The planar approximation. II

C. Itzykson and J.-B. Zuber
Division de la Physique, CEN-Saclay, Boite Postale No. 2, 91190 Gif-Sur-Yvette, France

(Received 18 July 1979; accepted for publication 11 October 1979)

The planar approximation is reconsidered. It is shown that a saddle point method is ineffective, due to the large number of degrees of freedom. The problem of eliminating angular variables is illustrated on a simple model coupling two $N \times N$ matrices.

1. INTRODUCTION

The idea that a large-N expansion in the theory of SU(N) gauge fields is a means to generate an approximation to the true system remains an attractive one. Since the work of 't Hooft[1] on two-dimensional QCD, there have been various attempts at developing a systematic treatment. Recent claims have been made that this limit enables one to understand the connection with the string formulation of the dual model[2] and that it provides a semiquantitative understanding of various selection rules in the framework of quark interactions.[3]

The zero-dimensional counting problem in the same approximation is related to the theory of random matrices and might find applications in models of disordered media.[4] The question has been considered by mathematicians[5] and physicists[6] using combinatorial methods, or analytical ones.[7-9] In Sec. 2 we shall present a short review of this subject. In [A] the method was applied to quantum mechanics, where it was shown to give accurate approximations.

The existence of a large parameter N, namely the order of the invariance group, suggests at first that some form of the saddle point method might apply to the path-integral representing the transition amplitude. This seems further confirmed by the observation that expectation values of products of invariant operators $A,B,\cdots$ factorize in the large N limit:

$$\langle AB\cdots\rangle \to \langle A\rangle\langle B\rangle\cdots . \tag{1.1}$$

Could there exist a classical fluctuationless configuration to describe the situation?

Unfortunately, this turns out to be rather illusory, as will be illustrated in the following. To be specific, we shall study a simple model involving finitely many degrees of freedom, each one represented by a Hermitian $N \times N$ matrix M. These finitely many degrees of freedom might be thought of as a finite lattice approximation to a genuine d-dimensional continuum. The integrals to be considered have the form:

$$Z = \int \prod_1^p dM_i \exp\Big[-\sum_{i=1}^p V(M_i) + \sum_{i,j=1}^p \beta_{ij}\,\mathrm{tr} M_i M_j \Big], \tag{1.2}$$

with $V(M)$ a potential term, typically

$$V(M) = \tfrac{1}{2}\mathrm{tr}M^2 + \frac{g}{N}\mathrm{tr}M^4, \tag{1.3}$$

inducing a quartic anharmonic term with strength g/N. Here β_{ij} is a short-range "kinetic" coupling among sites, for instance, $\beta_{ij} = \beta$ if i and j are nearest neighbors and zero otherwise. Finally dM is a U(N)-invariant volume element. By allowing a finite number of space–time points and letting $N \to \infty$, we interchange, of course, the infinite volume (thermodynamic) limit and the restriction to planar diagrams (large-N limit). However, as long as this procedure is thought of as a means of generating Feynman diagrams in a series expansion in g corresponding to processes without infrared divergences, it seems without danger.

Rescaling M into $N^{1/2}M$, we can look upon (1.2) as an integral involving an action multiplied by the large number N, which calls for the saddle-point method of evaluation. This is obviously too naive since it omits two aspects of crucial importance. The first is the contribution of the measure itself and the second the large degeneracy due precisely to the invariance group, here U(N) with N^2 parameters. The search for a saddle point can only be undertaken once these degenerate degrees of freedom have been eliminated. When this is done, one deals with a basis of group invariants. A sharp distinction appears here between the above planar problem, and the one encountered in a seemingly analogous situation involving vector instead of matrix variables, such as the classical Heisenberg O(N) ferromagnet. In this case, the variables attached to each of the p points of the lattice are N-dimensional vectors S_i^a, $a = 1,\ldots,N$, $i = 1,\ldots,p$. A basis of invariants under the real orthogonal group O(N) is given in terms of the $p(p+1)/2$ scalar products $\mathbf{S}_i\cdot\mathbf{S}_j$, $i \leqslant j$. (Since $p \ll N$ no quantity involving a determinant does occur.) In the measure, the O(N) degrees of freedom can be factored out, leaving as a result

$$\prod_{i<j} d(\mathbf{S}_i\cdot\mathbf{S}_j)\,|\det(\mathbf{S}_i\cdot\mathbf{S}_j)|^{(N-p-1)/2}.$$

As a consequence,

$$\begin{aligned} Z_{\text{vector}} &= \int \prod_{i=1}^p d^N S_i \exp\Big[-\sum_i V(\mathbf{S}_i^2) + \sum_{i,j}\beta_{ij}\mathbf{S}_i\cdot\mathbf{S}_j \Big] \\ &= Z_0 \int \prod_{i<j} d(\mathbf{S}_i\cdot\mathbf{S}_j)\,|\det(\mathbf{S}_i\cdot\mathbf{S}_j)|^{(N-p-1)/2} \\ &\quad\times \exp\Big[-\sum_i V(\mathbf{S}_i^2) + \sum_{i,j}\beta_{ij}\mathbf{S}_i\cdot\mathbf{S}_j \Big], \end{aligned} \tag{1.4}$$

with Z_0 a normalization constant independent of V and the last integral extending over the positivity domain of the matrix $(\mathbf{S}_i\cdot\mathbf{S}_j)$. This expression is suited to the application of the saddle-point method, which will lead in this case to the usual $1/N$ expansion of the classical ferromagnet. This success

 0022-2488/80/030411-11$1.00

may be attributed to the fact that we have found a choice of $p(p+1)/2$ invariants much smaller in number than the original Np variables, more specifically much smaller than the large parameter N. As in thermodynamics, each degree of freedom will contribute to a connected quantity a finite amount. If the total number of degrees of freedom is vanishingly small as compared to the large parameter, the saddle-point method is useful and will be the starting point of a systematic expansion.

The situation is not as good in the matrix case. Except in the case of a single matrix where the space of invariants is N-dimensional and hence much smaller than the original space (of dimension N^2), it is sufficient to look at the set of invariants for two matrices to see that its size is comparable to the size of the original space. Consequently, fluctuations to all orders will be essential in the evaluation of the integral and no simple saddle-point method will work. Does this really mean that the planar problem is totally untractable in general? We have no answer to this question, but the successful applications of the planar approximation to quantum mechanics (see [A]) leaves some hope that an appropriate trick works for each specific case.

It is therefore of interest to confront the type of difficulty discussed above on the first nontrivial instance, namely when the integral Z involves two matrices only. This is the main part of the present investigation, to which we devote the last two sections. The result of the integration over angular variables corresponding to the unitary group transformations is presented in Sec. 3, while in Sec. 4 we discuss two expansions of the planar limit.

The outcome of this investigation seems a little disappointing. We feel, nevertheless, that it is worth being reported since it illustrates the nontrivial character of the planar approximation. Moreover, some of the expressions derived below might turn out to be useful in another context. Finally, our incomplete solution might raise other people's interest in finding a more complete answer.

2. THE COUNTING PROBLEM REVISITED

We recall the results obtained elsewhere[7-9] on the counting of diagrams with a definite topology. We shall generalize the theory to include an arbitrary polynomial interaction $V(M)$, which we assume even for simplicity:

$$V(M) = \tfrac{1}{2}\,\mathrm{tr}M^2 + \sum_{p\geqslant 2} \frac{g_p}{N^{p-1}}\,\mathrm{tr}M^{2p}. \tag{2.1}$$

Let dM be the unitary invariant measure on Hermitian $N\times N$ matrices

$$dM = \prod_{i=1}^{N} dM_{ii} \prod_{i<j} 2d\,\mathrm{Re}M_{ij}\, d\,\mathrm{Im}M_{ij}. \tag{2.2}$$

We define

$$Z(g) = \int dM\, e^{-V(M)}, \tag{2.3}$$

which makes sense in some appropriate (complex) domain for the coupling constants. In Eq. (2.1), the coefficients of higher-order terms are weighted with inverse powers of N in such a way that the perturbative expansion of

$$E\left(g, \frac{1}{N}\right) = -\frac{1}{N^2}\ln\left(\frac{Z(g)}{Z(0)}\right), \tag{2.4}$$

will produce contributions of the form

$$E\left(g, \frac{1}{N}\right) = \sum_{H\geqslant 0} \frac{1}{N^{2H}} E_{(H)}(g), \tag{2.5}$$

with $E_{(H)}$ defined in terms of diagrams drawn on a surface with H handles (H = genus of the surface). $E_{(0)}$ corresponds to the planar (or spherical) topology, $E_{(1)}$ to the torus, and so on. In the sequel, we shall concentrate on the vacuum "energy" E and leave aside questions dealing with Green's functions. This generalization can be done along the lines of [A]. In the measure dM, the angular factors corresponding to the unitary transformation U to a diagonal form:

$$M = U\Lambda U^{\dagger}, \quad \Lambda \equiv \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & \lambda_n \end{pmatrix}, \tag{2.6}$$

can be factored out. When integrating over an invariant function $f(M)$, i.e., such that $f(M) \equiv f(UMU^{\dagger})$, we have

$$\int dM f(M) = \frac{(2\pi)^{N(N-1)/2}}{\prod_1^N p!} \int \prod_1^N d\lambda_i \Delta(\lambda)^2 f(\Lambda), \tag{2.7}$$

with $\Delta(\lambda)$ the Vandermonde determinant

$$\Delta(\lambda) = \prod_{i>j} (\lambda_i - \lambda_j) = \det(\lambda_i^{\,j-1}). \tag{2.8}$$

The numerical factor in (2.7) will follow from our subsequent arguments. The structure of this relation suggests a connection between the calculation of Z and the theory of orthogonal polynomials, as discussed by Bessis[8] and Parisi.[9] This goes as follows. First define $\bar{g}_p = g_p/N^{p-1}$ and, for the time being, let $\bar{g}_p$ be considered as fixed, real, and such that the measure

$$d\mu(\lambda) = d\lambda e^{-V(\lambda)}, \tag{2.9}$$

is integrable. Here $V(\lambda)$ stands for

$$V(\lambda) = \frac{\lambda^2}{2} + \sum_{p\geqslant 2} \bar{g}_p \lambda^{2p}. \tag{2.10}$$

We call $Z_n(g)$ what was previously called $Z(g)$ in the case $n = N$, i.e.,

$$Z_n(g) = \int d\mu(\lambda_1)\cdots d\mu(\lambda_n)[\Delta(\lambda_1,\ldots,\lambda_n)]^2, \tag{2.11}$$

and define the polynomial of degree n

$$\begin{aligned} P_n(\lambda) &= (-1)^n P_n(-\lambda) \\ &= Z_n^{-1}(g) \int d\mu(\lambda_1)\cdots d\mu(\lambda_n)[\Delta(\lambda_1,\ldots,\lambda_n)]^2 \\ &\quad \times \prod_{s=1}^{n} (\lambda - \lambda_s). \end{aligned} \tag{2.12}$$

