\zeta)G_{n+1}(\zeta),\quad m=n+1,\\ (n/2)^{1/2}G_n^*(\zeta)G_{n-1}(\zeta),\quad m=n-1,\\ 0,\quad m\neq n\pm 1.\end{cases} \tag{4.9}$$

For the purpose of obtaining an expression for the temperature field that is consistent with direct solutions[16] to the macroscopic Fourier-Helmholtz equation (2.3) in the limit of constant k, ρ, and c, we choose the particular eigenmodes $n=0$, $m=1$ in (4.9), and we get

$$\langle z\rangle_{0,1}=(1/\sqrt{2})G_0^*(\zeta)G_1(\zeta) \tag{4.10}$$

$$=(1/\sqrt{2})\exp[-e^{i\pi/4}|\Omega_\tau|\zeta], \tag{4.11}$$

and using (3.31), (2.9), and (2.10) in (4.11), the expectation (macroscopic) function for the temperature can be written as

$$T(x)\equiv\langle T(x)\rangle_{0,1}=\frac{|\hbar|^{1/2}e^{-i\pi/4}}{2[\omega_0k(x)\rho(x)c(x)]^{1/4}}$$
$$\times\exp\left[-\int_0^x\sigma_s(\omega_0,x')dx'\right], \tag{4.12}$$

where

$$\sigma_s(\omega_0,x')\equiv(1+i)a_s(\omega_0,x') \tag{4.13}$$

is the complex local thermal diffusion coefficient of the ex-

tended Rosencwaig–Gersho theory.[22] Equation (4.12) is valid for semi-infinite solids, for which the positive root for $\Omega_\tau(\omega_0)$ was rejected in (2.21). This situation reflects the use of thermal wave physics for quantitative analysis leading to depth-profiling at high source modulation frequencies ω_0, so that

$$1/|\sigma_s(\omega_0,L)| \ll L, \tag{4.14}$$

where L is the thickness of the propagation medium. If (4.14) is not satisfied, then the term $\exp[-\int_0^x \sigma_s(\omega_0,x')dx']$ in (4.12) must be replaced by $2\cosh[\int_0^x \sigma_s(\omega_0,x')dx']$, which is the result of the retention of the positive root for $\Omega_\tau(\omega_0)$ in (2.21). Using

$$E_0 = \tfrac{1}{2}|\mathcal{b}|\,|\Omega_\tau| = (1/J)|\Omega_\tau|^2\,\tau_0^2 = Jp_T^2(0) \tag{4.15}$$

for the average total macroscopic generalized energy of the THO in (4.12), with $\tau_0 = \tau(\xi = 0)$, the expectation function for the temperature becomes, after some algebraic manipulation and use of (2.6) and (2.8)–(2.11),

$$T(x) = \frac{Q(x)}{k(x)\sigma_s(\omega_0,x)} \exp\left[-\int_0^x \sigma_s(\omega_0,x')dx'\right], \tag{4.16}$$

where

$$\begin{aligned} Q(x) \equiv Q_0 &\left[\frac{k(x)\rho(x)c(x)}{k(0)\rho(0)c(0)}\right]^{1/4} \\ &\times\left\{1 + \left(\frac{e - i\pi/4}{4\omega_0^{1/2}}\right)\left[\frac{k(0)}{\rho(0)c(0)}\right]^{1/2} \right. \\ &\left.\times\left[\frac{d}{dx}\ln(k\rho c)\right]_x = 0\right\}^{-1}. \end{aligned} \tag{4.17}$$

Equation (4.16) reduces immediately to

$$\begin{aligned} T(x) = &[Q_0/(1+i)k(\omega_0\rho c/2k)^{1/2}] \\ &\times\exp[-(1+i)(\omega_0\rho c/2k)^{1/2}x], \end{aligned} \tag{4.18}$$

in the limit of constant k, ρ, and c. Equation (4.18) has been previously derived by a number of authors.[13,16,29,30]

C. Ehrenfest's theorems

These theorems can be easily formulated upon introducing quantum mechanical commutation relations to replace the Poisson brackets (2.26). For the conjugate variables τ and p_τ the following operator relations can be easily proven:

$$[\hat{\tau},\hat{p}_\tau] = i\mathcal{b}, \tag{4.19a}$$

$$[\hat{\tau},\hat{H}] = \mathcal{b}J\frac{\partial}{\partial\xi}, \tag{4.19b}$$

$$[\hat{p}_\tau,\hat{H}] = -\mathcal{b}\omega_0 J\tau. \tag{4.19c}$$

Ehrenfest's theorem for the generalized thermal momentum can be derived from consideration of the expectation value

$$\frac{d}{d\xi}\langle z\rangle_{n,m} = \frac{d}{\delta\xi}\int_{-\infty}^{\infty}\psi_n^*(z,\xi)z\psi_m(z,\xi)dz \tag{4.20}$$

$$= \frac{i}{\mathcal{b}}\left[\int_{-\infty}^{\infty}\psi_n^*(\hat{H}z - z\hat{H})\psi_m dz\right]$$

$$= (i/\mathcal{b})\langle[\hat{H},z]\rangle_{n,m}. \tag{4.21}$$

Equation (4.21) was obtained under the assumptions[31] that $\hat{H}$ is Hermitian and that z is not an explicit function of ξ. Equations (3.11), (3.13b), (4.19b), and (4.21) give

$$\frac{d}{d\xi}\langle z\rangle_{n,m} = \left(\frac{4\omega_0}{|\mathcal{b}|^2}\right)e^{-i\pi/4}\left\langle\frac{d}{d\xi}\tau(\xi)\right\rangle_{n,m}. \tag{4.22}$$

Using the defining equations (2.9)–(2.11) the term in the brackets can be calculated:

$$\begin{aligned} \left\langle\frac{d}{d\xi}\tau(\xi)\right\rangle_{n,m} &= \frac{1}{F^{1/4}(\xi)}\left[\frac{d}{d\xi}F^{1/4}(\xi)\right]\langle\tau(\xi)\rangle_{n,m} \\ &\quad + [J/F^{1/4}(\xi)]\,\langle p_\tau(\xi)\rangle_{n,m} \\ &= \frac{d}{d\xi}\langle\tau(\xi)\rangle_{n,m}, \end{aligned} \tag{4.23}$$

where

$$F(\xi) \equiv k(\xi)\rho(\xi)c(\xi). \tag{4.24}$$

The equality of expectation values,

$$\left\langle\frac{d}{d\xi}\tau(\xi)\right\rangle_{n,m} = \frac{d}{d\xi}\langle\tau(\xi)\rangle_{n,m}, \tag{4.25}$$

originates in the fact that the potential field (3.6) for the THO is harmonic and does not involve terms higher than second order.[32] Equations (4.22) and (4.23) yield the following thermal wave Ehrenfest equation, which relates the expectation values of the generalized thermal velocity $(d/d\xi)\langle\tau\rangle$ and the generalized thermal momentum $\langle p_\tau\rangle$:

$$\begin{aligned} \frac{d}{d\xi}\langle\tau(\xi)\rangle_{n,m} = \frac{1}{F^{1/4}(\xi)}&\left[J\langle p_\tau(\xi)\rangle_{n,m}\right. \\ &\left.+\left(\frac{d}{d\xi}F^{1/4}(\xi)\right)\langle\tau(\xi)\rangle_{n,m}\right]. \end{aligned} \tag{4.26}$$

For $F = \text{const}$, (4.26) reduces to

$$\langle p_T\rangle_{n,m} = k\frac{d}{dx}\langle T\rangle_{n,m}, \tag{4.27}$$

in agreement with (2.8).

Furthermore, an equation for the motion of the observable thermal ray center of gravity can be derived in the form of an Ehrenfest relation

$$\left\langle\frac{d^2}{d\xi^2}\tau(\xi)\right\rangle_{n,m} = \frac{d^2}{d\xi^2}\langle\tau(\xi)\rangle_{n,m} = J\frac{d}{d\xi}\langle\Pi_\tau(\xi)\rangle_{n,m}, \tag{4.28}$$

where $\Pi_\tau(\xi)$ is the effective generalized thermal momentum in a medium with variable thermal/thermodynamic parameters $k(x)$,$\rho(x)$, and $c(x)$:

$$\Pi_\tau(\xi) \equiv \frac{1}{F^{1/4}(\xi)}\left[p_\tau(\xi) + \frac{1}{J}\left(\frac{d}{d\xi}F^{1/4}(\xi)\right)\tau(\xi)\right]. \tag{4.29}$$

Ehrenfest's theorem (4.26) can now be written in a suggestive form as

$$\frac{d}{d\xi}\langle\tau\rangle_{n,m} = \frac{1}{J^{-1}}\langle\Pi_\tau\rangle_{n,m}, \tag{4.30}$$

where J^{-1} plays the role of a generalized mass of the system, in agreement with (2.17). Differentiating $\Pi_\tau(\xi)$ with respect to ξ and inserting the resulting expression in (4.28) gives

$$\begin{aligned} &\frac{d^2}{d\xi^2}\langle\tau(\xi)\rangle_{n,m} \\ &\quad + \left[\Omega_\tau^2 - \frac{1}{F^{1/4}(\xi)}\frac{d^2}{d\xi^2}F^{1/4}(\xi)\right]\langle\tau(\xi)\rangle_{n,m} = 0. \end{aligned} \tag{4.31}$$

Equation (4.31) is Ehrenfest's theorem, which governs the motion of the thermal ray heat centroid in the presence of a harmonic generalized potential energy field $V(\tau)$ and for general functional forms of $F(\zeta)$. An equation similar to (4.31) in structure has been derived in connection with the eigenvalue problem of the Liouville equation by Morse and Feshbach [Ref. (26), Eq. (6.3.22)]. The analogy of the Ehrenfest approach to the classical mechanical theory became apparent with (4.30), which involves generalized thermal displacement and momentum. This analogy becomes complete once (4.28) is written in terms of an integral over eigenfunctions:

$$\frac{d^2}{d\zeta^2}\langle\tau\rangle_{n,m} = \frac{e^{i\pi/4}}{(4\omega_0/|\mathcal{b}|^2)^{1/4}}\frac{d^2}{d\zeta^2}\int_{-\infty}^{\infty}\psi_n^*(z,\zeta)\,z\psi_m(z,\zeta)\,dz$$
$$= \frac{i}{\mathcal{b}}\left[\int_{-\infty}^{\infty}\psi_n^*(\hat{H}\hat{\Pi}_\tau - \hat{\Pi}_\tau\hat{H})\,\psi_m\,dz\right]$$
$$= (i/\mathcal{b})\langle[\hat{H},\hat{\Pi}_\tau]\rangle_{n,m}. \tag{4.32}$$

To evaluate this commutator, the commutation relation

$$\hat{\Pi}_\tau f(\tau) - f(\tau)\hat{\Pi}_\tau = -i\mathcal{b}\frac{\partial f(\tau)}{\partial\tau} \tag{4.33}$$

can be easily verified from the definition (3.2) of the generalized thermal momentum operator $\hat{p}_\tau$. Using (4.33) in (4.32) and comparing with (4.28) yields the following form of Ehrenfest's theorem (4.31):

$$\frac{d}{d\zeta}\langle\Pi_\tau\rangle_{n,m} = \left\langle -\frac{\partial U}{\partial\tau}\right\rangle_{n,m}, \tag{4.34}$$

where U is the effective potential energy

$$U(\tau) = \tfrac{1}{2}K_{\text{eff}}\tau^2, \tag{4.35}$$

with the generalized effective spring constant

$$K_{\text{eff}}(\zeta) \equiv K - \frac{1}{JF^{1/4}(\zeta)}\frac{d^2}{d\zeta^2}F^{1/4}(\zeta). \tag{4.36}$$

Equation (4.34) is Ehrenfest's approach equivalent to the classical relationship between a restoring force, $-\nabla_\tau U$, and the rate of change of momentum due to the action of the force. The equations of motion (4.30) and (4.34) of the heat centroid of the thermal rays demonstrate the correspondence between macroscopic heat conduction theory and the expectation values of our microscopic approach.

D. Position of heat centroid and uncertainty principle

For experimental purposes the position of the heat centroid at a given modulation frequency ω_0 is extremely important, as its determination substantiates the depth-profiling capacity of thermal wave physics. The observed (measured) wave centroid is mathematically a well-defined quantity for thermal waves, unlike unattenuated electromagnetic and other plane waves, due to the damped nature of the former:

$$\langle\zeta\rangle_{0,1} = \frac{\int_0^\infty \zeta\, G_0^*(\zeta)G_1(\zeta)d\zeta}{\int_0^\infty G_0^*(\zeta)G_1(\zeta)d\zeta} \tag{4.37}$$
$$= (\sqrt{2}/1+i)\,(|\Omega_\tau(\omega_0)|)^{-1},$$

so that

$$|\langle\zeta\rangle_{0,1}| = \frac{L}{\omega_0^{1/2}\int_0^L dx/\alpha_s^{1/2}(x)}, \quad \alpha_s(x) = \frac{k(x)}{\rho(x)c(x)}, \tag{4.38}$$

where $\alpha_s(x)$ is the local thermal diffusivity of the propagation medium.[22] If the medium is semi-infinite, (4.38) is understood to mean

$$\langle\zeta\rangle_{0,1} = \left(\frac{1}{\omega_0^{1/2}}\right)\lim_{L\to\infty}\left(\frac{L}{\int_0^L dx/\alpha_s^{1/2}(x)}\right). \tag{4.38'}$$

A combination of (4.38) or (4.38') and (4.16), together with a measurement of $T(x)$ or of a quantity proportional to $T(x)$, demonstrates the potential of our quantum theory for quantitative analysis of depth-profiling studies through thermal wave physics in media of arbitrarily variable $\alpha_s(x)$. This aspect of the present theory will be examined in detail in a future publication.

For completeness of the thermal wave quantum treatment an uncertainty principle with an interesting interpretation will be established. By analogy to quantum mechanics,

$$(\Delta\tau)_{n,n+1} \equiv [\langle\tau^2\rangle_{n,n+1} - \langle\tau\rangle_{n,n+1}^2]^{1/2}, \tag{4.39}$$

$$(\Delta p_\tau)_{n,n+1} \equiv [\langle p_\tau^2\rangle_{n,n+1} - \langle p_\tau\rangle_{n,n+1}^2]^{1/2}. \tag{4.40}$$

The expressions involve integrals of the Weber functions with powers or derivatives of τ (or z). They can be simplified upon noticing that

$$\langle z^2\rangle_{n,n+1} \propto \int_{-\infty}^{\infty} z^2 D_n^*(z)D_{n+1}(z)\,dz = 0 \tag{4.41}$$

and

$$\langle p_\tau^2\rangle_{n,n+1}$$
$$\propto \langle p_z^2\rangle_{n,n+1} \propto -\mathcal{b}^2\int_{-\infty}^{\infty} D_n^*(z)\left[\frac{d^2}{dz^2}D_{n+1}(z)\right]dz$$
$$= -\mathcal{b}^2\left\{(n+1)n\int_{-\infty}^{\infty} D_n^*(z)D_{n-1}(z)dz\right.$$
$$-(n+1)\int_{-\infty}^{\infty} z\,D_n^*(z)D_n(z)\,dz$$
$$-\frac{1}{2}\int_{-\infty}^{\infty} D_n^*(z)D_{n+1}(z)dz$$
$$\left.+\frac{1}{4}\int_{-\infty}^{\infty} z^2 D_n^*(z)D_{n+1}(z)dz\right\} = 0, \tag{4.42}$$

as each and every integral inside the brackets can be shown to vanish due to the orthogonality of the Weber polynomials.[24] Taking (4.41) and (4.42) into account, (4.39) and (4.40) give

$$(\Delta\tau)_{n,n+1}(\Delta p_\tau)_{n,n+1} = (\langle\tau\rangle_{n,n+1})(\langle p_\tau\rangle_{n,n+1}), \tag{4.43}$$

where

$$\langle\tau\rangle_{n,n+1} = \frac{e^{i\pi/4}}{(4\omega_0/|\mathcal{b}|^2)^{1/4}}\left(\frac{n+1}{2}\right)^{1/2} G_n^*(\zeta)G_{n+1}(\zeta) \tag{4.44}$$

and

$$\langle p_\tau\rangle_{n,n+1} = [(4\omega_0/|\mathcal{b}|^2)^{1/4}/e^{i\pi/4}]\langle p_z\rangle_{n,n+1} \tag{4.45}$$

and

$$\langle p_z\rangle_{n,n+1} = -i\mathcal{b}\int_{-\infty}^{\infty}\psi_n^*(z,\zeta)\frac{\partial}{\partial z}\psi_m(z,\zeta)dz$$
$$= -i\mathcal{b}\,N_nN_{n+1}\,G_n^*(\zeta)G_{n+1}(\zeta)$$
$$\times\int_{-\infty}^{\infty} D_n(z)\left[\frac{d}{dz}D_{n+1}(z)\right]dz.$$

It can be shown that

$$\int_{-\infty}^{\infty} D_n(z)\left[\frac{d}{dz}D_{n+1}(z)\right]dz=\left(\frac{\pi}{2}\right)^{1/2}(n+1)! \tag{4.46}$$

so that

$$\langle p_z\rangle_{n,n+1}=-(i\check{b}/2)(n+1)^{1/2}\exp\left[-e^{i\pi/4}|\Omega_\tau|\zeta\right]$$

and

$$\langle p_\tau\rangle_{n,n+1}=-\left[i\check{b}\,(4\omega_0/|\acute{b}|^2)^{1/4}/2e^{i\pi/4}\right]\times(n+1)^{1/2}\exp\left[-e^{i\pi/4}|\Omega_\tau|\zeta\right]. \tag{4.47}$$

Equations (4.43), (4.44), and (4.47) yield the uncertainty principle

$$(\Delta\tau)_{n,n+1}(\Delta p_\tau)_{n,n+1}=\frac{1}{2}(n+1)\check{b}\exp\left[-2\int_0^x\sigma_s(\omega_0x')dx'\right], \tag{4.48}$$

which indicates that the uncertainty in either the thermal wave temperature excursion or its momentum decreases with increasing modulation frequency ω_0. The form of the uncertainty principle is unlike that of ordinary quantum mechanics[23] due to the fact that the quantal thermal wave formalism requires coupling between the nth and $(n+1)$th eigenmodes to produce temperature expectation functions with the correct limiting forms. If coupling were considered entirely within the nth eigenmode, then (i) the relations $\langle\tau\rangle_{n,n}=\langle p_\tau\rangle_{n,n}=0$ would be true, also familiar from the quantum mechanical harmonic oscillator theory ($\langle x\rangle=\langle p\rangle=0$)[23]; and (ii) a more typical uncertainty relation would be obtained:

$$(\Delta\tau)_{n,n}\,(\Delta p_\tau)_{n,n}=(n+\tfrac{1}{2})\check{b}. \tag{4.49}$$

The thermal-wave uncertainty relation (4.48) is important in that it sets a lower limit in the precision with which the temperature of the heat centroid of the thermal wave packet can be measured, when the thermal flux p_τ is known with a precision $\Delta p_\tau\equiv p_0$:

$$|\Delta T(x)|\geqslant\frac{b}{4\pi p_0F^{1/4}(x)}\exp\left[-(2\omega_0)^{1/2}\int_0^x\frac{dx'}{\alpha_s^{1/2}(x')}\right]. \tag{4.50}$$

This inequality is a mathematical statement for the maximum depth resolution with thermal waves in a medium in which temperature is modulated at ω_0. Equation (4.48) implies the spread of thermal waves due to diffraction, as the wave packet travels away from the surface and into the medium. The rapid decrease in the uncertainty minimum with increasing ω_0 indicates a more precise thermal imaging and information transfer from subsurface features at high frequencies due to decreased diffraction limitations. This theoretical observation has been borne out in several experiments.[1,4,9,10]

V. CONCLUSIONS

The thermal wave quantum mechanical formalism developed in this work as the extension of the thermal-wave Hamilton–Jacobi theory has been shown to be capable of providing exact analytical expressions for the temperature distribution and heat flux for general solids with continuously variable thermal/thermodynamic parameters. These macroscopic expressions are in the form of Ehrenfest-type expectation functions and are expected to be useful to depth profiling analysis in solids with rapidly varying thermal parameters locally, especially close to the surface, for instance, to microelectronic processing (e.g., impurity doping, ion implantation, radiation damage).

ACKNOWLEDGMENTS

Thanks are due to Dr. L. C. M. Miranda for bringing to the author's attention the early manuscript by Arnold and Crandall.[30]

The author also wishes to acknowledge the financial support of the National Science and Engineering Council of Canada, which made this research possible.

[1] A. Rosencwaig and G. Busse, Appl. Phys. Lett. **36**, 725 (1980).
[2] E. Brandis and A. Rosencwaig, Appl. Phys. Lett. **37**, 98 (1980).
[3] R. L. Thomas, J. J. Pouch, Y. H. Wong, L. D. Favro, P. K. Kuo, and A. Rosencwaig, J. Appl. Phys. **51**, 1152 (1980).
[4] G. Busse and A. Rosencwaig, Appl. Phys. Lett. **36**, 815 (1980).
[5] T. Sawada, H. Shimizu, and S. Oda, Jpn. J. Appl. Phys. **20**, L25 (1981).
[6] A. Rosencwaig, J. Appl. Phys. **51**, 2210 (1980).
[7] A. Rosencwaig, Thin Solid Films **77**, L43 (1981).
[8] A. Mandelis and J. D. Lymer, Appl. Spectrosc. **39**, 473 (1985).
[9] A. Rosencwaig and R. M. White, Appl. Phys. Lett. **38**, 165 (1981).
[10] A. Rosencwaig, Solid State Technol. **25**, 91 (March 1982).
[11] A. Rosencwaig, J. Photoacoust. **1**, 371 (1983).
[12] A. Rosencwaig, J. Opsal, and D. L. Willenborg, Appl. Phys. Lett. **43**, 166 (1983).
[13] J. Opsal and A. Rosencwaig, J. Appl. Phys. **53**, 4240 (1982).
[14] S. D. Campbell, S. S. Yee, and M. A. Afromowitz, IEEE Trans. Biomed. Engn. **BME-26**, 220 (1979).
[15] M. A. Afromowitz, S. P. Yeh, and S. S. Yee, J. Appl. Phys. **48**, 209 (1977).
[16] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids* (Clarendon, Oxford, 1960), 2nd ed., Sec. 2.6.
[17] D. Gloge and D. Marcuse, J. Opt. Soc. Am. **59**, 1629 (1969).
[18] G. Eichmann, J. Opt. Soc. Am. **61**, 161 (1971).
[19] J. C. Murphy and L. C. Aamodt, Appl. Phys. Lett. **31**, 728 (1977).
[20] B. A. Findlayson, *The Method of Weighted Residuals and Variational Principles* (Academic, New York, 1972), Chap. 7.
[21] H. Goldstein, *Classical Mechanics* (Addison–Wesley, Reading, MA, 1965), Chap. 9.
[22] A. Rosencwaig and A. Gersho, J. Appl. Phys. **47**, 64 (1976).
[23] L. I. Schiff, *Quantum Mechanics* (McGraw–Hill, New York, 1968), Chap. 4.
[24] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge U. P., Cambridge, 1963), 4th ed., Sec. 16.5.
[25] I. Ye. Kireyeva and K. A. Karpov, *Tables of Weber Functions*, translated by P. Basu (Pergamon, New York, 1961).
[26] P. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw–Hill, New York, 1953), Sec. 5.3.
[27] A/ Yariv, *Quantum Electronics* (Wiley, New York, 1975), 2nd ed., Sec. 6.5.
[28] *Bateman Manuscript Project, Higher Transcendental Functions*, edited by A. Erdelyi (McGraw–Hill, New York, 1953), Vol. 2, Sec. 10.13.
[29] A. J. Angstrom, Ann. Phys. Lpz. **114**, 513 (1861).
[30] H. D. Arnold and I. B. Crandall, Phys. Rev. **10**, 22 (1917).
[31] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961) Vol. I, Chap. V.
[32] D. Marcuse, *Light Transmission Optics* (Van Nostrand, New York, 1982), Sec. 3.6.

Fermi states of Bose systems in three space dimensions. II. The solvable version of the Hubbard model

Piotr Garbaczewski
Institute of Theoretical Physics, University of Wrocław, Cybulskiego 36, 50-205 Wrocław, Poland

(Received 21 November 1984; accepted for publication 15 March 1985)

An exact spectral solution of the Hubbard problem on the tetrahedron is given. It is demonstrated that it provides at the same time a partial spectral solution for the Bose version of this (Fermi) model. A remarkable relationship of the Hamiltonians $H_B = H_F + (1-P)H_B(1-P)$ is established, where P is a projection in the state space of the interacting Bose system. This partial observation is then extended to the spatial lattice composed of the four-site Hubbard cells, which interact due to the superimposed perturbation of the hopping type, affecting all sites of the lattice; $H = \Sigma_{(ij)} \Sigma_{\sigma=\pm} b_{ij} c^*_{i\sigma} c_{j\sigma} + U \Sigma_i n_{i+} n_{i-}$ being the corresponding Hamiltonian. It is thus demonstrated that the joint Bose–Fermi spectral problem (previously established in $1+1$ dimensions) makes sense in $1+3$ dimensions as well.

I. MOTIVATION

Models are the necessary artifice in the study of any realistic physical problem. In both high-energy and many-body physics, the phenomenological analysis leads to model Hamiltonian systems, which are used to approximate reality. If one is happy to solve any such model or to solve at least its somehow simplified version, there is always a hope that some amount of knowledge about reality has been gained.

With respect to quantum field theory on a lattice and in continuum, modeling has proved successful in $1+1$ (and $1+2$) space-time dimensions. The situation is dramatic in $1+3$ dimensions, where the variety of solvable (at least perturbatively) models is severely limited. Therefore it seems advantageous either to propose a new solvable model in three space dimensions, or to demonstrate a solvability of any known model.

Since the ultraviolet and infrared cutoffs are the standard regularizations of any reasonable quantum field theory, it is not unnatural to confine our attention to the lattice models in three space dimensions.

The aim of the present paper is the direct solvability proof for a certain version of the well-known many-body problem, the so-called Hubbard model.[1–5] This particular goal arises as a by-product of our previous studies of the relationships between the quantized Bose and Fermi systems; see, e.g., Refs. 6–11 and references therein. Though we were mostly confined to $1+1$ dimensions, the $(1+3)$-dimensional case was investigated as well; see Refs. 8–11. The present paper can be viewed as a continuation of the paper[8] referred to as I in the present series. Previously we have accomplished the study of Fermi (state space) sectors in the Hilbert space of the Bose system, for the family of the Lee-type models; the pure Fermi model is known there to be an exactly solvable example.

The Bose–Fermi relationship we have in mind can be characterized as follows. Given the Hamiltonian Bose system H_B with a certain projection P in its state space, $\mathscr{H}_B : \mathscr{H}_F = P\mathscr{H}_B$, which commutes with H_B. If this projection coincides with the unit operator in the "bosonized" Fermi (CAR) algebra of Refs. 10 and 11, then $\mathscr{H}_F$ is exactly the (Hilbert) space of Fermi states of the Bose system H_B. Since H_B acts invariantly on $\mathscr{H}_F$ its restriction to $\mathscr{H}_F$, denoted PH_BP, can be used instead of H_B itself. For the Bose models we have studied before, the following exact equivalence on $\mathscr{H}_B$ has been established:

$$PH_BP = H_F, \tag{1.1}$$

where H_F is received from H_B through a simple replacement of canonical Bose variables (fields) by Fermi ones.

The situation became more involved once we passed to the solvable (fermionic Lee) model in $1+3$ dimensions.[8] Then the object PH_BP can be equivalently described by the pure fermionic operator

$$PH_BP = G_F, \tag{1.2}$$

but G_F is formally non-Hermitian. An analogous effect appears if we invert the procedure, and depart from the pure fermionic Lee model H_F. We have proved that the Bose–Fermi relationship of the type (1.1) exists, but then

$$H_F = PG_BP, \tag{1.3}$$

where the Bose operator G_B is formally non-Hermitian, and hence cannot play the role of the standard (physical) Hamiltonian. In the case of the Lee-type models it means that one can consistently introduce Fermi fields within the Bose model, but no reasonable physical Hamiltonian system can be attributed to them.

Since (1.1) was characteristic for Bose systems in $1+1$ dimensions, while features (1.2) and (1.3) were displayed in $1+3$, our aim is to verify whether (1.1) is merely the $(1+1)$-dimensional peculiarity, or rather a general feature of some Bose models independent of the space-time dimension. Our answer, as given in the present paper, is in favor of the second alternative. The problem of correlated electron motions has attracted the attention of solid state physicists for a long time, culminating in the so-called Hubbard model[1] whose simplified version of Ref. 2 is described by the Hamiltonian on the (one-, two-, or three-dimensional spatial) lattice

$$H = H_F = \sum_k E_k(n_{k+} + n_{k-}) + U\sum_g n_{g+} n_{g-}, \tag{1.4}$$

where n_k is an electron number operator in the momentum space, an arrow indicating the spin-up or down cases.