The term of highest degree has a coefficient equal to 1. The polynomials $P_n(\lambda)$ are orthogonal with respect to the measure $d\mu(\lambda)$. Indeed,

$$Z_n \int d\mu(\lambda_{n+1}) P_n(\lambda_{n+1}) \lambda_{n+1}^{s}$$

$$= \int \prod_1^{n+1} d\mu(\lambda_k)\Delta(\lambda_1,\ldots,\lambda_{n+1})\Delta(\lambda_1,\ldots,\lambda_n)\lambda_{n+1}^s$$

$$= \frac{1}{n+1}\int \prod_1^{n+1} d\mu(\lambda_k)\Delta(\lambda_1,\ldots,\lambda_{n+1})$$

$$\times \sum_{k=1}^{n+1} (-1)^{n+1-k}\lambda_k^s \Delta(\lambda_1,\ldots,\hat{\lambda}_k,\ldots,\lambda_{n+1}).$$

The sum inside the integrand is the expansion of the determinant

$$\begin{vmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} & \lambda_1^s \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} & \lambda_2^s \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \lambda_{n+1} & \cdots & \lambda_{n+1}^{n-1} & \lambda_{n+2}^s \end{vmatrix},$$

with respect to its last column. It vanishes for $s = 0,1,\ldots,n-1$, which proves the assertion. For $s = n$, we find

$$Z_n \int d\mu(\lambda) P_n(\lambda)\lambda^n = Z_n \int d\mu(\lambda) P_n^2(\lambda)$$
$$= Z_{n+1}/(n+1).$$

Hence

$$h_n = \int d\mu(\lambda) P_n^2(\lambda) = Z_{n+1}/(n+1)Z_n. \tag{2.13}$$

This relation shows that the knowledge of the orthogonal polynomials yields a handle on Z. A statistical interpretation can be given to Z as a partition of a one-dimensional repulsive Coulomb gas of particles interacting with a potential V.

The polynomials P_n satisfy a three-term recursion relation

$$\lambda P_n = P_{n+1} + R_n P_{n-1}. \tag{2.14}$$

Since

$$h_{n+1} = \int d\mu(\lambda) P_{n+1}\lambda P_n$$
$$= \int d\mu(\lambda)(P_{n+2} + R_{n+1}P_n)P_n = R_{n+1}h_n,$$

we have

$$R_n = \frac{h_n}{h_{n-1}} = \frac{n}{n+1}\frac{Z_{n+1}Z_{n-1}}{Z_n^2}. \tag{2.15}$$

Consequently,

$$Z_n = n!\, h_{n-1}h_{n-2}\cdots h_1 h_0 = n!\, R_{n-1}R_{n-2}^2\cdots R_1^{n-1}h_0^n, \tag{2.16}$$

with

$$Z_1 = h_0 = \int d\mu(\lambda). \tag{2.17}$$

Incidentally, this provides a justification for the factor occurring in Eq. (2.7). For choose there $f(M) = \exp(-\frac{1}{2}\operatorname{tr}M^2)$. The left-hand side is equal to $(2\pi)^{N^2/2}$. On the right-hand side, the integral is $Z_N(0)$, corresponding to Hermite polynomials with the measure $d\mu(\lambda) = e^{-\lambda^2/2}d\lambda$. In this case, $h_N = (2\pi)^{1/2}N!$ and hence $Z_N = N!\, h_{N-1}h_{N-2}\cdots h_0 = \Pi_1^N p!(2\pi)^{N/2}$. The factor in the right-hand side of (2.7) is just fitted to match these two results.

The preceding development follows from the standard textbook treatment. We now use an argument due to Bessis and Parisi to obtain a recursion formula on R_n. From Eqs. (2.13) and (2.14), it follows that

$$n h_n = \int d\lambda\, e^{-V}\lambda P_n' P_n$$
$$= \int d\lambda\, e^{-V}P_n'(P_{n+1} + R_n P_{n-1})$$
$$= R_n \int d\lambda\, e^{-V}P_n' P_{n-1}$$
$$= R_n \int d\lambda\, e^{-V}V' P_n P_{n-1},$$

where an integration by parts has been used to obtain the last equality. Now

$$\int d\lambda\, e^{-V}P_n(V'P_{n-1})$$
$$= \int d\lambda\, e^{-V}P_n\left[\lambda + \sum_{p\geq 1} 2(p+1)\bar{g}_{p+1}\lambda^{2p+1}\right]P_{n-1}$$
$$= h_n\left[1 + \sum_{p\geq 1} 2(p+1)\bar{g}_{p+1}\sum_{\text{paths}} R_{\alpha_1}\cdots R_{\alpha_p}\right]. \tag{2.19}$$

In this expression the coefficient of $2(p+1)\bar{g}_{p+1}$ is a sum over the $(2p+1)!/p!(p+1)!$ paths along a "staircase" from the stair at height $n-1$ to the one at height n, in $2p+1$ steps of one unit, $p+1$ up, p down. A factor R_α occurs when descending from the stair α down to stair $\alpha - 1$. For instance, we have

$$p = 1 \quad \sum_{\text{paths}} = R_{n-1} + R_n + R_{n+1},$$

$$p = 2 \quad \sum_{\text{paths}} = R_{n-2}R_{n-1} + R_{n-1}^2 + 2R_{n-1}R_n$$
$$+ R_{n-1}R_{n+1} + R_n^2 + 2R_nR_{n+1}$$
$$+ R_{n+1}^2 + R_{n+1}R_{n+2}, \tag{2.20}$$

and so on. We can, of course, express this result in terms of the $(n,n-1)$ matrix element of the Jacobi matrix λ in Eq. (2.14) raised to the power $(2p+1)$. Inserting this expression into (2.19) yields:

$$n = R_n\left(1 + \sum_{p\geq 1} 2(p+1)\bar{g}_{p+1}\sum_{\text{paths}} R_{\alpha_1}\cdots R_{\alpha_p}\right). \tag{2.21}$$

Since we are only interested here in the leading term of $Z(g)$, we shall only use the dominant estimate of R_n for n of order N. From Eq. (2.21), R_n is of order N, and we set

$$x = n/N, \quad R_n = Na^2(x). \tag{2.22}$$

This entails for $\bar{g}_{p+1} = g_{p+1}/N^p$,

$$x = a^2(x) + \sum_{p\geq 2} g_p \frac{(2p)!}{p!(p-1)!}a^{2p}(x). \tag{2.23}$$

The quantity of interest, namely the generating function for the number of planar diagrams, is $E_{(0)}(g)$ given by Eqs. (2.4), (2.5). Dropping the index (0), we find

$$E(g) = \lim_{N\to\infty} -\frac{1}{N^2}\left[\sum_1^{N-1}(N-n)\ln\left(\frac{R_n(g)}{R_n(0)}\right)\right.$$

$$+ N \ln\left(\frac{h_0(g)}{h_0(0)}\right)\Big]$$
$$= -\int_0^1 dx\,(1-x)\ln\left(\frac{a^2(x)}{x}\right). \tag{2.24}$$

To analyze these relations, we define the functions

$$v(\lambda) = \frac{\lambda^2}{2} + \sum_{p\geq 2} g_p \lambda^{2p},$$
$$w(\lambda) = \lambda^2 + \sum_{p\geq 2} \frac{(2p)!}{p!(p-1)!}\lambda^{2p}. \tag{2.25}$$

They are related through

$$w(\lambda) = \frac{1}{2\pi}\int_{-2\lambda}^{2\lambda} d\xi\,(4\lambda^2 - \xi^2)^{1/2} v''(\xi),$$
$$\frac{v(\lambda)}{\lambda} = \int_0^{\lambda/2} \frac{d\xi}{(\lambda^2/4 - \xi^2)^{1/2}}\frac{w(\xi)}{\xi}. \tag{2.26}$$

$E(g)$ can be written as

$$E(g) = \int_0^a d\lambda w'(\lambda)[1 - w(\lambda)]\ln\left(\frac{w(\lambda)}{\lambda^2}\right)$$
$$= -\tfrac{1}{2}\ln a^2 + \int_0^a \frac{d\lambda}{\lambda}(2-w)w - (g=0), \tag{2.27}$$

with $a \equiv a(1)$ defined through

$$1 = w(a) = \int_{-2a}^{2a} d\xi \frac{(4a^2 - \xi^2)^{1/2}}{2\pi} v''(\xi). \tag{2.28}$$

These expressions coincide, of course, with those given in [A] for the quartic potential. One can note that the condition $w(a) = 1$ follows from a variational principle. If a is left arbitrary in (2.27) without the subtraction term, then the relation between a and g expresses the stationnarity of E with respect to a.

The preceding development avoids completely the use of the saddle-point method as presented in [A]. Nevertheless, it contains implicitly the asymptotic distribution of eigenvalues of the matrix M. We recall that the original eigenvalues have been scaled down by a factor $\sqrt{N}$ to obtain the reduced variables. The density of eigenvalues, i.e., the distribution of roots of the polynomial $P_n(\lambda)$, is readily related to the Jacobi matrix for λ in a basis of orthonormal states

$$\mathscr{P}_n(\lambda) = \frac{1}{\sqrt{h_n}} P_n(\lambda). \tag{2.29}$$

From (2.14) and (2.15)

$$\lambda \mathscr{P}_n(\lambda) = \left(\frac{h_{n+1}}{h_n}\right)^{1/2}\mathscr{P}_{n+1} + R_n\left(\frac{h_{n-1}}{h_n}\right)^{1/2}\mathscr{P}_{n-1}$$
$$= \sqrt{R_{n+1}}\,\mathscr{P}_{n+1} + \sqrt{R_n}\,\mathscr{P}_{n-1}. \tag{2.30}$$

Consider now the quantity

$$\lim_{N\to\infty}\frac{1}{N^{p+1}}\int d\mu(\lambda)\sum_0^{N-1}\mathscr{P}_n(\lambda)\lambda^{2p}\mathscr{P}_n(\lambda) = \langle\lambda^{2p}\rangle. \tag{2.31}$$