 0022-2488/85/102684-09$02.50

The second term in H describes the repulsion between two electrons of opposite spin which happen to be in the same orbit around a particular lattice site; g is thus a configuration space index, and

$$n_{g\sigma} = c^*_{g\sigma}\, c_{g\sigma},\quad [c_{g\sigma}, c^*_{h\tau}]_+ = \delta_{\sigma\tau}\,\delta_{gh}, \tag{1.5}$$

$$c^*_{k\sigma} = \frac{1}{\sqrt{N}} \sum_g \exp(ikg)\cdot c^*_{g\sigma}.$$

As is well known, the exact solution (in the sense of solving the spectral problem for H) has been produced in one space dimension only.[3]

The obvious difficulty in (1.4) is that the Hamiltonian consists of the two parts: the momentum and the configuration ones. Each of them is separately solvable, but a simultaneous diagonalization of them is not that trivial, especially if one, following the solid-state physics tradition, translates the whole problem to the momentum space language; see, e.g., Ref. 4. Even in Ref. 3 where the original configuration space formula is reproduced for the 1 + 1 dimensional model

$$H = t \sum_{(ij)} \sum_\sigma c^*_{i\sigma}\, c_{j\sigma} + U \sum_i c^*_{i+}\, c_{i+}\, c^*_{i-}\, c_{i-}, \tag{1.6}$$

the use of the Bethe ansatz route leads to the (quasi) momentum space picture again.

We shall be interested in the less restrictive than (1.4) version of the Hubbard model, and admit[5]

$$H = \sum_{(ij)} \sum_\sigma b_{ij}\, c^*_{i\sigma}\, c_{j\sigma} + U \sum_i n_{i+}\, n_{i-}, \tag{1.7}$$

where the specification of hopping amplitudes b_{ij} will be made in the course of the paper.

II. THE HUBBARD MODEL ON THE TETRAHEDRON

We shall begin from the study of the Hubbard model on the simplest example of the three-dimensional spatial lattice. The underlying structure is the tetrahedron, whose sites we enumerate 0,1,2,3. The nearest neighbors form pairs $(i,j) = (0,1)$, $(0,2)$, (03), $(1,2)$, (13), $(2,3)$. The Hamiltonian reads

$$c^\#_{i+} \doteq a^\#_i,\quad c^\#_{i-} \doteq \tilde{a}^\#_i,$$

$$H = H_F = t \sum_{(ij)} (a^*_i\, a_j + \tilde{a}^*_i\, \tilde{a}_j) + U \sum_i n_i \cdot \tilde{n}_i, \tag{2.1}$$

$$\sum_{(ij)} a^*_i\, a_j = \sum_{i=1}^{3} (a^*_0 a_i + a^*_i\, a_0) + \sum_{k=2}^{3} (a^*_1\, a_k + a^*_k\, a_1) + (a^*_2\, a_3 + a^*_3\, a_2).$$

The particle number operators are here the constants of motion

$$M = \sum_k a^*_k\, a_k,\quad \tilde{M} = \sum_k \tilde{a}^*_k\, \tilde{a}_k,\quad N = M + \tilde{M}, \tag{2.2}$$

by virtue of

$$[a^\#_i, a^\#_j]_+ = 0,$$

$$[a_i, a^*_j]_+ = \delta_{ij} = [\tilde{a}_i, \tilde{a}^*_j]_+, \tag{2.3}$$

other anticommutators vanishing. We obviously admit the existence of the Fock vacuum $|0)$

$$a_j|0) = 0 = \tilde{a}_j|0). \tag{2.4}$$

If, following Ref. 4 we restrict considerations to the sectors $N = 4$, $M + \tilde{M} = N$, then one refers to the so-called half-filled band case of the model. We shall, however, not insist on this restriction, and our aim is to demonstrate that the particular four-site (tetrahedron) Hubbard problem is calculable in all its N, M, $\tilde{M}$ sectors.

A. The $M = \tilde{M} = 1$ sector

The basis system in this case reads

$$|i{:}j) = a^*_i\, \tilde{a}^*_j|0), \tag{2.5}$$

and the action of the pure hopping, (a^*,a) part of H, which we denote H_{hop}, reads

$$\begin{aligned} &t^{-1}|0{:}j) \to |1{:}j) + |2{:}j) + |3{:}j),\\ &t^{-1}|1{:}j) \to |0{:}j) + |2{:}j) + |3{:}j),\\ &t^{-1}|2{:}j) \to |0{:}j) + |1{:}j) + |3{:}j),\\ &t^{-1}|3{:}j) \to |0{:}j) + |1{:}j) + |2{:}j), \end{aligned} \tag{2.6}$$

and analogously for $\tilde{H}_{\text{hop}}$. It is not that useless to mention that the eigenvectors for the pure hopping (i.e., $U = 0$) case of (2.1) arise immediately by observing that (2.6) implies, for example,

$$H_{\text{hop}}\{|k{:}j) - |l{:}j)\} = -t\{|k{:}j) - |l{:}j)\}, \tag{2.7}$$

for all $k,l, = 0,1,2,3$.

For any vector $|\psi\rangle$ in the $M = \tilde{M} = 1$ sector we have

$$\begin{aligned} &|\psi) = \sum_{kj} \alpha_{kj}|k{:}j),\\ &H_{\text{hop}}|\psi) = t \sum_{ki} \alpha_{kj} \sum_{s\neq k} |s{:}j) = t \sum_{kj} \left\{\sum_l \alpha_{lj} - \alpha_{kj}\right\}|k{:}j),\\ &\tilde{H}_{\text{hop}}|\psi) = t \sum_{kj} \left\{\sum_l \alpha_{kl} - \alpha_{kj}\right\}|k{:}j),\\ &V|\psi) = U \sum_k \alpha_{kk}|k{:}k),\\ &H = H_{\text{hop}} + \tilde{H}_{\text{hop}} + V. \end{aligned} \tag{2.8}$$

It amounts to the following eigenvalue problem:

$$H|\psi) = E|\psi)$$

$$\Rightarrow \begin{cases} k = j, & t \sum_{l\neq k} (\alpha_{lk} + \alpha_{kl}) + (U - E)\alpha_{kk} = 0,\\ k \neq j, & t \sum_{l\neq k} \alpha_{lj} + t \sum_{l\neq j} \alpha_{kl} - E\alpha_{kj} = 0. \end{cases} \tag{2.9}$$

Such a homogeneous system of 16 linear equations is in principle solvable but the straight algebraic route is not economic. We shall use the ordering $(i,0)$, $(i,1)$, $(i,2)$, $(i,3)$, $i = 0,1,2,3$, to count the α's, which implies the following block form of the matrix of the system (2.9):

$$\mathscr{A} = \begin{bmatrix} A_1 & tI & tI & tI\\ tI & A_2 & tI & tI\\ tI & tI & A_3 & tI\\ tI & tI & tI & A_4 \end{bmatrix}, \tag{2.10}$$

where each of the 4×4 entries is either proportional to a unit

matrix I, or can be deduced from

$$A_1=\begin{bmatrix} U-E & t & t & t \\ t & -E & t & t \\ t & t & -E & t \\ t & t & t & -E \end{bmatrix}, \tag{2.11}$$

since A_i differs from A_1 by permuting the $(U-E)$ element along the diagonal to the ith place.

Due to the particular form of (2.10), we can look for the coefficient representative $|\alpha)$ of the vector $|\psi)$ given as a direct sum of the four-dimensional vectors, so that

$$\begin{aligned} &|\alpha)=\sum_{i=1}^{4} a_i, \quad a_1=(a,0,0,0), \\ &a_2=(0,b,0,0), \quad a_3=(0,0,c,0), \\ &a_4=(0,0,0,d), \quad a=(a_{11},a_{12},a_{13},a_{14}),\ldots, \end{aligned} \tag{2.12}$$

and consequently

$$\mathscr{A}|\alpha)=0$$

$$\Rightarrow\begin{cases} A_1 a+t(b+c+d)=0, \\ ta+A_2 b+t(c+d)=0, \\ t(a+b)+A_3 c+td=0, \\ t(a+b+c)+A_4 d=0, \end{cases} \tag{2.13}$$

which furthermore implies the identities

$$(A_1-tI)\ a=(A_2-tI)\ b=(A_3-tI)\ c=(A_4-tI)d, \tag{2.14}$$

and hence

$$a+b+c+d=0. \tag{2.15}$$

If we observe that

$$\det(A_1-tI)=\det(A_i-tI), \quad \forall\, i=2,3,4, \tag{2.16}$$

then it is natural to demand

$$\det(A_1-tI)=0, \tag{2.17}$$

which immediately leads to

$$(E+2t)^2(E^2-UE+t(U-4t))=0. \tag{2.18}$$

It provides us with the spectral solution for our problem

$$E_0=-2t, \tag{2.19}$$

and

$$\begin{aligned} &E^2-UE+t(U-4t)=0, \\ &\Delta=U^2-4t(U-4t) \\ &\quad\Rightarrow E_{\pm}=\tfrac{1}{2}(U\pm\Delta^{1/2}). \end{aligned} \tag{2.20}$$

Let us observe that since $\Delta>0$, then if either U or t differs from 0, the spectrum consists of the three real eigenvalues except for the so-called band ($t=0$) and atomic ($U=0$) limits of the Hubbard model[4] when

$$\begin{aligned} &t\to 0\Rightarrow E_+=U, \quad E_0=E_-=0, \\ &U\to 0\Rightarrow E_+=2t, \quad E_0=E_-=-2t. \end{aligned} \tag{2.21}$$

The eigenvectors of matrices (A_i-tI), $i=1,2,3,4$ can be easily found by means of the Cramer formulas. To obtain the eigenvector $|\alpha)$ one must satisfy the linear dependence condition (2.15). But it is in fact an automatic consequence of (2.17). For example, the three eigenvectors of (A_1-tI) have the form

$$\begin{aligned} &\alpha(x,y,y,1), \quad \beta(x,y,1,y), \quad \gamma(x,1,y,y), \\ &x=-\frac{t(E+2t)}{(U-E-t)E+2t^2}, \\ &y=x+\frac{Ut}{(U-E-t)E+2t^2}, \end{aligned} \tag{2.22}$$

α,β,γ being arbitrary real numbers.

B. The $N=4$, $M=1$, $\tilde{M}=3$ sector

We can here exploit (2.6). The basis system reads

$$|i;jkl)=a_i^*\,\tilde{a}_j^*\,\tilde{a}_k^*\,\tilde{a}_l^*|0), \tag{2.23}$$

and it is enough to consider vectors $|012)$, $|013)$, $|023)$, $|123)$ of the tilde contribution, which under the action of $t^{-1}\tilde{H}_{\rm hop}$ transform as follows:

$$\begin{aligned} &|123)\to|023)-|013)+|012), \\ &|012)\to|123)-|023)+|013), \\ &|023)\to|123)+|013)-|012), \\ &|013)\to-|123)+|023)+|012). \end{aligned} \tag{2.24}$$

If we relabel the basis vectors according to

$$\begin{aligned} &|012)\doteq|0), \quad |013)\doteq|1), \\ &|023)\doteq|2), \quad |123)\doteq|3), \end{aligned} \tag{2.25}$$

we find that

$$t^{-1}\tilde{H}_{\rm hop}:\begin{cases} |0)\to|3)-|2)+|1), \\ |1)\to-|3)+|2)+|0), \\ |2)\to|3)+|1)-|0), \\ |3)\to|2)-|1)+|0). \end{cases} \tag{2.26}$$

(Notice that, e.g., $[|0)+|1)]$ is an eigenvector of $\tilde{H}_{\rm hop}$ with the eigenvalue t.) If compared with the action of $t^{-1}H_{\rm hop}$ on the $|k;j)$ basis [see (2.6)] we encounter the change of sign in the even (odd) contribution when the even (odd) vector is acted upon by $t^{-1}\tilde{H}_{\rm hop}$.

The $M=1$, $\tilde{M}=3$ version of (2.6) thus reads

$$\begin{aligned} &|\psi)=\sum_{kj}\alpha_{kj}|k;j), \\ &\tilde{H}_{\rm hop}|\psi)=t\sum_{kj}\left\{\sum_l \alpha_{kl}(-1)^{j-l+1}+\alpha_{kj}\right\}|k;j), \\ &H_{\rm hop}|\psi)=t\sum_{kj}\left\{\sum_l \alpha_{lj}-\alpha_{kj}\right\}|k;j), \end{aligned} \tag{2.27}$$

while the action of V leads to

$$\begin{aligned} V|\psi)=U|\psi)-U\{&\alpha_{03}|0:3)+\alpha_{12}|1:2) \\ &+\alpha_{21}|2:1)+\alpha_{30}|3:0)\}. \end{aligned} \tag{2.28}$$

Consequently

$$\begin{aligned} H|\psi)=E|\psi)\Rightarrow t\sum_l[\alpha_{kl}(-1)^{j-l+1}+\alpha_{lj}]+(U-E)\alpha_{kj} \\ -U(\delta_{k0}\,\delta_{j3}+\delta_{k1}\,\delta_{j2}+\delta_{k2}\,\delta_{j1} \\ +\delta_{k3}\,\delta_{j0})\alpha_{kj}=0, \end{aligned} \tag{2.29}$$

which is the homogeneous system of 16 linear equations with respect to the coefficient representative $|\alpha)$ of the vector $|\psi)$

$$\mathscr{A}|\alpha)=0. \tag{2.30}$$

If we adopt the ordering previously used, the matrix $\mathscr{A}$ has

again the form (2.10) but the 4×4 matrices on its (block) diagonal are different from (2.11)

$$A_1=\begin{bmatrix} U-E & t & -t & t \\ t & U-E & t & -t \\ -t & t & U-E & t \\ t & -t & t & U-E \end{bmatrix}. \tag{2.31}$$

Here, A_i, $i>1$ differs from A_1 in that the diagonal element $(-E)$ is permuted from the fourth to the $[(4-i)+1]$th position on the diagonal.

All the arguments used in Sec. II A apply to the present case as well, and we arrive again at the demand

$$\det(A_1 - tI)=0, \tag{2.32}$$

implying

$$(U-E)^2[E^2+E(4t-U)-Ut]=0, \tag{2.33}$$

i.e.,

$$E_0=U,$$

or

$$E^2+E(4t-U)-Ut=0. \tag{2.34}$$

Upon

$$\Delta=16t^2-4Ut+U^2=12t^2+(2t-U)^2>0, \tag{2.35}$$

we get

$$E_{\pm}=\tfrac{1}{2}(U-4t\pm\Delta^{1/2}). \tag{2.36}$$

Let us also observe that $t\to0\Rightarrow E_0=U$, $E_+=U$, $E_-=0$, $U\to0\Rightarrow E_0=E_+=0$, $E_-=-4t$.

C. The $M=\tilde{M}=2$ sector

The basis system in this case reads

$$|ij{:}kl)=a_i^*\,a_j^*\,\tilde{a}_k^*\,\tilde{a}_l^*|0), \tag{2.37}$$

and the complicacy of the algebra involved in the eigenvalue problem grows essentially.

As previously, let us first investigate the action of H_{hop} on the basis vectors

$$\begin{aligned}
&t^{-1}|01{:}kl)\to|21{:}kl)+|31{:}kl)+|02{:}kl)+|03{:}kl),\\
&t^{-1}|02{:}kl)\to|12{:}kl)+|32{:}kl)+|01{:}kl)+|03{:}kl),\\
&t^{-1}|03{:}kl)\to|13{:}kl)+|23{:}kl)+|02{:}kl)+|01{:}kl),\\
&t^{-1}|12{:}kl)\to|13{:}kl)+|32{:}kl)+|02{:}kl)+|10{:}kl),\\
&t^{-1}|13{:}kl)\to|03{:}kl)+|23{:}kl)+|12{:}kl)+|10{:}kl),\\
&t^{-1}|23{:}kl)\to|13{:}kl)+|03{:}kl)+|20{:}kl)+|21{:}kl).
\end{aligned} \tag{2.38}$$

Hence

$$\begin{aligned}
&H_{\text{hop}}|ij{:}kl)=t\sum_{\substack{p\neq i\\ p\neq j}}|pj{:}kl)+t\sum_{\substack{q\neq i\\ q\neq j}}|iq{:}kl),\\
&\tilde{H}_{\text{hop}}|ij{:}kl)=t\sum_{\substack{p\neq k\\ p\neq l}}|ij{:}pl)+t\sum_{\substack{q\neq k\\ q\neq l}}|ij{:}kq),\\
&V|ij{:}kl)=U[(\delta_{ik}+\delta_{jl})+(\delta_{il}+\delta_{jk})]|ij{:}kl).
\end{aligned} \tag{2.39}$$

The eigenvalue problem is investigated on vectors of the form

$$|\psi)=\sum_{ijkl}\alpha_{ij:kl}\cdot|ij{:}kl), \tag{2.40}$$

so that $H|\psi)=E|\psi)$ implies (we admit $\alpha_{ii:kl}=\alpha_{ij:kk}=0$)

$$\begin{aligned}
&t\sum_{p\neq i}\alpha_{pj:kl}+t\sum_{q\neq j}\alpha_{iq:kl}+t\sum_{p\neq k}\alpha_{ij:pl}+t\sum_{q\neq l}\alpha_{ij:kq}\\
&\quad+[U(\delta_{ik}+\delta_{jl}+\delta_{il}+\delta_{jk})-E]\alpha_{ij:kl}=0.
\end{aligned} \tag{2.41}$$

Let us renumber the basis vectors (2.38) to form the sequence $|1{:}kl),|2{:}kl),\ldots,|6{:}kl)$ and then turn them to the form $|m{:}n)$, where both m and n vary from 1 to 6. Then (2.41) becomes the homogeneous system of linear equations whose square matrix $\mathscr{A}$ (36×36) has the following block form:

$$\mathscr{A}=\begin{bmatrix} A_1 & tI & tI & -tI & -tI & 0 \\ tI & A_2 & tI & tI & 0 & -tI \\ tI & tI & A_3 & 0 & tI & tI \\ -tI & tI & 0 & A_4 & tI & -tI \\ -tI & 0 & tI & tI & A_5 & tI \\ 0 & -tI & tI & -tI & tI & A_6 \end{bmatrix}, \tag{2.42}$$

where the 6×6 matrices A_i differ from each other by a permutation of the $(2U-t)$ entry along the diagonal to the ith place.

The form of A_1 reproduces that of $\mathscr{A}$

$$A_1=\begin{bmatrix} 2U-E & t & t & -t & -t & 0 \\ t & U-E & t & t & 0 & -t \\ t & t & U-E & 0 & t & t \\ -t & t & 0 & U-E & t & -t \\ -t & 0 & t & t & U-E & t \\ 0 & -t & t & -t & t & -E \end{bmatrix}. \tag{2.43}$$

For a coordinate representative of $|\psi)$ we adopt

$$\mathscr{A}|\alpha)=0,\quad |\alpha)=|a,b,c,d,e,f), \tag{2.44}$$

where each entry is the six-dimensional vector, and $|\alpha)$ is understood as an element of the direct sum of six six-dimensional spaces. Here, $\mathscr{A}|\alpha)=0$ implies

$$\begin{aligned}
&A_1a+t(b+c-d-e)=0,\\
&A_2b+t(a+c+d-f)=0,\\
&A_3c+t(a+b+e+f)=0,\\
&A_4d+t(-a+b+e-f)=0,
\end{aligned} \tag{2.45}$$

$$A_5 e + t(-a + c + d + f) = 0,$$
$$A_6 f + t(-b + c - d + e) = 0,$$

which is still too complicated unless a guess

$$A_1 a = \alpha a, \quad A_2 b = \alpha b, \quad \ldots, \quad A_6 f = \alpha f \tag{2.46}$$

is made. The system described by (2.46) admits nontrivial solutions only if

$$\det(A_i - \alpha I) = 0, \quad \forall\, i = 1,2,\ldots,6. \tag{2.47}$$

One should observe that each matrix $A_i - \alpha I$ can be mapped into $A_j - \alpha I$, $i \neq j$ through subsequent rows–columns–rows transpositions which are followed if necessary by a simultaneous multiplication by (-1) of the column–row pair. These operations do not affect the determinant value, hence it is enough to consider

$$\det(A_1 - \alpha I) = 0. \tag{2.48}$$

To further simplify our problem, let us admit that the first component of each vector a,b,c,d,e,f coincides with some component of the vector a according to

$$a_1, \quad b_1 = a_2, \quad c_1 = a_3, \quad d_1 = a_4, \quad e_1 = a_5, \quad f_1 = a_6. \tag{2.49}$$

By virtue of (2.45) these components must satisfy the system of six linear equations governed by the matrix X

$$Xa = 0, \tag{2.50}$$

$X \doteq A_1$(each element of the diagonal replaced by α).

Then $\det X = 0$ must hold true, while $\det(A_1 - \alpha I) = 0$ is valid simultaneously. We thus conclude that the vector a must solve

$$(X + A_1 - \alpha I)a = 0, \quad (X - A_1 + \alpha I)a = 0, \tag{2.51}$$

and the necessary condition for the solvability of the second vector equation is that

$$\det(X - A_1 + \alpha I) = [2\alpha - (2U - E)]\cdot[2\alpha - (U - E)]^4 (2\alpha + E) = 0, \tag{2.52}$$

i.e.,

$$\alpha = \tfrac{1}{2}(2U - E) \text{ or } \alpha = \tfrac{1}{2}(U - E) \quad \text{or} \quad \alpha = -\tfrac{1}{2}E. \tag{2.53}$$

Since (2.50) must be solvable as well, with a bit tedious algebra we finally arrive at

$$\det X = (2t - \alpha)^2(2t + \alpha)[\alpha^3 - \alpha 2t(1 + t) - 4t^3] = 0, \tag{2.54}$$

which implies

$$\alpha = 2t, \quad \alpha = -2t, \tag{2.55}$$
$$\alpha^3 + \alpha[2t(1 + t)] - 4t^3 = 0.$$

The third-order equation has a standard form $\alpha^3 + p\alpha + q = 0$ to which the Cardano formulas apply. The discriminant reads

$$\Delta = q^2 + 4p^3/27 = 16t^3[t^3 + \tfrac{2}{27}(1 + t)^3] \tag{2.56}$$

and is positive for $t > 0$. Then the third-order equation admits the single real root

$$\alpha = (q/2 + \Delta^{1/2})^{1/3} + (q/2 - \Delta^{1/2})^{1/3}. \tag{2.57}$$

Since we solved the problem for Hermitian matrices, and (2.53) holds true for all E, one of the α's should be eliminated. The study of the atomic ($U = 0$) limit leads here to the conclusion that

$$\alpha = \pm 2t \Rightarrow E = 2U \mp 4t, \quad E = U \mp 4t, \quad E = \mp 4t. \tag{2.58}$$

To demonstrate that this result is correct, we shall give the explicit examples of the eigenvectors of $H_{\text{hop}} + \tilde{H}_{\text{hop}}$ in the $N = 4, M = \tilde{M} = 2$ case (more detailed study of the issue will be given in Sec. III).

One immediately observes that

$$H_{\text{hop}}|\varphi) = 2t\,|\varphi),$$
$$|\varphi) = \varphi(a^*,a)|0) = (a_0^* a_1^* + a_0^* a_2^* + a_0^* a_3^*)|0), \tag{2.59}$$

and analogously for the vector $(a_1^* a_2^* + a_3^* a_1^* + a_2^* a_3^*)|0)$. Consistently

$$|\varphi,\tilde{\varphi}) \doteq \varphi(a^*,a)\cdot\varphi(\tilde{a}^*, \tilde{a})|0) \tag{2.60}$$

is a common eigenvector of H_{hop} and $\tilde{H}_{\text{hop}}$ corresponding to the eigenvalue $2t$. It implies that

$$(H_{\text{hop}} + \tilde{H}_{\text{hop}})|\varphi,\tilde{\varphi}) = 4t\,|\varphi,\tilde{\varphi}). \tag{2.61}$$

An analogous procedure for the vector $|\eta) = \eta(a^*, a)|0) = (a_2^* a_1^* + a_0^* a_2^* + a_1^* a_0^*)|0)$ demonstrates that

$$(H_{\text{hop}} + \tilde{H}_{\text{hop}})|\eta,\tilde{\eta}) = -4t\,|\eta,\tilde{\eta}),$$
$$|\eta,\tilde{\eta}) = \eta(a^*, a)\cdot\eta(\tilde{a}^*, \tilde{a})|0). \tag{2.62}$$

Consequently, the strikingly simple result (2.58) solves the spectral problem under consideration.

Getting the eigenvectors for $\mathscr{A}|\alpha) = 0$ reduces now to a simple though a bit tedious algebra with 6×6 matrices.

D. The $N = 3, M = 1, \tilde{M} = 2$ sector

Since the basis in the present case reads

$$|k{:}ij) = a_k^* \tilde{a}_i^* \tilde{a}_j^*|0),$$
$$|\psi) = \sum_{ijk} \alpha_{k:ij}\,|k{:}ij), \tag{2.63}$$

it is enough to combine the studies of previous sectors. The action of H_{hop} was considered in Sec. II A, while that of $\tilde{H}_{\text{hop}}$ in Sec. II B.

Since the tilde basis has six elements while the nontilde basis has four, the eigenvalue problem in the $N = 3$, $M = 1$, $\tilde{M} = 2$ sector will lead to the 24×24 matrix governing the homogeneous system of 24 linear equations with respect to 24 variables. We observe that

$$V|k{:}ij) = U(\delta_{ki} + \delta_{kj})|k{:}ij). \tag{2.64}$$

Hence $H|\psi) = E|\psi)$ implies

$$t\sum_{l\neq k} \alpha_{l:ij} + t\sum_{p\neq i} \alpha_{k:pj} + t\sum_{q\neq j} \alpha_{k:iq} + [U(\delta_{ki} + \delta_{kj}) - E]\alpha_{k:ij} = 0. \tag{2.65}$$

To construct the matrix $\mathscr{A}$ we use the $l = 0,1,2,3$ ordering with respect to the row index while that of

$(ij) = (01),(02),(03),(12),(13),(23)$ with respect to the column label.

The final ordering of rows and columns of $\mathscr{A}$ comes from the sequence (0:01),..., (0:23), (1:01),..., (3:23) of labels for $\alpha_{k:ij}$. $\mathscr{A}$ has the following 6×6 block form:

$$\mathscr{A} = \begin{bmatrix} A_1 & tI & tI & tI \\ tI & A_2 & tI & tI \\ tI & tI & A_3 & tI \\ tI & tI & tI & A_4 \end{bmatrix}, \tag{2.66}$$

where the matrices A_i differ in the diagonal elements only

$$\begin{aligned} A_1&: U-E, U-E, U-E, -E, -E, -E, \\ A_2&: U-E, -E, -E, U-E, U-E, -E, \\ A_3&: -E, U-E, -E, U-E, -E, U-E, \\ A_4&: -E, -E, U-E, -E, U-E, U-E, \end{aligned} \tag{2.67}$$

and the matrix A_1 is given as follows:

$$A_1 = \begin{bmatrix} U-E & t & t & -t & -t & 0 \\ t & U-E & t & t & 0 & -t \\ t & t & U-E & 0 & t & t \\ -t & t & 0 & -E & t & -t \\ -t & 0 & t & t & -E & t \\ 0 & -t & t & -t & t & -E \end{bmatrix}. \tag{2.68}$$

The problem of determining the coefficient representative $|\alpha)$ of the vector $|\psi)$ is here reduced to the method of Sec. II A.

$$\det(A_1 - tI) = 0, \tag{2.69}$$

with the difference that the computation involves 6×6 matrices and thus becomes a bit tedious. The final result, we arrive at, is

$$(t+U-E)\cdot(E+t)\cdot W(U,E,t) = 0, \tag{2.70}$$

where $W(U,E,t)$ is the fourth-order polynomial with respect to E

$$\begin{aligned} W(U,E,t) = &-3E^4 + E^3(4U-18t) + E^2(-4t-21t^2 \\ &+ 24Ut - 2U^2) + E(6t^2 + 9Ut^2 - 10t^2 \\ &- 4U^2t + 4Ut) \\ &+ (6t^3 - Ut^3 - 2t^2U + 2t^2U^2), \end{aligned} \tag{2.71}$$

which is in principle solvable. Namely, the use of the Descartes–Euler prescription to solve the equation $x^4 + ax^3 + bx^2 + cx + d = 0$ is that the substitution $x = y - (a/4)$ yields $y^4 + py^2 + qy + r = 0$, whose four roots are given by the formula $\pm\sqrt{z_1} \pm \sqrt{z_2} \pm \sqrt{z_3}$, provided the signs are reconciled with the demand $(\pm\sqrt{z_1})(\pm\sqrt{z_2})(\pm\sqrt{z_3}) = \sqrt{z_1 z_2 z_3} = -q/8$, and z_1, z_2, z_3 are the roots of the cubic equation

$$z^3 + \frac{p}{2}z^2 + \frac{p^2-4r}{16}z - \frac{q^2}{64} = 0.$$

Remark: We have thus concluded the study of nontrivially solvable sectors. The case of the pure hopping spectrum e.g., $M = 0, \tilde{M} \leqslant N$ or $\tilde{M} = 0$, $M \leqslant N$, we shall discuss in another context, in Sec. III.