In the reduced variables ($\lambda \to \lambda/\sqrt{N}$), we look for a positive density $u(\lambda)$ such that

$$\langle\lambda^{2p}\rangle = \int d\lambda\, u(\lambda)\lambda^{2p} = \int_0^1 dx\, a^{2p}(x)\sum_{q=0}^{p}\left[\binom{p}{q}\right]^2, \tag{2.32}$$

where the last expression results in this limit $N\to\infty$ from Eq. (2.30), using notations introduced in (2.22). This relation takes the form

$$\int d\lambda\, u(\lambda)\lambda^{2p}$$
$$= \int_0^1 dx\, a^{2p}(x)\int_{-1}^{1}\frac{dy}{\pi}\frac{(2y)^{2p}}{(1-y^2)^{1/2}}$$
$$= \int_{-2a(1)}^{2a(1)} d\lambda\,\lambda^{2p}\int_{|\lambda|/2}^{a(1)}\frac{d\mu}{\pi}\frac{w'(\mu)}{(4\mu^2-\lambda^2)^{1/2}},$$

and yields an even measure concentrated on the interval $(-2a, 2a)$, where $a \equiv a(1)$, equal to

$$u(\lambda) = \frac{1}{\pi}\int_{|\lambda|/2}^{a} d\xi \frac{w'(\xi)}{(4\xi^2 - \lambda^2)^{1/2}}$$
$$= \frac{(4a^2-\lambda^2)^{1/2}}{\pi}\frac{1}{2\pi}\oint_{-2a}^{2a}\frac{d\eta}{(4a^2-\eta^2)^{1/2}}\frac{v'(\eta)}{\eta-\lambda}. \tag{2.33}$$

This gives a distribution of the form $(1/\pi)(4a^2 - \lambda^2)^{1/2}$ times a polynomial in λ, equal for $\lambda^2 < 4a^2$ to the real part of an even analytic function

$$\frac{1}{2\pi}\int_{-2a}^{2a}\frac{d\eta}{(4a^2-\eta^2)^{1/2}}\frac{v'(n)}{\eta-\lambda},$$

vanishing faster than $1/\lambda$ as $|\lambda| \to \infty$, and with a discontinuity on the finite interval $\lambda^2 > 4a^2$ equal to $iv'(\lambda)(4a^2 - \lambda^2)^{-1/2}$. The condition $w(a) = 1$, is equivalent to the statement $\int_{-2a}^{2a} u(\lambda)d\lambda = 1$. This reproduces, of course, the result for the quartic interactions given in [A]. The extension of the previous analysis to include functions $v(\lambda)$ not necessarily even is, of course, possible. One can also proceed[8] to the systematic study of the corrections in powers of $1/N$, starting from the exact expression (2.21).

3. INTEGRATION OVER THE UNITARY GROUP

We return to the investigation of integrals of the type (1.2) over several $N \times N$ Hermitian matrices, in fact, to the simplest one involving two matrices

$$Z = \int dM_1\, dM_2 \exp\{ -[V(M_1) + V(M_2) - \beta \operatorname{tr}(M_1 M_2)]\}. \tag{3.1}$$

As explained in the Introduction it is important to integrate first over the angular variables. We are therefore led to study the expression

$$I(M_1, M_2; \beta) = \int dU \exp[\beta \operatorname{tr}(M_1 U M_2 U^\dagger)], \tag{3.2}$$

where dU is the normalized Haar measure on the unitary group $U(N)$. We can, in fact, restrict the integration to $SU(N)$ since this is the only part which acts effectively on the Hermitian matrices M in the adjoint representation. If Λ_1 and Λ_2 stand for the diagonal matrices of eigenvalues of M_1 and M_2 as in (2.6), I depends only on Λ_1 and Λ_2 and is, in fact, a symmetric function of each set. Then Z reduces to the form

$$Z = \frac{(2\pi)^{N(N-1)}}{(\Pi_1^N p!)^2} \int \prod d\lambda_{1,i} d\lambda_{2,i} \Delta^2(\Lambda_1)\Delta^2(\Lambda_2) \times \exp\{ -[V(\Lambda_1) + V(\Lambda_2)]\} I(\Lambda_1, \Lambda_2), \tag{3.3}$$

due to the invariance of the measure $dM\, e^{-V(M)}$ under unitary transformations.

We will now show that

$$I(\Lambda_1, \Lambda_2; \beta) = \beta^{-N(N-1)/2} \prod_1^{N-1} p! \frac{\det(e^{\beta \lambda_{1,i} \lambda_{2,j}})}{\Delta(\Lambda_1)\Delta(\Lambda_2)}. \tag{3.4}$$

Let D be the unitary invariant Laplacian operator on Hermitian matrices

$$D \equiv \sum_i \frac{\partial^2}{\partial M_{ii}^2} + \frac{1}{2} \sum_{i<j} \left[\frac{\partial^2}{(\partial \operatorname{Re} M_{ij})^2} + \frac{\partial^2}{(\partial \operatorname{Im} M_{ij})^2} \right]. \tag{3.5}$$

Consider the propagator

$$f(t; M_1, M_2) = \langle M_1 | e^{-tD/2} | M_2 \rangle = \frac{1}{(2\pi t)^{N^2/2}} \exp\left[-\frac{1}{2t} \operatorname{tr}(M_1 - M_2)^2 \right], \tag{3.6}$$

a solution for t positive of the heat equation

$$\left(\frac{\partial}{\partial t} - \frac{1}{2} D_1 \right) f(t; M_1, M_2) = 0, \tag{3.7}$$

which reduces when $t \to 0$ to a δ function with respect to the measure introduced above. If $g(t, M)$ is a solution of the above equation for $t > 0$, which coincides for $t = 0$ with a given function $g(M)$ invariant under unitary transformations, i.e., a symmetric function of the eigenvalues of M, then

$$g(t, \Lambda_1) = C \int dU \int d\Lambda_2\, \Delta^2(\Lambda_2) f(t; \Lambda_1, U\Lambda_2 U^\dagger) g(\Lambda_2). \tag{3.8}$$

The constant C corresponds to the value appearing on the right-hand side of Eq. (2.7)

$$C = (2\pi)^{N(N-1)/2} \prod_1^N p!. \tag{3.9}$$

Consequently,

$$\Delta(\Lambda_1) g(t, \Lambda_1) = \int d\Lambda_2\, K(t; \Lambda_1, \Lambda_2)[\Delta(\Lambda_2) g(\Lambda_2)], \tag{3.10}$$

$$K(t; \Lambda_1, \Lambda_2) = C\Delta(\Lambda_1)\Delta(\Lambda_2) \int dU f(t; \Lambda_1, U\Lambda_2 U^\dagger),$$

which means that K is the evolution kernel for antisymmetric functions of the form

$$\xi(\Lambda) = \Delta(\Lambda) g(\Lambda). \tag{3.11}$$

The function ξ satisfies the equation

$$\frac{\partial \xi}{\partial t} = \frac{1}{2} \left(\frac{1}{\Delta(\Lambda)} \sum_{i=1}^N \frac{\partial}{\partial \lambda_i} \Delta^2(\Lambda) \frac{\partial}{\partial \lambda_i} \frac{\xi}{\Delta(\Lambda)} \right) = \frac{1}{2} \sum_i \left(\frac{\partial}{\partial \lambda_i} + \sum_{k \neq i} \frac{1}{\lambda_i - \lambda_k} \right) \left(\frac{\partial}{\partial \lambda_i} - \sum_{k \neq i} \frac{1}{\lambda_i - \lambda_k} \right) \xi = \frac{1}{2} \sum_i \frac{\partial^2 \xi}{\partial \lambda_i^2} - \sum_{k \neq l \neq i} \frac{1}{\lambda_i - \lambda_k} \frac{1}{\lambda_i - \lambda_l} \xi. \tag{3.12}$$

The last sum vanishes owing to the identity

$$\frac{1}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} + \frac{1}{(\lambda_2 - \lambda_3)(\lambda_2 - \lambda_1)} + \frac{1}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)} = 0.$$

Therefore, ξ fulfills

$$\frac{\partial \xi}{\partial t} = \frac{1}{2} \sum_{i=1}^N \frac{\partial^2}{\partial \lambda_i^2} \xi,$$

and is required to be antisymmetric. The kernel K of the corresponding evolution is then

$$K(t; \Lambda_1, \Lambda_2) = \frac{1}{(2\pi t)^{N/2}} \frac{1}{N!} \sum_{\mathscr{P}} (-1)^{\mathscr{P}} \times \exp\left[-\frac{1}{2t} \sum_i (\lambda_{1,i} - \lambda_{2,\mathscr{P}_i}) \right] = \frac{1}{(2\pi t)^{N/2}} \frac{1}{N!} \det\left\{ \exp\left[-\frac{1}{2t} (\lambda_{1,i} - \lambda_{2,j})^2 \right] \right\}. \tag{3.13}$$

If we compare this with (3.10) and (3.6), we find

$$\int dU \exp\left[-\frac{1}{2t} \operatorname{tr}(\Lambda_1 - U\Lambda_2 U^\dagger)^2 \right] = t^{N(N-1)/2} \prod_1^N p! \frac{\det[\exp -(1/2t)(\lambda_{1,i} - \lambda_{2,j})^2]}{\Delta(\Lambda_1)\Delta(\Lambda_2)}, \tag{3.14}$$

a formula equivalent to (3.4). The reader will recognize in

this derivation the features that made the planar approximation to quantum mechanics very simple in terms of fermionic wavefunctions (see [A]).

Let us now derive a series expansion for $I(M_1,M_2;\beta)$ in terms of the characters of the linear (or unitary) group, using a device due to Weyl. We recall that the irreducible representations of the group U(N) are characterized by a sequence of nondecreasing integers $n_0 \leqslant n_1 \leqslant \cdots n_{N-1}$, which for $n_0 \geqslant 0$ can be attached to a Young tableau.[10] We will consider here only polynomial representations, i.e., those such that the group factor U(N)/SU(N) is represented by $(\det U)^{n_0}$, $n_0 \geqslant 0$. The complete set of representations is obtained by relaxing the positivity condition on n_0. Let $U \to \mathscr{D}^{\{n\}}_{aa'}(U)$ be the corresponding representation with character $\chi_{\{n\}}$

$$\chi_{\{n\}}(U) = \sum_a \mathscr{D}^{\{n\}}_{aa}(U) = \frac{\det(\delta_i^{n_j+j})}{\det(\delta_i^j)}, \tag{3.15}$$

where δ_i are the eigenvalues of U. Denote by $d_{\{n\}}$ the dimension of this representation:

$$d_{\{n\}} = \chi_{\{n\}}(I). \tag{3.16}$$

Let first restrict our attention to the SU(N) group. This is, anyhow, the only part that enters the integral (3.2). In this case $n_0 = 0$. We have the orthogonality and completeness relations

$$\int dU\, \mathscr{D}^{\{n\}}_{a_1a_1'}(U)\mathscr{D}^{\{n'\}*}_{a_2a_2'}(U) = \frac{1}{d_{\{n\}}}\delta^{\{n\},\{n'\}}\delta_{a_1a_2}\delta_{a_1'a_2'}, \tag{3.17}$$

$$\sum_{\{n\},a,a'} d_{\{n\}}\mathscr{D}^{\{n\}}_{aa'}(U)\mathscr{D}^{\{n\}*}_{aa'}(U') = \delta(U,U').$$

Let U_1 and U_2 stand for two arbitrary elements in SU(N). We have

$$\int dU \exp(\beta\, \mathrm{tr} U_1 U U_2 U^\dagger)$$

$$= \int dV\, e^{\beta\, \mathrm{tr}V} \int dU \delta(V, U_1 U U_2 U^\dagger)$$

$$= \int dV\, e^{\beta\, \mathrm{tr}V} \sum_{\{n\},a,a'} \mathscr{D}^{\{n\}}_{aa'}(U_1)\chi_{\{n\}}(U_2)\mathscr{D}^{\{n\}*}_{aa'}(V).$$

The integral over V being invariant under the adjoint action $V \to UVU^\dagger$, we can replace $\mathscr{D}^{\{n\}*}_{aa'}(V)$ by $(\delta_{aa'}/d_{\{n\}})\chi^*_{\{n\}}(V)$.