III. BOSE VERSION OF THE HUBBARD MODEL ON THE TETRAHEDRON AND ITS FERMI STATES

Let us consider the four-site spatial problem of the previous section upon the change of the canonical anticommutation relations into the canonical commutation ones, which we implement through changing the notation from $a^{\#}, \tilde{a}^{\#}$ to $b^{\#}, \tilde{b}^{\#}$

$$\begin{aligned} &[b_i^{\#}, \tilde{b}_j^{\#}]_- = 0, \\ &[b_i, b_j^{*}]_- \equiv \delta_{ij} \equiv [\tilde{b}_i, \tilde{b}_j^{*}]_-, \end{aligned} \tag{3.1}$$

other commutators vanishing, and the notation $\equiv$ indicating that a proper domain must here be chosen in the state space of the system. The Fock-ness property is assumed as well:

$$b_j|0)_{\mathrm{B}} = \tilde{b}_j|0)_{\mathrm{B}} = 0, \quad \forall_j. \tag{3.2}$$

To distinguish between the Fermi and Bose Hamiltonians we shall use the abbreviation H_{F} and H_{B}, respectively. H_{F} coincides with the previous notation H and $H_{\mathrm{B}} = H_{\mathrm{F}}$ $(a^{\#}, \tilde{a}^{\#} \to b^{\#}, \tilde{b}^{\#})$ while $|0)_{\mathrm{F}}$ coincides with the previous $|0)$.

In the $M = \tilde{M} = 1$ sector it is trivial to observe that the mere replacement $a^{\#}, \tilde{a}^{\#} \to b^{\#}, \tilde{b}^{\#}$ transfers all the formulas to the Bose case, with the result that both the spectrum and the coefficient representative $|\alpha)$ of the eigenvector (2.12) are the same for the Bose and Fermi cases

$$|\psi)_{\mathrm{B}} = \sum_{ij} \alpha_{i:j}\, a_i^{*}\, \tilde{a}_j^{*}\, |0)_{\mathrm{B}},$$

$$|\psi)_{\mathrm{F}} = \sum_{ij} \alpha_{i:j}\, b_i^{*}\, \tilde{b}_j^{*}\, |0)_{\mathrm{F}}.$$

The situation becomes less trivial in other sectors when the Pauli principle at each lattice site may nontrivially intervene. To understand the issue let us consider the pure hopping nontilde term of (2.1) which, irrespective of the definite choice of statistics, has the form

$$\begin{aligned} \frac{1}{t}H_{\mathrm{hop}} &= \sum_{i=1}^{3} (A_0^{*} A_i + A_i^{*} A_0) \\ &\quad + \sum_{k=2}^{3} (A_1^{*} A_k + A_k^{*} A_1) + (A_2^{*} A_3 + A_3^{*} A_2) \\ &= \sum_{i,j=0}^{3} A_i^{*}\, W_{ij} A_j, \end{aligned} \tag{3.3}$$

where the square 4×4 matrix W reads as follows[6]:

$$W = (W_{ij}),$$

$$W_{ij} = 1 - \delta_{ij} \Rightarrow \begin{cases} W = \sum_{l=0}^{N-1} c_l \gamma^l, \\ c_l = (1-\delta_{0l}), \\ \gamma^4 = I, N = 4, \end{cases}$$

$$\gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{3.4}$$

It implies, furthermore

$$\begin{aligned} &\gamma f_k = \lambda_k f_k, \quad \lambda_k = \exp i\pi k/2 \doteq \varphi^k, \\ &\varphi = \exp i(\pi/2), \quad k = 0,1,2,3, \\ &f_k = \{f_{k\alpha}\}_{\alpha=0,1,2,3}, \\ &f_{k0} = 1,\ f_{k1} = \varphi^k,\ f_{k2} = \varphi^{2k},\ f_{k3} = \varphi^{3k}. \end{aligned} \tag{3.5}$$

Here

$$g_k = \frac{1}{2} f_k, \quad (g_k, g_l) = \sum_{\alpha=0}^{3} \bar{g}_{k\alpha} g_{k\alpha} = \delta_{kl}, \tag{3.6}$$

and, consistently,

$$B_k = \sum_{\alpha=0}^{3} \bar{g}_{k\alpha} A_\alpha, \quad B_k^* = \sum_{\alpha=0}^{3} g_{k\alpha} A_\alpha^* \tag{3.7}$$

preserve the (anti) commutation relations of A, A^*. Moreover (we use $N = 4$)

$$\begin{aligned} \sum_{\alpha\beta} A_\alpha^* W_{\alpha\beta} A_\beta &= \sum_k \left(\sum_l c_l \varphi^{kl} \right) B_k^* B_k \\ &= - \sum_{k \neq 0} B_k^* B_k + (N-1) B_0^* B_0 \\ &= - \sum_{k=0}^{3} B_k^* B_k + 4 B_0^* B_0. \end{aligned} \tag{3.8}$$

It means accordingly that in the pure Bose case

$$\begin{aligned} &H^{\mathrm{B}}_{\mathrm{hop}} = - t \sum_{k=0}^{3} \xi_k^* \xi_k + 4t \xi_0^* \xi_0, \\ &[\xi_k, \xi_l^*]_- \equiv \delta_{kl}, \end{aligned} \tag{3.9}$$

while in the pure Fermi case

$$\begin{aligned} &H^{\mathrm{F}}_{\mathrm{hop}} = - t \sum_{k=0}^{3} \eta_k^* \eta_k + 4t \eta_0^* \eta_0, \\ &[\eta_k, \eta_l^*]_+ = \delta_{kl}. \end{aligned} \tag{3.10}$$

The respective pure hopping eigenvectors have the form

$$\begin{aligned} &\xi_0^{*n_0} \xi_1^{*n_1} \xi_2^{*n_2} \xi_3^{*n_3} |0\rangle_{\mathrm{B}} = |n_0, n_1, n_2, n_3\rangle_{\mathrm{B}}, \\ &\eta_0^{*p_0} \eta_1^{*p_1} \eta_2^{*p_2} \eta_3^{*p_3} |0\rangle_{\mathrm{F}} = |p_0, p_1, p_2, p_3\rangle_{\mathrm{F}}, \\ &p_i = 0 \text{ or } 1, \end{aligned} \tag{3.11}$$

n_i's being the arbitrary positive integers.

At this point let us follow our previous analysis of Ref. 6 (there being applied to the harmonic chain in one space dimension) to demonstrate that both Bose and Fermi versions of the pure hopping problem (3.3) can be solved in the common state space as the joint Bose–Fermi spectral problem; further studies in this connection can be found in Refs. 7 and 8.

In the state space $\mathscr{H}_{\mathrm{B}}$ of the Bose system let us introduce the two-level projection for each k th mode

$$p_k = :\exp(-\xi_k^* \xi_k): + \xi_k^* :\exp(-\xi_k^* \xi_k): \xi_k, \tag{3.12}$$

and then form the product

$$P = \prod_{k=0}^{3} p_k, \tag{3.13}$$

which is a projection in $\mathscr{H}_{\mathrm{B}}$ with properties

$$\begin{aligned} &[H^{\mathrm{B}}_{\mathrm{hop}}, P]_- = 0, \quad P \xi_k^* P = \sigma_k^+, \quad P \xi_k P = \sigma_k^-, \\ &[\sigma_k^-, \sigma_k^+]_+ = p_k, \quad [\sigma_k^\#, \sigma_l^\#]_- = 0, \quad k \neq l, \\ &H^{\mathrm{B}}_{\mathrm{hop}} = P H^{\mathrm{B}}_{\mathrm{hop}} P + (1-P) H^{\mathrm{B}}_{\mathrm{hop}} (1-P). \end{aligned} \tag{3.14}$$

In particular

$$P \xi_0^{*n_0} \cdots \xi_3^{*n_3} |0\rangle_{\mathrm{B}} = (\sigma_0^+)^{n_0} \cdots (\sigma_3^+)^{n_3} |0\rangle \tag{3.15}$$

vanishes if any $n_i > 1$.

In the terminology of Refs. 6–8, P selects the subspace $\mathscr{H}_{\mathrm{F}}$ of Fermi states in $\mathscr{H}_{\mathrm{B}}$. Indeed by using the Jordan–Wigner formulas we immediately realize that Pauli operators both in (3.15) and in

$$P H^{\mathrm{B}}_{\mathrm{hop}} P = - t \sum_{k=0}^{3} \sigma_k^+ \sigma_k^- + 4t \sigma_0^+ \sigma_0^- \tag{3.16}$$

can be replaced by Fermi operators

$$\begin{aligned} &\eta_k^* = \exp(i\pi \sum_{l \neq k} \sigma_l^+ \sigma_l^-) \cdot \sigma_k^+, \\ &\eta_k = \exp(i\pi \sum_{l \neq k} \sigma_l^+ \sigma_l^-) \cdot \sigma_k^- \end{aligned} \tag{3.17}$$

acting on the proper subspace $\mathscr{H}_{\mathrm{F}}$ of the state space $\mathscr{H}_{\mathrm{B}}$ for the Bose system. It proves that if we solve the spectral problem for $H^{\mathrm{B}}_{\mathrm{hop}}$ and $H^{\mathrm{F}}_{\mathrm{hop}}$ in the Bose state space, then each Fermi eigenstate (3.11) is as well an exact eigenstate with the same eigenvalue for $H^{\mathrm{B}}_{\mathrm{hop}}$.

The same analysis applies to the pure hopping tilde contribution to (2.1).

Since

$$\begin{aligned} P \xi_0^{*n_0} \cdots \xi_3^{*n_3} |0\rangle_{\mathrm{B}} &= (\sigma_0^+)^{n_0} \cdots (\sigma_3^+)^{n_3} |0\rangle_{\mathrm{B}} \\ &= \eta_0^{*n_0} \cdots \eta_3^{*n_3} |0\rangle_{\mathrm{B}}, \end{aligned} \tag{3.18}$$

and $\eta_k^\#$'s are linear combinations of the former Fermi operators $a_k^\#$ we have also a guarantee that

$$[P, n_\alpha^{\mathrm{B}}]_- = 0, \quad n_\alpha^{\mathrm{B}} = b_\alpha^* b_\alpha, \tag{3.19}$$

i.e.,

$$n_\alpha^{\mathrm{B}} = P n_\alpha^{\mathrm{B}} P + (1-P) n_\alpha^{\mathrm{B}} (1-P) \quad \forall \alpha, \tag{3.20}$$

and analogously for $\tilde{n}_\alpha^{\mathrm{B}}$, which proves furthermore that for a complete Hubbard problem on the tetrahedron we have

$$\begin{aligned} &[H_{\mathrm{B}}, P]_- = 0, \\ &H_{\mathrm{B}} = H_{\mathrm{F}} + (1-P) H_{\mathrm{B}} (1-P), \quad H_{\mathrm{F}} = P H_{\mathrm{B}} P. \end{aligned} \tag{3.21}$$

Hence the solution of the standard Fermi model provides us with the partial spectral solution for the associated Bose model in all $N, M, \tilde{M}$ sectors. If compared with Ref. 4, where the half-filled-band case was considered, one should realize that presently both models are governed by the same (in form) Hamiltonian, while in Ref. 4 the so-called pair Hamiltonian (which is to be a pure Bose reconstruction of the original Fermi model), significantly differs from the Fermi one. In fact, the result of Ref. 4 comes from the traditional "fusion" ideology of incorporating two Fermi modes in a single Bose one, albeit the domain constraints were necessary to establish the equivalence. In our case each Fermi mode has its unique Bose image; compare, e.g., also Chapter 10 of Ref. 9, where a more detailed discussion of the Bose–Fermi relationship can be found.

IV. THE SOLVABLE HUBBARD-TYPE PROBLEM IN THREE SPACE DIMENSIONS

Let us turn back to the general form of the Hubbard Hamiltonian in its Anderson's version[5]

$$H=\sum_{ij\sigma} b_{ij}\, c^{*}_{i\sigma}\, c_{j\sigma} + U\sum_{i} n_{i+}\, n_{i-}\,. \tag{4.1}$$

Our aim is to incorporate the previously solved tetrahedron problem into this general lattice framework. The appropriate definition of b_{ij}'s follows from the ansatz we make about the form of $H=H_{\mathrm{F}}$

$$H=\sum_{k} H_k + \tau \sum_{(ij)} \sum_{\sigma} c^{*}_{i\sigma}\, c_{j\sigma} \doteq H_0 + \tau V. \tag{4.2}$$

Here each H_k represents the four-site (tetrahedron) problem of Sec. II and k is used to indicate the pairwise disjoint quartets of sites (independent tetrahedrons)

$$[H_k,H_l]_{-}=0, \quad \forall k,l. \tag{4.3}$$

In principle k is the three-dimensional lattice label. The strength of hopping within each tetrahedron we admit to equal $t>0$ while the external pure hopping "noise" is understood to provide the very weak perturbation $0<\tau\ll 1$ of the set of tetrahedrons.

Since τ is very small, the problem (4.2) can be solved by means of the standard stationary state perturbation theory for the case of the degenerate spectrum. Here H_0 is the solvable unperturbed problem, while the τV term is the underlying perturbation.

Due to the previous considerations we have in hand the complete spectrum and (up to algebraic manipulations) eigenfunctions of H_0. Presently we shall follow a bit simpler route by explicitly considering the half-filled-band variant of the model, where the overall number of electrons coincides with the (actually large) number of sites N. Then in agreement with Ref. 4 the $M,\tilde{M}$ must satisfy $N=M+\tilde{M}$. In our case the additional conservation rules should be observed. Namely, $\forall k$, $N_k=4$, and $M_k,\tilde{M}_k$ are the constants of motion.

We admit that the set of tetrahedrons exhausts the whole of the lattice, which is thus a net of $n=N/4$ Hubbard cells

$$\begin{aligned} &H_0=\sum_{k} H_k, \quad [H_k,H_l]_{-}=0,\\ &H_k=H(a^{*}_k,a_k,\tilde{a}^{*}_k,\tilde{a}_k). \end{aligned} \tag{4.4}$$

We use the notation $\mathscr{H}_k$ for the k th (cell) state space so that the state space of H_0 emerges as the direct tensor product of distinct $\mathscr{H}_k$'s

$$\mathscr{H}=\prod_{k}{}^{\otimes}\mathscr{H}_k, \quad |0)=\prod_{k}{}^{\otimes}|0)_k, \tag{4.5}$$

$|0)_k$ being the Fock vacuum in $\mathscr{H}_k$. Let us recall that, traditionally, (4.5) is meant to describe the lattice of elementary (pointlike) quantum systems, each one being attributed to its (k th) site. In our case each "site" stands for a four-site cell which has a nontrivial spatial structure. Upon introducing the τV perturbation, the interaction between the Hubbard cells is generated by the interaction of their constituents.