Now

$$\int dV\, \mathrm{tr}V)^p \chi^*_{\{n\}}(V) = \delta_{p,|n|}\sigma_{\{n\}}, \tag{3.18}$$

where $|n| = \Sigma_0^{N-1} n_i$ and $\sigma_{\{n\}}$ is the number of times the representation $\mathscr{D}^{\{n\}}(U)$ occurs in the tensor product $\overset{|n|}{\otimes} U$.

This can also be interpreted as the number of distinct ways of constructing piece by piece the Young tableau for the representation $\mathscr{D}^{\{n\}}$ while respecting the rules for such tableaux. (Therefore $\sigma_{\{n\}}$ is nothing but the dimension of the representation of the permutation group on $|n|$ objects pertaining to the same tableau. For a proof see Ref. 10.) It follows that

$$= \sum_{\{n\}} \frac{\beta^{|n|}}{|n|!}\frac{\sigma_{\{n\}}}{d_{\{n\}}}\chi_{\{n\}}(U_1)\chi_{\{n\}}(U_2). \tag{3.19}$$

This result has been derived for U_1, $U_2 \in$ SU(N), and the sum on the right-hand side runs only over representations with $n_0 = 0$. It can readily be extended to U_1, $U_2 \in$ U(N) provided we reintroduce all representations with $n_0 \geqslant 0$. $\chi_{\{n\}}(U)$ is a polynomial in the matrix elements of U and therefore can be continued as a function of an arbitrary $N \times N$ matrix. By analytic continuation, we therefore reach the conclusion that

$$I(M_1,M_2;\beta) = \sum_{\{n\}} \frac{\beta^{|n|}}{|n|!}\frac{\sigma_{\{n\}}}{d_{\{n\}}}\chi_{\{n\}}(M_1)\chi_{\{n\}}(M_2). \tag{3.20}$$

A similar formula could in fact be directly obtained by expanding the numerator of the right-hand side of Eq. (3.4) in powers of the eigenvalues. Comparison with (3.20) yields

$$\sigma_{\{n\}} = |n|! d_{\{n\}} \prod_0^{N-1} p! \Big/ \prod_0^{N-1} (n_p + p)!, \tag{3.21}$$

and we recall the Weyl formula

$$d_{\{n\}} = \Delta(n_{N-1} + N - 1, n_{N-2} + N - 2, \ldots, n_0) \Big/ \prod_0^{N-1} p!,$$

where Δ is the discriminant used througout our previous discussion. Table I gives explicit formulas for the characters up to $|n| = 4$ in terms of traces of powers of the matrix M. We check, of course, that

$$d_{\{n\}} = \chi_{\{n\}}(I) \quad \text{and} \quad (\mathrm{tr}M)^p = \sum_{\substack{\{n\} \\ |n|=p}} \sigma_{\{n\}}\chi_{\{n\}}(M). \tag{3.22}$$

TABLE I. Characters of the linear group up to $|n| = 4$.

Young tableau	$\chi_{\{n\}}(A)$	$d_{\{n\}}$	$\sigma_{\{n\}}$
□	$\mathrm{tr}A$	N	1
□□	$\frac{1}{2}[(\mathrm{tr}A)^2 + \mathrm{tr}A^2]$	$\frac{1}{2}N(N+1)$	1
▯	$\frac{1}{2}[(\mathrm{tr}A)^2 - \mathrm{tr}A^2]$	$\frac{1}{2}N(N-1)$	1
□□□	$\frac{1}{6}[(\mathrm{tr}A)^3 + 2\,\mathrm{tr}A^3 + 3\mathrm{tr}A\,\mathrm{tr}A^2]$	$\frac{1}{6}N(N+1)(N+2)$	1
▯□	$\frac{1}{3}[(\mathrm{tr}A)^3 - \mathrm{tr}A^3]$	$\frac{1}{3}N(N+1)(N-1)$	2
▯	$\frac{1}{6}[(\mathrm{tr}A)^3 + 2\,\mathrm{tr}A^3 - 3\mathrm{tr}A\,\mathrm{tr}A^2]$	$\frac{1}{6}N(N-1)(N-2)$	1
□□□□	$\frac{1}{24}[(\mathrm{tr}A)^4 + 6\,\mathrm{tr}A^4 + 3(\mathrm{tr}A^2)^2 + 6\,\mathrm{tr}A^2(\mathrm{tr}A)^2 + 8\,\mathrm{tr}A^3\,\mathrm{tr}A]$	$\frac{1}{24}N(N+1)(N+2)(N+3)$	1
▯□□	$\frac{1}{8}[(\mathrm{tr}A)^4 - 2\,\mathrm{tr}A^4 - (\mathrm{tr}A^2)^2 + 2\,\mathrm{tr}A^2(\mathrm{tr}A)^2]$	$\frac{1}{8}N(N+1)(N+2)(N-1)$	3
▯▯	$\frac{1}{12}[(\mathrm{tr}A)^4 - 4\,\mathrm{tr}A^3\,\mathrm{tr}A + 3(\mathrm{tr}A^2)^2]$	$\frac{1}{12}N^2(N+1)(N-1)$	2
▯□	$\frac{1}{8}[(\mathrm{tr}A)^4 + 2\,\mathrm{tr}A^4 - (\mathrm{tr}A^2)^2 - 2\,\mathrm{tr}A^2(\mathrm{tr}A)^2]$	$\frac{1}{8}N(N+1)(N-1)(N-2)$	3
▯	$\frac{1}{24}[(\mathrm{tr}A)^4 - 6\,\mathrm{tr}A^4 + 3(\mathrm{tr}A^2)^2 - 6\,\mathrm{tr}A^2(\mathrm{tr}A)^2 + 8\,\mathrm{tr}A^3\,\mathrm{tr}A]$	$\frac{1}{24}N(N-1)(N-2)(N-3)$	1

TABLE II. The coefficients $X_k(A,B)$.

k	$X_k(A,B)$
1	$\langle A\rangle\langle B\rangle$
2	$f_2(A)f_2(B)$
3	$f_3(A)f_3(B)$
4	$f_4(A)f_4(B) - 4\dfrac{f_2^2(A)}{2}\dfrac{f_2^2(B)}{2}$
5	$f_5(A)f_5(B) - 5f_2(A)f_3(A)f_2(B)f_3(B)$
6	$f_6(A)f_6(B) - 6\left\{f_4(A)f_2(A)\left[f_4(B)f_2(B) + \dfrac{f_3^2(B)}{2}\right] + \dfrac{f_3^2(A)}{3!}\left[f_4(B)f_2(B) + 2\dfrac{f_3^2(B)}{2!} + 2\dfrac{f_2^3(B)}{3!}\right] + \dfrac{f_3^2(A)}{3!}\left[2\dfrac{f_2^3(B)}{2!} - 12\dfrac{f_2^3(B)}{3!}\right]\right\}$
7	$f_7(A)f_7(B) - 7\left\{f_5(A)f_2(A)[f_5(B)f_2(B) + f_4(B)f_3(B)] + f_4(A)f_3(A)\left[f_5(B)f_2(B) + 2f_4(B)f_3(B) + 2f_3(B)\dfrac{f_2^2(B)}{2!}\right] + f_3(A)\dfrac{f_2^2(A)}{2!}\left[2f_4(B)f_3(B) - 12f_3(B)\dfrac{f_2^2(B)}{2!}\right]\right\}$
8	$f_8(A)f_8(B) - 8\left\{f_6(A)f_2(A)\left[f_6(B)f_2(B) + f_5(B)f_3(B) + \dfrac{f_4^2(B)}{2!}\right] + f_5(A)f_3(A)\left[f_6(B)f_2(B) + 2f_5(B)f_3(B) + 2\dfrac{f_4^2(B)}{2!} + 2\dfrac{f_3^2(B)}{2!}f_2(B) + 2f_4(B)\dfrac{f_2^2(B)}{2}\right] + \dfrac{f_4^2(A)}{2!}\left[f_6(B)f_2(B) + 2f_5(B)f_3(B) + 3\dfrac{f_4^2(B)}{2!} + 4\dfrac{f_3^2(B)}{2!}f_2(B) + 2f_4(B)\dfrac{f_2^2(B)}{2!} + 6\dfrac{f_2^4(B)}{4!}\right] + \dfrac{f_3^2(A)}{2!}f_2(A)\left[2f_5(B)f_3(B) + 4\dfrac{f_4^2(B)}{2!} - 20\dfrac{f_3^2(B)}{2!}f_2(B) - 12f_4(B)\dfrac{f_2^2(B)}{2!} - 48\dfrac{f_2^4(B)}{4!}\right] + f_4(A)\dfrac{f_2^2(A)}{2!}\left[2f_5(B)f_3(B) + 2\dfrac{f_4^2(B)}{2!} - 12\dfrac{f_3^2(B)}{2!}f_2(B) - 14f_4(B)\dfrac{f_2^2(B)}{2!}\right] + \dfrac{f_4^2(A)}{4!}\left[6\dfrac{f_4^2(B)}{2!} - 48\dfrac{f_3^2(B)}{2!}f_2(B) + 360\dfrac{f_2^4(B)}{4!}\right]\right\}$

We have now two exact expressions for the kernel $I(M_1,M_2;\beta)$ given in Eqs. (3.4) and (3.20). For our purpose, we are also interested in an expansion of $\ln[I(M_1,M_2;\beta)]$ for large N assuming the eigenvalues of M_1 and M_2 to be of order $N^{1/2}$. Without loss of generality we take M_1 and M_2 diagonal and rescale them as $M_1 = \sqrt{N}A$ and $M_2 = \sqrt{N}B$ with A and B of order unity. We look for the dominant term

$$X(A,B;\beta) = \lim_{N\to\infty}\frac{1}{N^2}\ln\left[I(\sqrt{N}A,\sqrt{N}B,\beta)\right] = \lim_{N\to\infty}\frac{1}{N^2}\ln\left[\int dU\, e^{N\beta\,\mathrm{tr}(AUBU^+)}\right], \tag{3.23}$$

with

$$A = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & & 0 \\ & \ddots & \\ 0 & & b_n \end{pmatrix}.$$

The quantity X admits a series expansion in powers of β

$$X(A,B;\beta) = \sum_1^\infty \frac{\beta^k}{k} X_k(A,B), \tag{3.24}$$

where $X_k(A,B) = X_k(B,A)$ is a symmetric function of the a_i and b_i, homogeneous of degree k. It is given in terms of the quantities $\langle A^p\rangle \equiv (1/N)\,\mathrm{tr}A^p$, $\langle B^p\rangle$. By singling out $\langle A\rangle$ which can readily be factored in I, it will be useful to use rather the mean values $e_{(n)}(A)$, $e_n(B)$:

$$\begin{aligned} e_0(A) &= 1, \quad e_1 \equiv 0, \\ e_p(A) &= \langle (A - \langle A\rangle)^p\rangle \\ &= \frac{1}{N}\mathrm{tr}\left(A - \frac{1}{N}\mathrm{tr}A\right)^p, \quad p \geqslant 2, \end{aligned} \tag{3.25}$$

or even better the "connected" ones $f_n(A), f_n(B)$. The relation between these two basis, already discussed in [A] for Green's functions is most easily expressed through the generating functions

$$\begin{aligned} \phi(j;A) &= 1 + \sum_2^\infty j^k e_k(A), \\ \psi(z;A) &= 1 + \sum_2^\infty z^k f_k(A), \\ \phi(j;A) &= \psi(z[j;A];A), \\ z[j;A] &= j\phi(j;A), \end{aligned} \tag{3.26}$$

or, more explicitly,

$$e_k(A) = \sum_{\substack{\{r_q\geqslant 0\} \\ \sum qr_q = k}} \frac{k!}{(k+1-\sum r_q)!}\frac{[f_2(A)]^{r_2}}{r_2!}\frac{[f_3(A)]^{r_3}}{r_3!}\cdots, \tag{3.27}$$

$$f_k(A) = -\sum_{\substack{\{r_q\geqslant 0\} \\ \sum qr_q = k}} \frac{(k+\sum r_q - 2)!}{(k-1)!}\frac{[-e_2(A)]^{r_2}}{r_2!} \times \frac{[-e_3(A)]^{r_3}}{r_3!}\cdots.$$

We can start grinding the coefficients $X_k(A,B)$ using the expansion given in (3.20). The results up to order 8 are displayed in Table II. To expose some properties of this expansion, we shall write for X differential equations similar to those discussed at the beginning of this section. Let $X^{(N)}$ be equal to $(1/N^2)\ln[I(\sqrt{NA},\sqrt{NB};\beta)]$, i.e., to the same quantity as X before going to the limit $N\to\infty$. We have

$$e^{N^2X^{(N)}} = \frac{\Pi_1^{N-1} p!}{(\beta N)^{N(N-1)/2}} \frac{\det(e^{N\beta a_i b_j})}{\Delta(A)\Delta(B)}. \tag{3.28}$$

The quantity

$$[\Delta(A)e^{N^2X^{(N)}}]^{-1}\sum_{i=1}^{N}\left(\frac{\partial}{\partial a_i}\right)^p \Delta(A)e^{N^2X^{(N)}},$$

is obviously equal to $N^p\beta^p\Sigma\, b_i^p$ and we therefore derive the identity

$$\beta^p\langle B^p\rangle = \frac{1}{\Delta(A)}\frac{1}{N}\sum_i\left(\frac{1}{N}\frac{\partial}{\partial a_i} + N\frac{\partial X^{(N)}}{\partial a_i}\right)^p \Delta(A).$$

Since $N\partial X^{(N)}/\partial a_i$ is of order unity, we may omit in the large-N limit the action of derivatives on it, when expanding this pth power. To leading order,

$$\beta^p\langle B^p\rangle = \sum_{s=1}^{p}\frac{p!}{s!(p-s)!}\frac{1}{N^{p-s+1}} \times \sum_{i\neq j_1\neq\cdots\neq j_p}\frac{(N\partial X/\partial a_i)^s}{(a_i-a_{j_1})\cdots(a_i-a_{j_p})}, \tag{3.29}$$

where the term $s=0$ is absent since $\Delta^{-1}\Sigma_i\partial_i^p\Delta \equiv 0$ for $p<N$. For $p=1$ this yields

$$\beta\langle B\rangle = \frac{1}{N}\sum_i N\frac{\partial X}{\partial a_i},$$

which means

$$X_1 = \beta\langle A\rangle\langle B\rangle, \qquad \frac{1}{N}\sum_i N\frac{\partial X_k}{\partial a_i} = 0, \quad k>1. \tag{3.30}$$

Thus, for $k>1$, X_k which is a homogeneous function of A of degree k may be written in terms of the $e_s(A)$, $s\leqslant k$, which all satisfy

$$\frac{1}{N}\sum_i N\frac{\partial e_s(A)}{\partial a_i} = 0.$$

Defining $\tilde{X}$ through

$$X = \tilde{X} + \beta\langle A\rangle\langle B\rangle, \tag{3.31}$$

we find

$$\beta^p e_p(B) = \sum_{s=1}^{p}\frac{p!}{s!(p-s)!}\frac{1}{N^{p-s+1}} \times \sum_{i\neq j_1\neq\cdots\neq j_p}\frac{(N\partial\tilde{X}/\partial a_i)^s}{(a_i-a_{j_1})\cdots(a_i-a_{j_p})}. \tag{3.32}$$

This infinite set of equations determines the functions X_k recursively. The algebra becomes rapidly quite cumbersome, and we did not succeed in finding a simple algorithm. We shall, however, indicate some simple features. Let us first focus on the first terms ($s=1$) of the right-hand side of Eq. (3.32). It reads

$$\Delta^{(p)}\tilde{X} \equiv \frac{p}{N^p}\sum_{i\neq j_1\neq\cdots\neq j_{p-1}}\frac{N\partial\tilde{X}/\partial a_i}{(a_i-a_{j_1})\cdots(a_i-a_{j_{p-1}})}. \tag{3.33}$$

One may show that

$$\Delta^{(p)}e_k(A) = k\sum_{r_1+r_2+\cdots+r_p=k} e_{r_1}\cdots e_{r_p}, \quad k=2,3,\cdots, \tag{3.34}$$

where the right-hand side is zero for $k<p$. On the generating function $\phi(j;A)$ of Eq. (3.26)

$$\Delta^{(p)}\phi(j;A) = j\frac{\partial}{\partial j}(j^p\phi^p). \tag{3.35}$$

Let us now show that $\Delta^{(p)}$ has a very simple action on the connected $f_k(A)$. Since it is a derivative, we have

$$\Delta^{(p)}\phi(j;A) = \Delta^{(p)}\psi(z[j;A];A) = \frac{\partial\psi}{\partial z}\Delta^{(p)}z[j;A] + \Delta^{(p)}\psi(z;A)\Big|_{z=z[j;A]},$$

but from (3.26)

$$\Delta^{(p)}z[j;A] = j\Delta^{(p)}\phi(j;A)$$

and

$$j\frac{\partial z}{\partial j}\left(1-j\frac{\partial\psi}{\partial z}\right) = j\frac{\partial z}{\partial j}\frac{\partial}{\partial z}(z-j\psi)+j\psi = z.$$

Hence

$$\Delta^{(p)}\psi(z;A) = \left(1-j\frac{\partial\psi}{\partial z}\right)\Delta^{(p)}\phi(j;A) = \left(1-j\frac{\partial\psi}{\partial z}\right)j\frac{\partial}{\partial j}(j\phi)^p = z\frac{\partial}{\partial z}z^p = pz^p, \tag{3.36}$$

or equivalently

$$\Delta^{(p)}f_k(A) = p\delta_{kp}. \tag{3.37}$$

This now suggests to rewrite Eq. (3.32) as

$$\beta^p\sum_{\substack{\{r_q\}\\ \Sigma qr_q=p}}\frac{p!}{(p+1-\Sigma r_q)!}\prod_{q\geqslant 2}\frac{(f_q(B))^{r_q}}{r_q!} = \Delta^{(p)}\tilde{X} + \sum_{s=2}^{p}\binom{p}{s}\frac{1}{N^{p-s+1}} \times \sum_{i\neq j_1\neq\cdots\neq j_p}\frac{(N\partial\tilde{X}/\partial a_i)^s}{(a_i-a_{j_1})\cdots(a_i-a_{j_p})},$$

and so solve for $\tilde{X}$ according to its increasing degree in the $f_k(B)$. To lowest order (linear terms) one has

$$\Delta^{(p)}\tilde{X}^{(1)} = \beta^p f_p(B), \tag{3.38}$$

and hence

$$\tilde{X}^{(1)} = \frac{\beta^p}{p}f_p(A)f_p(B)+\cdots, \quad p=2,3,\cdots,$$

where the triple dots stand for terms independent of $f_p(A)$. $\tilde{X}^{(1)}$ is necessarily of the form

$$\tilde{X}^{(1)} = \sum_{p=2}\frac{\beta^p}{p}f_p(A)f_p(B). \tag{3.39}$$

β^2 can be extracted self-consistently from the above equations.