Since the half-filled-band case is presently of interest for us, only the $4=M+\tilde{M}$ sectors are of interest at each cell. It means that each k th cell may contribute the following nonzero energies to the spectrum of H_0:

$$\begin{aligned} &M=1, \quad \tilde{M}=3, \quad (M=3,\tilde{M}=1),\\ &E^{13}_{0}=U, \end{aligned} \tag{4.6}$$

$$E^{13}_{\pm}=\tfrac{1}{2}\left[U-4t\pm(12t^2+(2t-U)^2)\right]^{1/2},$$

$$\begin{aligned} &M=\tilde{M}=2,\\ &E^{22}_{1\pm}=2U\pm 4t, \end{aligned} \tag{4.7}$$

$$\begin{aligned} &E^{22}_{2\pm}=U\pm 4t,\\ &E^{22}_{3\pm}=\pm 4t. \end{aligned}$$

Since the eigenvalues of $M=\Sigma_k M_k$ do not exceed the value $N/2$ we can easily reconstruct the spectrum of the N-site lattice $(N=4n)$ by considering the sequence $m=1,2,3=(1+2)$, $4=(2+2)$ or $(1+1+2)$ or $(1+1+1+1)$, $5=(2+2+1)$, and so on, up to $2n$. If we denote

$$\epsilon_1=E^{22}_{1+}, \quad \epsilon_2=E^{22}_{1-}, \quad \ldots, \quad \epsilon_6=E^{22}_{3-}, \tag{4.8}$$

$$\epsilon_7=E^{13}_{0}, \quad \epsilon_8=E^{13}_{+}, \quad \epsilon_9=E^{13}_{-}, \tag{4.9}$$

then the energy of H_0 is given by the formula

$$\epsilon_{(\alpha)}=\sum_{i=1}^{9}\alpha_i\epsilon_i, \tag{4.10}$$

where α_i's are the positive integers satisfying

$$2\sum_{i=1}^{6}\alpha_i+\sum_{i=7}^{9}\alpha_i=m, \quad 1\leqslant m\leqslant\frac{N}{2}, \tag{4.11}$$

m being the eigenvalue of $M=\Sigma_k M_k$.

Obviously this spectrum is degenerate. Let $\epsilon_{(\alpha)}$ be the degenerate eigenvalue of H_0, the respective eigenvectors we denote $|1,\alpha)$, $|2,\alpha),\ldots,|k,\alpha)$.

As is well known, the effect of the perturbation τV is the fine splitting of each energy level $\epsilon_{(\alpha)}$ into (the first-order contribution is accounted for)

$$E_{(\alpha),\beta}=\epsilon_{(\alpha)}+\tau(\beta\,|V|\beta), \tag{4.12}$$

where $(\beta\,|V|\beta)$ stands for the eigenvalue of the square $k\times k$ matrix $(i,\alpha|V|j,\alpha)$, $i,j,=1,2,\ldots,k$.

We claim that the spectrum (4.12) of the just-introduced Fermi model is also an exact (albeit partial only) spectrum of the related pure Bose model, the respective eigenvectors being the Fermi states of this Bose system. To have this conjecture justified it is enough to exploit the discussion of Sec. III and to observe the following property of the pure hopping τ perturbation. We have N large, hence up to the boundary terms, the simple momentum space version of τV can be introduced[4]

$$\tau V=\sum_{p}\epsilon_p(a^{*}_p\,a_p+\tilde{a}^{*}_p\,\tilde{a}_p), \tag{4.13}$$

provided

$$\epsilon_p=-\tau\sum_{\gamma}\exp(ip\cdot R_\gamma), \tag{4.14}$$

where R_γ denotes the nearest-neighbor lattice vector, and the passage to the configuration space is realized via

$$a_s = \frac{1}{\sqrt{N}} \sum_p a_p \exp(ip \cdot R_s), \tag{4.15}$$

where s is the site label.

By making use of the general construction of Refs. 10 and 11 allowing us to embed the canonical Fermi algebra (CAR) in the canonical Bose algebra (CCR), one finds the trivially realized relations

$$\begin{aligned} &\tau V_B = \sum_p \epsilon_p (b_p^* b_p + \tilde{b}_p^* \tilde{b}_p), \\ &[P, V_B]_- = 0, \\ &P V_B P = V_F = V, \end{aligned} \tag{4.16}$$

where the projection P in the state space of the Bose system coincides with the operator unit 1_F of the (bosonized) Fermi algebra.

We cannot completely debate on the utility of the just-constructed model in the solid state physics, though its insulating–conducting properties might deserve further study. Let us, however, emphasize a feature of the model which seems to be of interest from the point of view of high-energy physics.

Namely, in our case we deal with the three-dimensional "gas" of four-site "composites," each one being nontrivially structured in space. The interaction between the separate Hubbard cells is then mediated (generated) through interactions of the constituents. It makes a significant difference if compared with the standard concept of interaction of elementary pointlike quantum objects widely used in the literature.

Since in $1 + 1$ the Hubbard model can be related to the continuum field theory model, the chiral-invariant Gross–Neveu one, it could be quite instructive to investigate the (continuum) limiting behavior of our model.

Note added in proof: After completing the present paper I have realized that an independent group theoretical analysis of the tetrahedral Hubbard problem was made in L. M. Falicov and R. H. Victora [Phys. Rev. B **30**, 1695 (1984)]. The secular equations for the energy values were found in all $0 \leqslant N \leqslant 8$ sectors, albeit with no details of the derivation. Since we did not achieve complete agreement with these results, and since the method of Sec. I allows to get an insight into the structure of eigenvectors (which was also the problem of Ref. 4) the interested reader should view both approaches as complementary up to possible inaccuracies in computing the energy spectrum (e.g., determinants).

ACKNOWLEDGMENTS

I am sending words of gratitude to Dr. Ewa Ziemniak and Professor Heinz Barentzen from Mülheim a.d. Ruhr for bringing the family of Hubbard models to my attention.

[1] M. Cyrot, Physica B,C **91**, 141 (1977); see also A. A. Ovcinnikov, I. I. Ukrainski, and G. V. Kventsel, Usp. Fiz. Nauk. **108**, 81 (1972).
[2] M. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).
[3] E. Lieb and F. Wu, Phys. Rev. Lett. **20**, 1445 (1968).
[4] H. Barentzen, Phys. Rev. B **28**, 4143 (1983).
[5] P. W. Anderson, Rev. Mod. Phys. **50**, 191 (1978).
[6] P. Garbaczewski, J. Math. Phys. **26**, 2039 (1985).
[7] P. Garbaczewski, J. Math. Phys. **25**, 862 (1984).
[8] P. Garbaczewski, J. Math. Phys. **26**, 490 (1985).
[9] P. Garbaczewski, *Classical and Quantum Field Theory of Exactly Soluble Nonlinear Systems* (World Scientific, Singapore, 1985).
[10] P. Garbaczewski, Commmun. Math. Phys. **43**, 131 (1975).
[11] P. Garbaczewski, J. Math. Phys. **19**, 642 (1978).

Hydrodynamic self-similar cosmological models

Abhik Kumar Sanyal and A. Banerjee
Department of Physics, Jadavpur University, Calcutta-700032, India
Dipankar Ray
Department of Mathematics, Jadavpur University, Calcutta-700032, India

(Received 24 August 1984; accepted for publication 21 March 1985)

Hydrodynamic self-similar solutions, as obtained by Chi [J. Math. Phys. **24**, 2532 (1983)] have been generalized by introducing new variables in place of the old space and time variables. A systematic procedure of obtaining a complete set of solutions has been suggested. The Newtonian analogs of all homogeneous isotropic Friedmann dust universes with spatial curvature $k = 0$ and ± 1 have been given.

I. INTRODUCTION

In hydrodynamics the idea of self-similarity has been exploited with considerable success to simplify time-dependent problems. Self-similar solutions are often the leading terms in an asymptotic expansion of a non-self-similar evolution, in a regime far from the initial conditions and far from the influence of boundary conditions. The behaviors of the self-similar solutions are likely to be encountered in nature and as such they have greater physical interest than merely being a special class of mathematical solutions. This concerns the concept of intermediate asymptotics as reviewed by Barenblatt and Zeldovich.[1]

In a recent paper Henriksen and Wesson[2] have discussed some Newtonian and relativistic self-similar cosmological models. They have presented some solutions by considering two constants of nature, viz., the Newtonian gravitational constant "G" and a characteristic velocity. Chi[3] has extended their work giving a set of new solutions considering G as the only constant in nature. One of the interesting solutions obtained by Chi[3] is the Newtonian analog of the Einstein–de Sitter cosmological model with vanishing spatial curvature.

In the present paper, while going for a complete set of hydrodynamic self-similar cosmological model solutions, we find the Newtonian analogs of all homogeneous isotropic Friedmann[4] dust universes with spatial curvature $k = 0$ and ± 1.

The complete set of hydrodynamic equations that governs spherically symmetric, isentropic fluid flow in an internal gravitational field includes the continuity equation, the momentum equation, the energy equation, and the gravitational equation. These equations are

$$\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial r} = -\frac{\rho}{r^2}\frac{\partial}{\partial r}(r^2 v), \tag{1.1}$$

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial r} = g - \frac{1}{\rho}\frac{\partial p}{\partial r}, \tag{1.2}$$

$$\frac{\partial}{\partial t}(p\rho^{-\gamma}) + v\frac{\partial}{\partial r}(p\rho^{-\gamma}) = 0, \tag{1.3}$$

$$\frac{\partial g}{\partial t} + v\frac{\partial g}{\partial r} = -\frac{2vg}{r}, \tag{1.4}$$

where ρ, p, and v are, respectively, the local density, pressure, and velocity; g is the local gravitational acceleration; and γ is the polytropic index.

Now, a one-parameter self-similar solution (which represents the intrinsic evolution behavior of a system not dependent on the incidental details of particular initial or boundary conditions) is sought for the density, velocity, pressure, and gravitational acceleration. If the Newtonian gravitational constant G is assumed to be the only constant in nature and further if there is no characteristic velocity, one can then introduce the dimensionless quantities η, V, and ϕ, and Q to obtain the general forms of the physical quantities as follows:

$$\begin{aligned} &\rho = (\lambda/r^{\theta})\eta(\xi), \quad v = (r/t)V(\xi), \\ &g = (r/t^2)\phi(\xi), \quad p = (\lambda/r^{\theta-2}t^2)Q(\xi), \end{aligned} \tag{1.5}$$

with

$$\xi = G\lambda t^2/r^{\theta}. \tag{1.6}$$

Here θ is a constant of the order unity. The parameter λ is introduced to make ξ dimensionless.

Now Eqs. (1.1)–(1.4) can be simplified in the following manner without assuming self-similarity. Since r and t are independent variables, Eq. (1.1) can be rewritten as

$$\frac{\partial}{\partial t}(r^2\rho) + \frac{\partial}{\partial r}(r^2\rho v) = 0,$$

and so there must exist a function ψ, such that

$$r^2\rho = \frac{\partial \psi}{\partial r}, \quad r^2\rho v = -\frac{\partial \psi}{\partial t}. \tag{1.7}$$

The relations (1.7) lead us to write the equation

$$\frac{\partial \psi}{\partial t} + v\frac{\partial \psi}{\partial r} = 0. \tag{1.8}$$

Comparing Eq. (1.8) with Eq. (1.3) it is possible to obtain the relation

$$\frac{\partial \psi/\partial t}{\partial \psi/\partial r} = -v = \frac{\partial(p\rho^{-\gamma})/\partial t}{\partial(p\rho^{-\gamma})/\partial r}, \tag{1.9}$$

which implies that the quantity $(p\rho^{-\gamma})$ must be a function of ψ. We therefore write

$$p\rho^{-\gamma} = f(\psi). \tag{1.10}$$

Further in view of (1.9) one can express the velocity v as

 0022-2488/85/102693-07$02.50

$$v = \left.\frac{\partial r}{\partial t}\right|_{\psi = \text{const.}} \tag{1.11}$$

Equation (1.11) implies that the new variable ψ characterizes the individual fluid particles.

In the following sections we have given a number of solutions of the set of hydrodynamic equations in the presence of gravity. These equations are written in terms of new variables ψ and t in place of r and t, so that on introducing the self-similar variables it is possible to get more general solutions in different cases. The technique followed in this paper leads us to a systematic derivation of different solutions, instead of obtaining only a few as special cases in an *ad hoc* manner from the equations given by Chi.[3] It is to be noted that the solutions are finally expressed in terms of the old variables r and t.

In Sec. II we introduce the new variables in place of the old space and time variables r and t in the hydrodynamic equations, without introducing the assumption of self-similarity at this stage. Finally, in Sec. III the self-similar variables are introduced in the equations. Explicit solutions in different cases are given in Secs. IV and V. Some of these are identical to those given by Chi.[3] In Sec. V, we discuss the isotropic homogeneous dust universes with spatial curvature $k = 0, \pm 1$.

II. CHANGE OF VARIABLES

In this section we shall change the set of independent variables from (r, t) to (ψ, t) in the following manner:

$$\left.\frac{\partial}{\partial t}\right|_{\text{old}} = \left.\frac{\partial}{\partial t}\right|_{\text{new}} + \frac{\partial \psi}{\partial t}\frac{\partial}{\partial \psi}, \tag{2.1}$$

$$\left.\frac{\partial}{\partial r}\right|_{\text{old}} = \frac{1}{r_\psi}\frac{\partial}{\partial \psi},$$

where $(\partial/\partial t)|_{\text{old}}$ implies a time derivative keeping r constant, $(\partial/\partial t)|_{\text{new}}$ implies a time derivative keeping ψ constant, and $(\partial/\partial r)|_{\text{old}}$ implies a derivative with respect to the radial vector r keeping t constant.

So,

$$\left.\frac{\partial v}{\partial t}\right|_{\text{old}} = v_{t_{\text{new}}} + v_\psi \frac{\partial \psi}{\partial t},$$

where $v_{t_{\text{new}}}$ implies $(\partial v/\partial t)|_{\text{new}}$, i.e., at constant ψ. We also have

$$\left.\frac{\partial v}{\partial r}\right|_{\text{old}} = v_\psi \frac{\partial \psi}{\partial r},$$

where v_ψ implies $\partial v/\partial \psi$. Combining these two equations with Eq. (1.8) we obtain quite naturally (from now on we shall refer to $v_{t_{\text{new}}}$ at constant ψ as v_t) the relation

$$\left.\frac{\partial v}{\partial t}\right|_{\text{old}} + v\left.\frac{\partial v}{\partial r}\right|_{\text{old}} = v_t. \tag{2.2}$$

Again,

$$\left.\frac{\partial p}{\partial r}\right|_{\text{old}} = \frac{p_\psi}{r_\psi}. \tag{2.3}$$

So combining Eqs. (1.2), (2.2), and (2.3) we obtain

$$r_{tt} + (1/\rho)(p_\psi/r_\psi) - g = 0, \tag{2.4}$$

where we have written r_t for $(\partial r/\partial t)|_\psi$.

Again by the above transformation rule given by Eq. (2.1) we further obtain

$$\left.\frac{\partial g}{\partial t}\right|_{\text{old}} = g_{t_{\text{new}}} + g_\psi \frac{\partial \psi}{\partial t},$$

where $g_{t_{\text{new}}}$ is $\partial g/\partial t$ at constant ψ. Also $(\partial g/\partial r)|_{\text{old}} = g_{\psi_{\text{new}}} \times (\partial \psi/\partial r)$, where $g_{\psi_{\text{new}}}$ is $\partial g/\partial \psi$. Using these relations in Eq. (1.8) we get

$$\left.\frac{\partial g}{\partial t}\right|_{\text{old}} + v\left.\frac{\partial g}{\partial r}\right|_{\text{old}} = g_t.$$

Here and in what follows we write only g_t for $(\partial g/\partial t)|_\psi$ and similarly for other variables. The above equation together with Eqs. (1.4) and (1.11) yields

$$g_t/g + 2r_t/r = 0,$$

which on integration gives us the relation

$$gr^2 = C(\psi). \tag{2.5}$$

Equation (1.7) further implies that

$$\rho = 1/r^2 r_\psi. \tag{2.6}$$

Using the relation (2.5), Eq. (2.4) can be written as

$$r_{tt} + \frac{1}{\rho}\frac{\partial p}{\partial r} - \frac{C}{r^2} = 0. \tag{2.7}$$

We make it clear at this point that the use of new variables, that is (ψ, t) instead of the old (r, t), will help us to derive many solutions systematically. But sometimes we use the old variables as well for convenience. To find the pressure, for example, we use Eq. (2.7), where p is written as a function of the old variable r. Finally, of course, all the quantities have to be obtained as functions of r and t.

III. SELF-SIMILAR SOLUTIONS

In view of Eqs. (1.5), (1.6), (1.9), and (2.5) we readily obtain the following relations:

$$p\rho^{-\gamma} = f(\psi) = G\lambda^{(2-\gamma)}r^{(\theta(\gamma-2)+2)}\eta^{-\gamma}Q\xi^{-1}, \tag{3.1}$$

and

$$gr^2 = C(\psi) = G\lambda r^{(3-\theta)}\phi\xi^{-1}. \tag{3.2}$$

One can conclude from (3.1) that unless $\theta(\gamma - 2) + 2 = 0$, the expression for the spatial coordinate r must be in the form of a product of a function of ψ and a function of ξ. The same conclusion can be made without any loss of generality from the relation (3.2) unless $(3 - \theta) = 0$. These results lead us to consider Eqs. (3.1) and (3.2) only for two distinct cases, viz., $\theta = 3$ and $\theta \neq 3$. Again for the case $\theta = 3$ it is found from the Eq. (3.2) that, if $C(\psi)$ is not a constant, ξ must be a function of ψ. This in turn yields, in view of (3.1), the result $\theta(\gamma - 2) + 2 = 0$ or the $\gamma = \frac{4}{3}$ polytropic relation.

IV. SOLUTIONS FOR $\theta = 3$

A. Case 1. $\theta = 3$, $C(\psi) \neq$ const, i.e., $\xi = \xi(\psi)$ and $\gamma = \frac{4}{3}$

In view of Eqs. (1.5) and (1.11) we find

$$v = r_t = (r/t)V. \tag{4.1}$$

Finding r_t from Eq. (1.6) and substituting its value in the above equation one finds

$$V = \tfrac{2}{3}. \tag{4.2}$$

Let us consider here a special case for the matter content, that is, for dust with $p = 0$. Using in this case Eq. (4.2) in (4.1) one can calculate r_{tt} as

$$r_{tt} = -\tfrac{2}{9}(r/t^2),$$

which, when substituting in Eq. (2.7), gives

$$C(\psi) = -\tfrac{2}{9}(G\lambda/\xi). \tag{4.3}$$

The above expression for $C(\psi)$ being substituted in (3.2) immediately yields

$$\phi = -\tfrac{2}{9}. \tag{4.4}$$

Again Eqs. (1.5) and (2.6) are combined to give

$$\rho = 1/r^2 r_\psi = \lambda\eta/r^3.$$

Computing r_ψ from Eq. (1.6) and using it in the above equation we finally obtain the relation

$$\eta = -(3/\lambda)\xi\psi_\xi. \tag{4.5}$$

So ultimately with the values of V and ϕ given by Eqs. (4.2) and (4.4), it is quite easy to find v and g from Eq. (1.5). Also since η is given as a function of ξ by Eq. (4.5) it is possible to find ρ from Eq. (1.5) as a function of ξ, where ξ is given by Eq. (1.6). Thus the solutions are obtained for the dust case $p = 0$. It is interesting to note that for the special choice of $\theta = 3$ and $C(\psi) \neq$ const, the values of the parameters γ and V are uniquely determined, viz. $\gamma = \frac{4}{3}$ and $V = \frac{2}{3}$.

B. Case 2. $\theta = 3$, $C(\psi) =$ const

In this case g is a function of r only, as can be quite easily observed from Eq. (2.5). Also from Eq. (3.2) it is found that

$$\phi = (C/G\lambda)\xi. \tag{4.6}$$

In view of Eqs. (1.5), (1.11), and (2.6) we obtain the two relations

$$r_\psi = r/\lambda\eta, \quad r_t = (r/t)V,$$

which in turn further yield

$$-\left.\frac{\partial\psi}{\partial t}\right|_{r=\text{ const}} = \frac{\lambda\eta V}{t}.$$

The above relation can be written in a useful form like

$$\left.\frac{\partial\psi}{\partial(G\lambda t^2/r^3)}\right|_{r=\text{ const}} = \left.\frac{\partial\psi}{\partial\xi}\right|_{r=\text{ const}} = -\frac{\lambda\eta V}{2\xi},$$

the integration of which can be immediately made to yield

$$\psi = -\int\frac{\lambda\eta V}{2\xi}d\xi + b(r) = a(\xi) + b(r), \tag{4.7}$$

where $b(r)$ is an arbitrary function of r. We should note that $\xi(\psi, t)$ and $r(\psi, t)$ are related to each other in the form given in (4.7). From Eq. (4.7) one can obtain

$$\frac{da(\xi)}{d\xi}(\xi)\left.\frac{\partial\xi}{\partial t}\right|_{\psi=\text{ const}} + \frac{db(r)}{dr}\left.\frac{\partial r}{\partial t}\right|_{\psi=\text{ const}} = 0.$$

Using the definition of ξ given in (1.6) and (1.11) in the above equation, we get for $V \neq 0$

$$\frac{da(\xi)}{d\xi}\xi\left(\frac{2}{V} - 3\right) = -r\frac{db(r)}{dr}.$$

Since the left-hand side is a function of ξ and the right-hand side is a function of r, the above equation holds in view of (1.6) only if both the sides are separately equal to a constant, say D. Then

$$r\frac{db(r)}{dr} = D = \xi\frac{da(\xi)}{d\xi}\left(3 - \frac{2}{V}\right). \tag{4.8}$$

The relation (4.8) on integration yields

$$b(r) = D\ln r + \text{ const},$$

and Eq. (4.7) takes the form

$$\psi = a(\xi) + D\ln r + \text{ const}. \tag{4.9}$$

Equation (4.9) has been obtained with the only assumption that $V \neq 0$. If we consider $V = \frac{2}{3}$, then from Eq. (4.8) we get $D = 0$ and as such $\psi = \psi(\xi)$, which has already been considered in case 1. So in the following section, the $V \neq \frac{2}{3}$ cases are discussed. Let us also consider dust for which $p = 0$. Equations (2.4) and (2.2) together yield

$$\left.\frac{\partial v}{\partial t}\right|_{r=\text{ const}} + v\left.\frac{\partial v}{\partial r}\right|_{t=\text{ const}} = g.$$

Finding the values of $(\partial v/\partial t)|_r$ and $v(\partial v/\partial r)|_t$ from Eq. (1.5) and substituting these values together with that of g from Eq. (2.5), in the above equation, we readily obtain the relation

$$\xi\frac{dV}{d\xi}(2 - 3V) + V^2 - V = \frac{C}{G\lambda}\xi.$$

This equation has the solution

$$V = \sqrt{\xi}, \quad C/G\lambda = -\tfrac{1}{2}. \tag{4.10}$$

Using Eqs. (1.5), (3.2), and (4.10) we arrive at the two relations

$$v = \sqrt{\xi}(r/t) \tag{4.11}$$

and

$$\phi = -\tfrac{1}{2}\xi. \tag{4.12}$$

Further finding the values of $\partial\psi/\partial t$ and $\partial\psi/\partial r$ from Eq. (4.9) and substituting these values in Eq. (1.8) together with the value of v from Eq. (4.11) we obtain

$$da(\xi) = [D/(3\xi - 2\sqrt{\xi})]d\xi,$$

which on integration yields

$$a(\xi) = \tfrac{2}{3}D\ln(\sqrt{\xi} - \tfrac{2}{3}) + \text{ const}. \tag{4.13}$$

With this value of a, ψ of Eq. (4.9) can be explicitly written in the form

$$\psi = \tfrac{2}{3}D\ln(\sqrt{\xi} - \tfrac{2}{3}) + D\ln r + \text{ const}. \tag{4.14}$$

Again a combination of Eqs. (1.5) and (2.6) yields

$$\frac{1}{r^2}\left.\frac{\partial\psi}{\partial r}\right|_{t=\text{ const}} = \frac{\eta\lambda}{r^3}.$$

We can obtain $(\partial\psi/\partial r)|_t$ from Eq. (4.14), which, when used in the above relation, gives

$$\eta = \eta_1/(2 - 3\sqrt{\xi}), \tag{4.15}$$

where η_1 is a constant, given by $\eta_1 = 2D/\lambda$.

So ultimately we obtain the complete solution for this case with $\theta = 3$, $C =$ const, and $V \neq \frac{2}{3}$. All the dimensionless quantities are known, viz. $V(\xi)$ from Eq. (4.10), $\phi(\xi)$ from Eq. (4.12), $\eta(\xi)$ from Eq. (4.15), and $Q = 0$, since $p = 0$. This solu-

tion has already been obtained by Chi[3] [see solution (9)], But here in Eq. (4.8) one can choose V arbitrarily as a function of ξ or a constant, so that $a(\xi)$ can be found from Eq. (4.8) and thus ψ is known explicitly from Eq. (4.9). Thus other quantities can also be easily obtained and hence all possible solutions for different values of V can be generated, whereas Chi[3] could obtain only a few special solutions. We shall now discuss some of the solutions with a special choice of V. Since the solution for $V=1$ has already been obtained by Chi[3] [see Eq. (5)] we shall consider in the following examples some other values of V.

Example 1. $V=2$: With this value of V, it is possible to find $a(\xi)$ from Eq. (4.8) as

$$a(\xi) = (D/2)\ln \xi + \text{const.} \tag{4.16}$$

With the help of Eq. (4.16), ψ can be found from Eq. (4.9) as

$$\psi = (D/2)\ln \xi + D \ln r + \text{const.} \tag{4.17}$$

So now η can be found out from Eqs. (1.5), (2.6), and (4.17) as

$$\eta = -\tfrac{1}{2}(D/\lambda) = \eta_0, \tag{4.18}$$

where η_0 is a constant. We obtain ϕ from Eq. (3.2) as

$$\phi = (C/G\lambda)\xi = \phi_0 \xi, \tag{4.19}$$

where $\phi_0 = C/G\lambda$ is a constant. For $V=2$ it is found from (1.5) that

$$v = r_t = 2(r/t),$$

and a second differentiation with respect to time coordinate at constant ψ yields

$$v_t = r_{tt} = 2(r/t^2).$$

Further using (1.5) in the relation (4.9) one finds also

$$g = \phi_0 (r/t^2)\xi.$$

With the help of the above two equations one can directly integrate Eq. (2.7) and get the solution for p in the form

$$p = 2\lambda\eta_0/t^2 r - c\lambda\eta_0/4r^4 + T(t), \tag{4.20}$$

where $T(t)$ is an arbitrary function of t. So all the quantities are known with this special choice of $V=2$. This is a new solution. It is to be noted that to find p, Eq. (2.7), instead of Eq. (2.4), has been utilized, because it gives directly the solution for p as a function of r and t.

Example 2. $V=\xi$: In this example, instead of considering V to be a constant, we have chosen V as a simple function of ξ. The solution can be obtained by the above procedure. We obtain $a(\xi)$ from Eq. (4.8) as

$$a(\xi) = (D/3)\ln(3\xi - 2). \tag{4.21}$$

We can obtain ψ from Eqs. (4.9) and (4.21) as

$$\psi = (D/3)\ln(3\xi - 2) + D \ln r + \text{const.} \tag{4.22}$$

We obtain η from Eqs. (1.5), (2.6), and (4.22) as

$$\eta = \eta_0/(2 - 3\xi), \tag{4.23}$$

where $\eta_0 = 2D/\lambda$ is a constant quantity. We find ϕ from Eq. (3.2) as

$$\phi = \phi_0 \xi. \tag{4.24}$$

The pressure p can be obtained again in the same way as in the previous example. Also, ρ, g, and v can be known from the value of ξ given in Eq. (1.6). All these give another set of new solutions.