An alternative method of evaluation reveals the connection of our problem with matrix elements of the free evolution operator between Slater determinants, i.e., wavefunctions for large Fermi systems. To this end, we make the following change of variables:

$$\beta = \frac{1}{1+t/2}, \quad g_p = \frac{g'_p}{2(\tfrac{1}{2}+1/t)^p}, \qquad M_{1,2} = (\tfrac{1}{2}+1/t)^{1/2} M'_{1,2}. \tag{4.9}$$

In this way $Z(g,\beta)$ takes the form

$$Z(g,\beta) = (\tfrac{1}{2}+1/t)^{N^2}(2\pi t)^{N^2/2}\int dM\, e^{-V(M)} W(g',t), \tag{4.10}$$

$$W(g',t) = (2\pi t)^{-N^2/2}\left(\int dM\, e^{-V(M)}\right)^{-1}\int dM_1\, dM_2 \times \exp\{-\tfrac{1}{2}[V(M_1,g') + V(M_2,g')] - (1/2t)\,\mathrm{tr}(M_1-M_2)^2\}. \tag{4.11}$$

Henceforth we drop the prime on the coupling constants. After integration over the unitary group we have

$$\begin{aligned} W(g,t) &= K\frac{1}{(2\pi t)^{N/2}}\int d\Lambda_1\, d\Lambda_2\, \Delta(\Lambda_1)\Delta(\Lambda_2) \times \exp[-\tfrac{1}{2}V(\Lambda_1) - \tfrac{1}{2}V(\Lambda_2)] \times \det\left\{\exp\left[-\frac{1}{2t}(\lambda_{1,i}-\lambda_{2,j})^2\right]\right\} \\ &= K'\frac{1}{(2\pi t)^{N/2}}\int d\Lambda_1\, d\Lambda_2 \det[\psi_k(\lambda_{1,l})] \times \det\left\{\exp\left[-\frac{1}{2t}(\lambda_{1,i}-\lambda_{2,j})^2\right]\right\} \times \det[\psi_m(\lambda_{2,n})], \end{aligned} \tag{4.12}$$

with constants K and K' independent of t adjusted to insure that $W(g,0) = 1$. We have introduced the orthonormal functions of Sec. 2

$$\psi_k(\lambda) = \mathscr{P}_k(\lambda) e^{-1/2V(\lambda)}, \tag{4.13}$$

with V as in (2.10), and the normalized polynomials $\mathscr{P}_k$ defined in (2.29). The determinants are $N\times N$ with the index of the orthogonal functions running from 0 to $N-1$. Finally we find

$$\begin{aligned} W(g,t) &= \det\int d\lambda_1\, d\lambda_2\, \psi_k(\lambda_1) \times \frac{\exp[-(\lambda_1-\lambda_2)^2/2t]}{(2\pi t)^{1/2}}\psi_l(\lambda_2) \\ &= \det(\langle k|e^{-ht}|l\rangle). \end{aligned} \tag{4.14}$$

Here h is the free Hamiltonian

$$h = -\frac{1}{2}\frac{d^2}{d\lambda^2}, \tag{4.15}$$

$W(g,t)$ has been written as the matrix element of the free evolution operator in the ground state of N "fermions" occupying the levels $\psi_0,\psi_1,\ldots,\psi_{N-1}$. In the large-$N$ limit, we define

$$\varphi(g,t) = -\lim_{N\to\infty}\frac{1}{N^2}\ln[W(g,t)] \equiv \sum_1^\infty \frac{(-1)^{n-1}}{n!}\varphi_n(g)t^n. \tag{4.16}$$

For $t\to\infty$, φ behaves as $\frac{1}{2}\ln t$, while for $g=0$

$$\varphi(0,t) = \tfrac{1}{2}\ln(1+t/4). \tag{4.17}$$

Again, we are unable to obtain $\varphi(g,t)$ except as a power series in t. To see this, we introduce the projectors

$$P = \sum_0^{N-1}\psi_k\otimes\psi_k, \quad Q = I - P, \tag{4.18}$$

in the Hilbert space $\mathscr{L}^2(R)$, which enable us to express W as an infinite determinant

$$W = \det(Q + Pe^{-ht}P). \tag{4.19}$$

We then decompose h in a block form adapted to the ψ_k basis:

$$h = PAP + PBQ + QB^\dagger P + QCQ. \tag{4.20}$$

Thus[11]

$$\begin{aligned} -\ln W &= -\mathrm{tr}\{\ln[I - P(I-e^{-ht})P]\} \\ &= t\,\mathrm{tr}A - \frac{t^2}{2!}\mathrm{tr}BB^\dagger + \frac{t^3}{3!}\mathrm{tr}B(CB^\dagger - B^\dagger A) \\ &\quad - \frac{t^4}{4!}\mathrm{tr}\{B[C(CB^\dagger - B^\dagger A) - (CB^\dagger - B^\dagger A)A] - 2BB^\dagger BB^\dagger\} + \cdots. \end{aligned} \tag{4.21}$$

Except for the first, all terms in this series involve for large N the matrix elements of h close to the "Fermi level" N. To obtain $\varphi(g,t)$, we divide the above expression by N^2 and look for the limiting behavior. With an implicit limit sign, the first term reads

$$\begin{aligned} \varphi_1(g) &= \frac{1}{N^2}\mathrm{tr}A = \frac{1}{N^2}\sum_{k=0}^{N-1}\langle k|h|k\rangle \\ &= \frac{1}{N^2}\sum_{k=0}^{N-1}\frac{1}{2}\int d\lambda\left[\frac{d}{d\lambda}(e^{-1/2V(\lambda)}\mathscr{P}_k(\lambda))\right]^2 \\ &= \frac{1}{8N^2}\sum_{k=0}^{N-1}\int d\lambda\,(V'\psi_k)^2. \end{aligned} \tag{4.22}$$

With the notations of (2.33), it follows that

$$\varphi_1(g) = \tfrac{1}{8}\int_{-2a}^{2a} d\lambda\, u(\lambda)[v'(\lambda)]^2, \tag{4.23}$$

where $Nv(\lambda) = V(N^{1/2}\lambda)$. Explicitly for the quartic interaction

$$\begin{aligned} \varphi_1(g) &= \frac{1}{8}\int_{-2a}^{2a}\frac{d\lambda}{2\pi}(4a^2-\lambda^2)^{1/2} \times (1+8ga^2+4g\lambda^2)(\lambda+4g\lambda^3)^2 \\ &= \frac{1}{8}\frac{(1-a^2)(4-a^2)}{36a^2}. \end{aligned} \tag{4.24}$$

The computation can be carried further. For instance, for the same interaction to second order in t, we find after tedious calculations

$$\varphi_2(g) = \frac{1}{32}\left\{1 + \frac{1-a^2}{a^4}\left[1 - \frac{1}{3}a^2 + \left(\frac{1-a^2}{3}\right)^3\right]\right\}. \tag{4.25}$$

This program may be pursued order by order: The term $\tilde{X}^{(n)}$ of degree n in $f(B)$ satisfies a set of equations

$$\Delta^{(p)}\tilde{X}^{(n)} = \mathscr{F}^{(p,n)}(\tilde{X}^{(1)},\ldots,\tilde{X}^{(n-1)}, f(A), f(B)),$$

which may be integrated owing to (3.37). In particular a compact expression may be given to the terms quadratic in *both* $f(A)$ and $f(B)$. To summarize

$$X(A,B;B)$$

$$= \beta\langle A\rangle\langle B\rangle + \sum_{n=2}^{\infty}\beta^n\Bigg\{\frac{1}{n}f_n(A)f_n(B) - \sum_{\substack{\{r_p\geqslant 0\}\\ \sum pr_p = n\\ \sum r_p = 2}}\sum_{\substack{\{s_q\geqslant 0\}\\ \sum qs_q = n\\ \sum s_q = 2}}\Bigg[\prod_{p\geqslant 2}\frac{f_p^{r_p}(A)}{r_p!} \times\prod_{q\geqslant 2}\frac{f_q^{s_q}(B)}{s_q!}\min(p-1,q-1)\Bigg] + \cdots\Bigg\}, \tag{3.40}$$

where $\min(p-1,q-1)$ runs over the indices p or q appearing in the term at hand, and the three dots stand for terms at least cubic in $f(A)$ or $f(B)$. Of course, this general expression coincides with the first few terms listed in Table II.

As a last remark, we observe that the Cauchy determinant

$$\det\left(\frac{1}{1-x_i y_j}\right) = \frac{\Delta(x)\Delta(y)}{\Pi_{i,j}(1-x_i y_j)}, \tag{3.41}$$

can be used to obtain a reproducing kernel for $I(A,B;\beta)\sim\exp[N^2X(A,B;\beta)]$ in the form

$$I(A,B;\beta) = \frac{1}{N!}\oint\prod_1^N\left(\frac{dz_k}{2\pi i z_k}\right)\Delta(Z)\Delta(Z^{-1}) \times\exp\left\{N^2\sum_1^{\infty}\frac{1}{s}\langle Z^{-s}\rangle\langle A^s\rangle\right\}I(Z,B;\beta),$$

where Z is a diagonal matrix. The reader will recognize that the integral runs over the equivalence classes of the unitary group $U(N)$.

4. THE TWO-MATRIX PROBLEM

We now focus our attention on the quantity Z of Eqs. (3.1) or (3.3) using the closed form obtained in (3.4) for the integral over the unitary group

$$Z = \int dM_1\, dM_2 \exp[-V(M_1) - V(M_2) + \beta\,\mathrm{tr}M_1M_2]$$
$$= \frac{(2\pi)^{N(N-1)}}{N\Pi_1^N p!}\beta^{-N(N-1)/2}\int d\Lambda_1\, d\Lambda_2\,\Delta(\Lambda_1)\Delta(\Lambda_2) \times\exp[-V(\Lambda_1) - V(\Lambda_2)]\det[\exp(\beta\lambda_{1,i}\lambda_{2,j})], \tag{4.1}$$

with

$$V(\Lambda) = \frac{1}{2}\sum_i\lambda_i^2 + \sum_{p\geqslant 2}\frac{g_p}{N^{p-1}}\sum_i\lambda_i^{2p}. \tag{4.2}$$

We can deal with this expression in two ways. The first one is a small β expansion where we substitute in the exponent the series in β discussed at the end of the previous section. The alternative strong coupling expansion will be presented afterwards.

Thus we write

$$Z = \frac{(2\pi)^{N(N-1)}}{(\Pi_i^N p!)^2}\int d\Lambda_1\, d\Lambda_2\,\Delta^2(\Lambda_1)\Delta^2(\Lambda_2) \times\exp\left[-V(\Lambda_1) - V(\Lambda_2) + N^2X\left(\frac{\Lambda_1}{N^{1/2}}, \frac{\Lambda_2}{N^{1/2}};\beta\right)\right]. \tag{4.3}$$

For fixed Λ_2, this is an integral over Λ_1 with an "effective potential" of a generalized type involving only symmetric functions. We can therefore use the techniques of Sec. 2, which are equivalent to the saddle-point method of [A]. Symmetry under the interchange $\Lambda_1\longleftrightarrow\Lambda_2$ implies that the coefficients of this generalized potential which depends only on Λ_2 can be determined self-consistently by requiring that the symmetric functions of both matrices be equal at the saddle point. We use the work "generalized potential" since it contains arbitrary powers of the symmetric functions, in contrast with the original one (4.2). We may speak in that case of "nonlocality" in the index of eigenvalues. To illustrate this point, we shall compute

$$\mathscr{E}(g,\beta) = -(1/N^2)\ln[Z(g,\beta)/Z(g,0)], \tag{4.4}$$

to fourth order in β. Rescaling Λ into $\Lambda\sqrt{N}$, we find that $\mathscr{E}(\beta)$ is the saddle-point value of the functional:

$$\mathscr{E} = \left\{\int_0^1 dx\, v[\lambda(x)] - \int_0^1\int_0^1 dx\, dy\ln|\lambda(x) - \lambda(y)|\right\} + [\lambda(x)\longleftrightarrow\mu(x)] - \beta\langle\lambda\rangle\langle\mu\rangle - \frac{\beta^2}{3}f_3(\lambda)f_3(\mu) - \frac{\beta^3}{3}f_3(\lambda)f_3(\mu) - \frac{\beta^4}{4} \times[f_4(\lambda)f(\mu) - f_2^2(\lambda)f_2^2(\mu)] - \cdots. \tag{4.5}$$

The rescaled eigenvaues have been rearranged as increasing functions of the reduced index $x = i/N$. A continuous limit as $N\to\infty$ is understood.

The notations of Sec. 3 have been generalized to mean $\langle\lambda^s\rangle = \int_0^1 dx\lambda^s(x)$, $e_s(\lambda) = \langle(\lambda - \langle\lambda\rangle)^s\rangle$, and $f_s(\lambda)$ is related to $e_s(\lambda)$ as in (3.26)–(3.27). If $u(\lambda)$ denotes the density of eigenvalues, we obtain the saddle-point equation

$$0 = -v'(\lambda) + 2\fint\frac{d\lambda' u(\lambda')}{\lambda - \lambda'} + \beta\langle\mu\rangle + \beta^2 f_2(\mu)(\lambda - \langle\lambda\rangle) + \beta^3 f_3(\mu)[(\lambda - \langle\lambda\rangle)^2 - \langle(\lambda - \langle\lambda\rangle)^2\rangle] + \beta^4 f_4(\mu)[(\lambda - \langle\lambda\rangle)^3 - \langle(\lambda - \langle\lambda\rangle)^3\rangle] - \beta^4[2f_4(\mu) + f_2^2(\mu)]f_2(\lambda)[\lambda - \langle\lambda\rangle] - \cdots.$$

For definiteness, let us consider the φ^4 theory with $v(\lambda) = \lambda^2/2 + g\lambda^4$. A consistent Ansatz assumes the odd mean values to vanish, viz., $\langle\lambda^{2s+1}\rangle = \langle\mu^{2s+1}\rangle = 0$. The lowest order in β for $\mathscr{E}(g,\beta)$ is β^2, and we readily find

$$\mathscr{E}(g,\beta) = -\frac{\beta^2}{2}\left[1 - 4g\frac{dE(g)}{dg}\right]^2 + O(\beta^4) = -\frac{\beta^2}{8}[a^2(4-a^2)]^2 + \cdots, \tag{4.7}$$

with $E(g)$ given by (2.27)

$$E(g) = \tfrac{1}{24}(a^2-1)(9-a^2) - \tfrac{1}{2}\ln a^2, \qquad 12ga^4 + a^2 - 1 = 0. \tag{4.8}$$

This can be checked diagrammatically to the first few orders in g. With more algebra, the coefficients of higher powers in

Even though this direct method lacks some elegance, it is, however, very effective.[12]

The conclusions to be drawn from this large amount of algebra were already presented in the Introduction. The planar approximation seems a very nontrivial one, and even in the simplest case discussed in this paper, no simple algorithm was found. But it could well be that, for deeper geometric reasons, the same approximation is more tractable in the case of gauge fields.

ACKNOWLEDGMENTS

We gratefully acknowledge numerous discussions with G. Parisi and E. Brézin, who informed us of similar considerations by E. Witten.

[1]G.'t Hooft, Nucl. Phys. B **72**, 461 (1974); B **75**, 461 (1974).
[2]G. Veneziano in Proc. 19th Int. Conf. High Energy Physics, Tokyo 1978, A. Migdal, unpublished.
[3]E. Witten, Harvard Preprint HUTP 79/A007.
[4]H. Wegner, to be published.
[5]W.T. Tutte, Can. J. Math. **14**, 21 (1962).
[6]J. Koplik, A. Neveu, and S. Nussinov, Nucl. Phys. B **123**, 109 (1977).
[7]E. Brézin, I. Itzykson, G. Parisi, and J.-B. Zuber, Commun. Math. Phys. **59**, 35 (1978), hereafter referred to as [A].
[8]D. Bessis, Saclay Preprint 79, 23, to be published in Comm. Math. Phys.
[9]G. Parisi, Private communication.
[10]H. Weyl, *The Classical Groups* (Princeton Univ. Press, Princeton, N.J., 1946).
[11]A complete expression may be found in R. Balian and E. Brézin, Nuov. Cimento B **64**, 37 (1969).
[12]Note added in proof: The two matrix problem has finally been solved by M.L. Mehta, Saclay preprint **79**, 124.

On the SU_2 unit tensor[a)]

Maurice Kibler[b)]
Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

Geneviève Grenet
Institut de Physique Nucléaire, Université Claude Bernard Lyon-I et IN2P3, 69621 Villeurbanne, France

(Received 1 August 1979; accepted for publication 19 October 1979)

This paper deals with the $SU_2 \supset G^*$ unit tensor operators $t_{k\mu\alpha}$. In the case where the spinor point group G^* coincides with U_1, then $t_{k\mu\alpha}$ reduces (up to a constant) to the (Wigner–Racah–Schwinger) tensor operator $t_{kq\alpha}$, an operator which produces an angular momentum state $|j+\alpha, m+q\rangle$ when acting upon the state $|jm\rangle$. We first investigate those general properties of $t_{k\mu\alpha}$ which are independent of their realization. We then turn our attention to realizations of $t_{k\mu\alpha}$ in terms of two pairs of boson creation and annihilation operators. This leads us to look at the Schwinger calculus (found to be connected to the de Sitter algebra $so_{3,2}$) relative to one angular momentum or two coupled angular momenta. As a by-product, we give a procedure for producing recursion relationships between $SU_2 \supset U_1$ Wigner coefficients. Expressions for $t_{kq\alpha}$, which cover the cases k integer and half-an-odd-integer, are derived in terms of boson operators. When k is integer, the latter expressions can be rewritten in the enveloping algebra of so_3 or $so_{3,2}$ according to as $\alpha = 0$ or $\alpha \neq 0$. Finally, we study in two appendices some of the properties of (i) the Wigner and Racah operators for an arbitrary compact group and (ii) the $SU_2 \supset G^*$ coupling coefficients.

1. INTRODUCTION

The concept of a unit tensor is not specific to the (angular momentum) group SU_2.[1-4] As a matter of fact, a unit tensor can be defined for any (compact) topological group J as a set of irreducible unit tensor operators acting on the (pre-Hilbert) representation space $\mathscr{E}$ of J (cf. Appendix A). The latter operators are frequently referred to as Wigner operators because their matrix elements between vectors of $\mathscr{E}$ transforming irreducibly under J are nothing but Wigner (or Clebsch–Gordan) coefficients of J (see also Ref. 1 and the introductory notes in Ref. 2). Therefore, they constitute basic ingredients for the Wigner–Racah algebra of J (see Refs. 1 and 3 for $J = U_n$). In the case where J is a compact Lie group, the unit tensor operators comprise the unit adjoint tensor operators which play a key role in the solution of the inner multiplicity problem for the Wigner coefficients of J (see Ref. 3 for $J = U_n$ and Ref. 4 for $J = O_n$).

We shall be mainly concerned in this paper with unit tensor operators for $J = SU_2$ in a $SU_2 \supset U_1$ basis and more generally in a $SU_2 \supset G^*$ basis, where G^* stands for (according to molecular physics notation) the spinor group of a molecular or crystallographic point symmetry group G.

We define the $SU_2 \supset U_1$ unit tensor operator $t_{kq\alpha}$ by its $j'm' - jm$ matrix elements through

$$\langle j'm' | t_{kq\alpha} | jm \rangle = \delta(j', j+\alpha)(-1)^{2k}(2j'+1)^{-1/2} \langle jkmq | j'm' \rangle . \quad (1)$$

It clearly appears that the action of $t_{kq\alpha}$ on the state vector $|jm\rangle$ leads to a state vector characterized by $j' = j + \alpha$ and $m' = m + q$.

Operators proportional (as far as matrix elements are concerned) to $t_{kq\alpha}$ enter various fields of physics and chemistry. The operator $t_{kq\alpha}$ resembles the operator $t(kq\alpha)$ introduced by Schwinger[5] in his famous treatment of angular momentum. In fact, we have

$$t(kq\alpha) = (-1)^{k+\alpha} \left(\frac{(2j+\alpha+k+1)!}{(2j+\alpha-k)!} \right)^{1/2} t_{kq\alpha} . \quad (2)$$

The case $\alpha = 0$ deserves special consideration because it corresponds to the Racah[6] unit tensor operator u^k_q. Indeed, our normalization for $t_{kq\alpha}$ is chosen in a way that

$$u^k_q = t_{kq0} , \quad (3)$$

as far as (shell model $l'm' - lm$) matrix elements are concerned. This choice justifies itself by the importance of the Racah unit tensor u^k for nuclear, atomic, molecular, and solid state spectroscopy. For instance, let us mention the growing interest of u^k for crystal- and ligand-field theory[7] and for the coupling between conduction electrons and moments of $3d$ and $4f$ ions in metals.[8] Another useful normalization is carried out by the Judd[9] tensor v^k:

$$v^k_q = (2k+1)^{1/2} t_{kq0} . \quad (4)$$

Furthermore, we have the following correspondences:

$$\begin{aligned} O^0_k &\sim t_{k00} , \\ O^q_k(c) &\sim \frac{1}{\sqrt{2}} [t_{k-q0} + (-1)^q t_{kq0}], \quad q \neq 0, \\ O^q_k(s) &\sim \frac{i}{\sqrt{2}} [t_{k-q0} - (-1)^q t_{kq0}], \quad q \neq 0, \\ \tilde{O}_{kq} (\equiv R_{kq}) &= \frac{1}{2^k} \left(\frac{(2j+k+1)!}{(2j-k)!} \right)^{1/2} t_{kq0} , \end{aligned} \quad (5)$$

[a)]Work supported in part by the France–Canada exchange programme, the Quantum Theory Group of the University of Waterloo, and the Institut National de Physique Nucléaire et de Physique des Particules (IN2P3). This work constitutes a part of Chapter 1 of the State Doctorate Thesis in preparation by G. Grenet.

[b)]Permanent address: Institut de Physique Nucléaire, 43 Bd du 11 Novembre, 69621 Villeurbanne, France.