V. SOLUTION FOR $\theta \neq 3$

We have already seen that if $\theta \neq 3$, the variable r, which is actually a function of the new variables ψ and t, can be expressed in view of (3.1) and (3.2) in the form

$$r = \alpha(\psi)\,\beta(\xi), \tag{5.1}$$

where $\alpha(\psi)$ and $\beta(\xi)$ are functions of ψ and ξ, respectively. Equations (1.5) and (2.6) together give a relation like

$$\rho = 1/r^2 r_\psi = \lambda\eta/r^\theta, \tag{5.2}$$

so that one can write

$$r_\psi = r^{(\theta-2)}/\lambda\eta.$$

Again from Eq. (1.5) we know

$$v = r_t = (r/t)V,$$

and as such one gets

$$\frac{r_t}{r_\psi} = \left.\frac{\partial\psi}{\partial t}\right|_{r=\text{const}} = -\frac{\lambda\eta V}{t r^{(\theta-3)}},$$

from which it immediately follows that

$$\left.\frac{\partial\psi}{\partial\xi}\right|_{r=\text{const}} = \left.\frac{\partial\psi}{\partial(G\lambda t^2/r^\theta)}\right|_{r=\text{const}} = -\frac{\lambda\eta V}{2\xi(\alpha\beta)^{(\theta-3)}}. \tag{5.3}$$

Again from Eq. (5.1) one gets

$$dr = \beta\alpha_\psi\, d\psi + \alpha\beta_\xi\, d\xi,$$

which enables us to write

$$\left.\frac{\partial\psi}{\partial\xi}\right|_{r=\text{const}} = -\frac{\alpha}{\alpha_\psi}\frac{\beta_\xi}{\beta}. \tag{5.4}$$

Comparing Eqs. (5.3) and (5.4), we arrive at the following results:

$$\alpha_\psi/\alpha = \alpha^{(\theta-3)}/H \tag{5.5}$$

and

$$\beta_\xi/\beta = (\lambda/H)[\eta V/2\xi\beta^{(\theta-3)}], \tag{5.6}$$

where H is a constant. Integrating Eq. (5.5), one further gets for $\theta \neq 3$

$$\alpha^{(3-\theta)} = [(3-\theta)/H]\psi + \text{const.} \tag{5.7}$$

Again Eq. (5.1), in view of (1.6), gives

$$r_\psi = \beta\alpha_\psi/[1 + \theta\xi\beta_\xi/\beta].$$

We use this expression for r_ψ in (5.2) and further utilizing (5.5) we obtain a relation

$$\beta^{(\theta-3)}/\eta[1 + \theta\xi\beta_\xi/\beta] = \lambda/H. \tag{5.8}$$

Equations (5.6) and (5.8) together now yield

$$V = 2\xi\beta_\xi/(\beta + \theta\xi\beta_\xi). \tag{5.9}$$

Also with the help of Eq. (5.1), Eq. (3.2) can be written in the form

$$C(\psi)/\alpha^{(3-\theta)} = G\lambda\beta^{(3-\theta)}\phi\xi^{-1}.$$

Since the left-hand side of the above equation is a function of ψ and the right-hand side is a function of ξ, the equation holds only if both sides are equal to a constant, that is

$$C(\psi) = E\alpha^{(3-\theta)} \tag{5.10}$$

and

$$\phi(\xi) = (E/G\lambda)\xi\beta^{(\theta-3)}, \tag{5.11}$$

where E is a constant.

Now the complete solutions for the $\theta \neq 3$ case are obtained from the set of equations (5.7)–(5.11). If, for example, V is given as a function of ξ, then from Eq. (5.9) β can be found as a function of ξ and as such β can be written explicitly as a function of r and t. Hence η and ϕ can be obtained from Eqs. (5.8) and (5.11), respectively. As β is now known, α can also be found as a function of r and t from Eq. (5.1). So from Eq. (5.7), ψ can be determined as a function of r and t. All the quantities being known, it is now easy to find the solution from the pressure p from Eq. (2.7). At this point we proceed to find the solutions identical with those appearing in the case of a Friedmann universe with zero cosmological constant, spatially uniform density distribution [$\rho = \rho(t)$], and $p = 0$. We define the function $R(t)$ as (see Chi[3])

$$v/r = V/t = \dot{R}/R, \tag{5.12}$$

where an overdot represents the time derivative. In Eq. (5.12), $R(t)$ is analogous to the scale factor of the Friedmann universe and is dependent on time alone. Therefore, V must be either a constant or a function of time alone. In the case $V = V(t)$, ξ is a function of time only, which implies $\theta = 0$ from the definition of ξ. Again spatially, uniform density distribution is a necessary condition for the Friedmann universe. Equation (1.5) implies that $\rho = \rho(t)$ is satisfied either for $\theta = 0$ or for $\eta(\xi) \simeq \xi^{-1}$. So in order to deal with the Friedmann model we have to consider either $V =$ const and $\eta \simeq \xi^{-1}$ or $\theta = 0$. The second case, that is, $\theta = 0$, however, gives only a limiting value for $\xi = \xi(t)$.

A. Case 1. $\theta \neq 3$, $V = V_0$ (constant), and $\eta = \eta_0\xi^{-1}$

Under these conditions, Eq. (5.9) can be easily integrated and one gets

$$\beta^{(2-\theta V_0)/V_0} = A\xi, \tag{5.13}$$

where A is the constant of integration. Since $\eta = \eta_0\xi^{-1}$, Eq. (5.8) can now be easily simplified by using Eq. (5.13) to yield

$$\beta^{(2-3V_0)/V_0} = (A\lambda\eta_0/H)[(2-\theta V_0)/2].$$

Since the right-hand side of the above equation is a constant, it is satisfied only if the left-hand side is also equal to a constant, that is, when

$$(2-3V_0)/V_0 = 0.$$

This immediately gives us

$$V = V_0 = \tfrac{2}{3}, \tag{5.14}$$

which implies, in view of Eq. (1.5),

$$v = \tfrac{2}{3}(r/t). \tag{5.15}$$

From Eq. (1.5) one can find the value of ρ by using Eq. (1.6) and it is given by

$$\rho = \eta_0/Gt^2. \tag{5.16}$$

Combining Eqs. (5.13) and (5.14) one gets

$$\xi\beta^{(\theta-3)} = 1/A.$$

Applying the above result in Eq. (5.11) it is not difficult to find that

$$\phi = \phi_0, \tag{5.17}$$

where $\phi_0 = E/AG\lambda$, E being a constant quantity. Therefore from Eq. (1.5) one finds

$$g = \phi_0(r/t^2). \tag{5.18}$$

Since $v = r_t$ is known from Eq. (5.15) and also, since g has already been given by Eq. (5.18), it is now possible to solve easily Eq. (2.7) for $p = 0$ to obtain

$$\phi_0 = -\tfrac{2}{3}, \tag{5.19}$$

which, when applied in Eq. (5.18), finally yields

$$g = -\tfrac{2}{3}(r/t^2). \tag{5.20}$$

Now with $V = \frac{2}{3}$, Eq. (5.12) becomes

$$\dot{R}/R = 2/3t, \tag{5.21}$$

and it is thus possible to express the fluid density ρ in the form

$$8\pi G\rho/3 = (\dot{R}/R)^2, \tag{5.22}$$

by suitably choosing the constant η_0. This is simply the Friedmann differential equation with zero cosmological constant and zero curvature constant ($k = 0$). It is a case of the Einstein–de Sitter universe as given by Chi[3] [see the solution (13) of Chi[3]].

B. Case 2. $\theta = 0$

In this case it is obvious from Eq. (1.5) that $\rho = \rho(t) = \lambda\eta$. Now since ξ is a function of t alone, so undoubtedly $\eta(\xi)$ is also a function of t only. So in view of Eqs. (5.11), (1.6), and (1.5), one gets

$$g = E\beta^{-3}r. \tag{5.23}$$

Again, since $\beta_t = \beta_\xi\xi_t$, one can find ξ_t from Eq. (1.6) and hence $\xi\beta_\xi$ can be obtained as

$$\xi\beta_\xi = t\beta_t/2,$$

which, being used in Eq. (5.9), yields

$$V = t\beta_t/\beta. \tag{5.24}$$

Using Eqs. (5.12) and (5.24) we find

$$\dot{R}/R = \dot{\beta}/\beta, \tag{5.25}$$

which on integration yields

$$\beta = \beta_0 R. \tag{5.26}$$

Here, β_0 is the constant of integration, and η can be found from Eq. (5.8), which, again being applied in the equation (1.5), gives

$$\rho\beta^3 = H,$$

and the above equation, together with Eq. (5.26) finally expresses the density ρ in the form

$$\rho = \rho_0/R^3, \tag{5.27}$$

where ρ_0 is constant. This is a well-known result for the Friedmann dust universe. Again from Eqs. (1.5) and (5.11) we have

$$g = E\beta^{-3}r. \tag{5.28}$$

Also using Eqs. (1.5) and (5.24) one finds

$$r_t = r\beta_t/\beta,$$

from which r_{tt} can be easily found. With the value of r_{tt} being calculated from the above expression and using the result of Eqs. (5.28) and (2.5) it is possible now to solve Eq. (2.7) for dust, that is, $p = 0$. We get

$$\beta^2\beta_{tt} = E.$$

Integrating the above equation we find

$$\dot{\beta}/\beta = \pm(1/\beta)(-E/\beta + E_1)^{1/2},$$

where E_1 is a constant of integration. Combining the above equation with Eq. (5.25) and using the result of Eq. (5.26) we find

$$\dot{R}^2 = \left(-\frac{E}{\beta_0^3}\frac{1}{R} + \frac{E_1}{\beta_0^2}\right).$$

Assuming $(-E\beta_0^{-3})$ to be positive everywhere, i.e., $(-E\beta_0^{-3}) = m^2$, where m is any real number, and also writing $E_1\beta_0^{-2} = -k$, we find

$$\dot{R}^2 = (m^2/R - k),$$

where k may be positive negative or zero. The above equation leads to the result

$$\dot{R}^2/R^2 + k/R^2 = m^2/R^3.$$

Using Eq. (5.27) in the above relation, we have

$$\dot{R}^2/R^2 + k/R^2 = (m^2/\rho_0)\rho. \tag{5.29}$$

Identifying m^2/ρ_0 with $8\pi G/3$, it is now possible to recognize the above equation to be the differential equation for the Friedmann universe with zero cosmological constant and nonvanishing spatial curvature ($k = \pm 1$). So the integration constant E_1 is associated with the space curvature k of the Friedmann universe. The vanishing of the spatial curvature ($k = 0$) leads to $R \sim t^{2/3}$ or $\beta \sim t^{2/3}$ and $\eta = \eta_0\xi^{-1}$, which is exactly the Einstein–de Sitter solution obtained by Chi,[3] except for the fact that here θ assumes only a fixed value zero.

We shall now give more solutions with different choices of V for a perfect fluid ($p \neq 0$) and for a spatially uniform density [$\rho = \rho(t)$] with $\theta = 0$.

1. Case 2a. $V = V_0 = const$, $\theta = 0$

So from Eq. (1.5) we have

$$v = V_0(r/t). \tag{5.30}$$

Equation (1.6) yields

$$\xi = G\lambda t^2. \tag{5.31}$$

Equation (5.9) can be easily integrated to yield

$$\beta = (\beta_0\xi)^{V_0/2}, \tag{5.32}$$

where β_0 is a constant of integration. We can obtain η from Eqs. (5.2) and (5.32) as

$$\eta = \eta_0\xi^{-3V_0/2}, \tag{5.33}$$

where $\eta_0 = (H/\lambda)\beta_0^{-3V_0/2}$ is a constant. So ρ can be found from Eqs. (1.5), (5.31), and (5.33) as

$$\rho = \lambda\eta_0(G\lambda)^{-3V_0/2}t^{-3V_0}. \tag{5.34}$$

We can obtain ϕ from Eq. (5.11) as

$$\phi = (E/G\lambda)\xi\beta^{-3}. \tag{5.35}$$

Equations (1.5), (5.31), and (5.35) together give

$$g = E\beta^{-3}r. \tag{5.36}$$

With these values of v, ρ, and g, Eq. (2.7) takes the form

$$\frac{\partial p}{\partial r} = \lambda\eta_0(G\lambda)^{-(3/2)V_0}t^{-3V_0}\left[\frac{Er}{(\beta_0 G\lambda)^{(3/2)V_0}t^{3V_0}} - (V_0^2 - V_0)\frac{r}{t^2}\right]. \tag{5.37}$$

If we now consider the case $V = 1$, with the help of the set of equations (5.30)–(5.37), we arrive at the following set of solutions:

$$v = r/t, \quad \rho = \lambda\eta_0(G\lambda)^{-3/2}t^{-3},$$
$$g = Er(\beta_0\xi)^{-3/2}, \quad p = (E\lambda\eta_0/2\beta^{3/2}(G\lambda)^3)(r^2/t^6) + T(t). \tag{5.38}$$

Here the constant of integration $T(t)$ must be equal to zero, since $\xi = \xi(t)$ and $p = (\lambda r^2/t^2)Q(\xi)$ [from Eq. (1.5)]. Finding $(\partial/\partial t)(p\rho^{-\gamma})$ and $(\partial/\partial r)(p\rho^{-\gamma})$ and using Eqs. (1.3) and (5.38) one finds immediately that $\gamma = \frac{4}{3}$. So the above uniform density solution (5.38), with $\theta = 0$ and $V = 1$, is valid only for a $\gamma = \frac{4}{3}$ polytrope.

If we now consider $V_0 = 2$ using the set of relations (5.30)–(5.37), we can again obtain the following solutions:

$$v = 2(r/t), \quad \rho = (H/\beta_0^3)(1/\xi^3) \propto (1/t^6),$$
$$g = (Er/\beta_0^3 G\lambda t^2\xi^2) \propto (r/t^6), \tag{5.39}$$

and

$$p = \frac{2}{3a_2}\frac{r^3}{t^8} - \frac{a_1}{2a_2}\frac{r^2}{t^6} + \text{const.}$$

In the above a_1 and a_2 are two constants given by

$$a_1 = E/(\beta_0 G\lambda)^3, \quad a_2 = (\beta_0 G\lambda)^3/H.$$

This is again another new set of solutions.

2. Case 2b. $V = 2\xi$, $\theta = 0$

In this case instead of considering V to be a constant, we have considered a simple functional relation of V and ξ. It is possible to get the relevant solutions by the same procedure and these are given by

$$v = 2G\lambda rt, \quad \rho = H\beta_0^{-3}e^{-3G\lambda t^2},$$
$$g = E\beta_0^{-3}re^{-3G\lambda t^2}, \quad p = F(t)r^2. \tag{5.40}$$

The solutions (5.40) are also another new set of solutions obtained by our procedure. So in this way we can now choose any arbitrary form for V to generate all possible solutions.

VI. CONCLUSION

In effect we have systematically derived a complete set of solutions for the hydrodynamic self-similar cosmological

models with the help of a new set of independent variables (ψ, t) in place of the old variables. Finally all the physical quantities have been expressed in terms of the old variables (r, t). Some of these solutions are in agreement with those previously obtained by Chi in an *ad hoc* manner. We have further obtained the Newtonian analogs of all homogeneous isotropic Friedmann dust universes with spatial curvature $k = 0$ as well as ± 1.

ACKNOWLEDGMENT

The work has been carried out under the financial support of the "University Grants Commission," India.

[1]G. I. Barenblatt and Y. B. Zeldovich, Ann. Rev. Fluid Mech. **4**, 285 (1972).
[2]R. N. Henriksen and P. S. Wesson, Astrophys. Space Sci. **53**, 429 (1978).
[3]L. K. Chi, J. Math. Phys. **24**, 2532 (1983).
[4]S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).