 0022-2488/80/030422-18$1.00

$$T_{kq} = k!\left(\frac{(2j+k+1)!}{2^k(2k)!(2j-k)!}\right)^{1/2} t_{kq0},$$

where O, $\tilde{O}(\equiv R)$, and T denotes the operators defined by Stevens,[10] Buckmaster (Lindgård),[11] and Buckmaster *et al.*,[12] respectively. These latter operators are of central importance in the theory of magnetic resonance (electron paramagnetic resonance as well as electronic and nuclear double resonance) and related phenomena.[10-13] Note that when k is integer, all those of the preceding operators that are proportional to $t_{kq\alpha}$ transform under rotation like the spherical harmonic Y_{kq} while $O_k^q(c)$ and $O_k^q(s)$ transform like the so-called tesseral harmonics Z_{kq}^c and Z_{kq}^s, respectively.

There exist numerous realizations for the operators of type $t_{kq\alpha}$, which turn out to be useful in physical applications. We shall consider in turn the particular cases $\alpha = 0$ and $\alpha =$ integer, and the general case $\alpha =$ integer or half-integer.

The case $\alpha = 0$

First, t_{kq0} can be realized in terms of (two) Bose operators. This yields (generally infinite) expansions that are of interest in the theory of magnetism (especially for spin wave calculations) and in nuclear physics (especially for a description of collective motions in nuclei). The best known ways to obtain such Bose operator expansions are probably through the use of the Holstein–Primakoff transformation and the Dyson and Maleev transformation.[14] Along this line, let us also mention the recent MME method[15] from which a Bose expansion of any $SU_2 \supset U_1$ tensor operator can be obtained by matching pertinent matrix elements.

Second, t_{kq0} can be realized in the enveloping algebra of su_2. This yields realizations which are known as polarized harmonics in nuclear physics[16] and as (diagonal) operator equivalents in solid state physics.[10-13] (The word diagonal refers to $\alpha = 0$.) The polarized form of t_{kq0} can be obtained in principle from

$$t_{kq0} = \frac{2^k}{k!}\left(\frac{4\pi}{2k+1}\,\frac{(2j-k)!}{(2j+k+1)!}\right)^{1/2}(\mathbf{J}\cdot\mathbf{grad})^k\,\mathscr{Y}(\mathbf{r}), \quad (6)$$

i.e., by polarization of the solid harmonic $\mathscr{Y}_{kq}(\mathbf{r})$. However, the obtention of the operator equivalents form of t_{kq0} is easier in many respects. The operator equivalents have proved to be extremely fruitful for the understanding of the magnetic and optical properties of d^N and f^N partly-filled shell ions plunged into crystalline materials. Therefore, it is perhaps worthwhile to briefly discuss the state-of-the-art in the operator equivalents and the operator equivalents method originally introduced by Stevens[10] and used in molecular and solid state physics since more than 25 years.[10-13]

In many fields we need calculate matrix elements of interactions involving (or transforming like) harmonic polynomials $r^k Y_{kq}(\theta,\varphi)$ and $r^{-k-1} Y_{kq}(\theta,\varphi)$, or more generally quantities of type $f(r)Y_{kq}(\theta,\varphi)$. The Steven prescription to get the matrix elements of $f(r)Y_{kq}(\theta,\varphi)$ within a space $\epsilon(j)$ of given angular momentum j proceeds as follows: (i) Express $f(r)Y_{kq}(\theta,\varphi)$ in terms of x, y, z (or $x \pm iy, z$) and symmetrize the obtained expression. (ii) Make the substitution $u \to J_u$ with $u = x, y, z$ and transform the obtained expression by using the angular momentum commutation rules. This leads to a diagonal operator equivalent for $f(r)Y_{kq}(\theta,\varphi)$ the (easily obtainable) matrix elements of which are (by virtue of the Wigner–Eckart theorem for the chain $SU_2 \supset U_1$) proportional within $\epsilon(j)$ to those of $f(r)Y_{kq}(\theta,\varphi)$. (iii) Evaluate the proportionality constant by working out twice one single matrix element.

The case $\alpha =$ integer

A prescription similar to the Stevens one exists when α is an integer different from zero and permits one to get the matrix elements of $f(r)Y_{kq}(\theta,\varphi)$ connecting states of angular momenta j and $j+\alpha$.[17] This yields $f(r)Y_{kq}(\theta,\varphi)$ to be mimicked by an off-diagonal operator equivalent, a quantity defined in the enveloping algebra of a 10-dimensional Lie algebra, namely, the Schwinger algebra s (to be explicited below).

Although the two afore-mentioned prescriptions to derive realizations of $t_{kq\alpha}$ with k, and therefore α, integer in the enveloping algebra of either su_2 ($\alpha = 0$) or s ($\alpha \neq 0$) both lie on the $SU_2 \supset U_1$ Wigner–Eckart theorem, it is to be noted that they lead to calculations of the matrix elements of $f(r)Y_{kq}(\theta,\varphi)$ which are far more tedious than the ones involved with a direct application of this theorem. It is the opinion of the authors that the development and the use of the operator equivalents method partly prevented and obscured the penetration of the Wigner–Racah (angular momentum) calculus in EPR and ENDOR spectroscopies and, to a less extent, in crystal-field theory. However, as a nice counterpart, it is probably true that the operator equivalents method largely contributed to the development of the (static and dynamic) spin Hamiltonian formalism originally introduced by Abragam and Pryce (cf. Ref. 13) and which is so useful for phenomenological descriptions of microwave resonance data.

There are many other ways to obtain realizations of t_{kq0} and $t_{kq\alpha}$, with α being integer, in the enveloping algebras of su_2 and s, respectively.[12,16,18-22] Let us mention, among others, the algorithms based on: (i) Step-up procedures with shift (lowering or raising) operators.[12,16] (ii) Polynomial methods as, e.g., in the approaches starting from the Molien generating function[18] or from the construction of orthogonal polynomials of a discrete real variable.[19] (iii) Analyses of formulas for Clebsch–Gordan coefficients.[20,22] (iv) Finite-difference and commutator calculus.[19,21]

The case $\alpha =$ integer and half-integer

In the general case, a boson realization of $t_{kq\alpha}$ can be obtained in principle from the Schwinger generating function for $t(kq\alpha)$.[5] An alternative realization can be excerpted from the boson representation of the $SU_2 \supset U_1$ Wigner–Eckart theorem recently proposed by Yamamura *et al.*[23] in their investigation of the Schwinger representation of the quantized rotator and its application to nuclear structure theory. Finally, the present authors have recently reported a preliminary account of another boson realization of $t_{kq\alpha}$.[22]

It is one of the aims of this work to discuss realizations of $t_{kq\alpha}$ which cover both the case k (and therefore q and α) integer and half-an-odd-integer. We start in Sec. 2 with those

properties of the operators $t_{kq\alpha}$ (and more generally of the G^* symmetry adapted operators $t_{ka\Gamma\gamma\alpha}$) which do not depend on their realizations. We then turn our attention to the problem of constructing realizations of $t_{kq\alpha}$ in terms of boson operators. For that purpose, we devote Sec. 3 to (some aspects of) the boson representation of angular momentum introduced by Jordan and fully developed by Schwinger.[5] This representation is by now used in physical fields as distant as: the theory of magnetism,[24] the collective model and the shell model of the nucleus,[25] and the elementary particle physics.[26] A Lie-like approach to s, relevant for Schwinger's angular momentum calculus, is developed in Sec. 3. It is proved that the $O_{3,2}$ de Sitter algebra constitutes the most general frame for an investigation of Schwinger's calculus. The pseudo-orthogonal groups $O_{p,q}$ have received considerable attention in the last 15 years. In particular, the de Sitter group $O_{3,2}$ has been recently discussed in various contexts: rigid rotator,[27,28] hydrogen atom,[29] symmetric top,[30] and mathematical physics.[31,32] It turns out to be one of the maximal subgroups of the conformal group $SO_{4,2}$ (locally isomorphic to $SU_{2,2}$ and $Sp_{4;\mathbf{R}}$). In addition, contraction of the de Sitter group $O_{3,2}$ (or $O_{4,1}$) yields the Poincaré group. All the closed connected subgroups of $O_{3,2}$ have been recently exhibited by Patera *et al.*[31] and the reader is referred to their work for the material concerning $O_{3,2}$ relevant to the present paper. We show in Sec. 4 how, when applied to a composite system (the Schwinger algebra of which is described by the chain $SO_{3,2} \otimes SO_{3,2} \supset SO_{3,2}$), Schwinger's calculus allows one to systematically derive families of recursion relationships for the $SU_2 \supset U_1$ Clebsch–Gordan coefficients. In particular, recurrence relations between 3-jm symbols recently (re) derived[33] from considerations on hypergeometric functions or recoupling formulas are obtained through simple manipulations of ladder operators. Section 5 goes back to our first preoccupation: It gives some realizations of the unit tensor operator $t_{kq\alpha}$. Several general formulas for $t_{kq\alpha}$ valid in the cases k integer and half-an-odd-integer are obtained in the boson representation of angular momentum. For the special (nevertheless physically important) case where k is integer, the formulas can be rewritten in the enveloping algebra of $so_{3,2}$. Many examples are given throughout the paper and useful formulas related to: (1) the Lie algebra it is possible to associate with the Wigner–Racah algebra of any compact group and (2) some properties of the coupling coefficients for the chain $SU_2 \supset G^*$, are relegated in Appendices A and B, respectively.

2. GENERAL PROPERTIES OF THE SU_2 UNIT TENSOR

Notation: We start from the Hilbert space $\mathscr{E} = \oplus_j \epsilon(j)$, where

$$\epsilon(j) = \{ | jm\rangle : m \text{ ranging}\} \tag{7}$$

is an irreducible subspace of $\mathscr{E}$ associated with the IRC (irreducible representations class) j of SU_2 and spanned by the eigenvectors $|jm\rangle$ of the square J^2 and the 3-component J_3 (or J_z) of a generalized angular momentum operator. By introducing the subduction $SU_2 \downarrow G^*$, the space $\epsilon(j)$ decomposes as $\epsilon(j) = \oplus_{a\Gamma} \epsilon(ja\Gamma)$, where Γ stands for an IRC of G^*, Γ_0 being the identity IRC, and a denotes a branching multiplicity label to be used when Γ occurs several times in j. Indeed, we have

$$\epsilon(ja\Gamma) = \{ | ja\Gamma\gamma\rangle : \gamma \text{ ranging}\}, \tag{8}$$

where the $SU_2 \supset G^*$ symmetry adapted vector

$$| ja\Gamma\gamma\rangle = \sum_m |jm\rangle \langle jm | ja\Gamma\gamma\rangle \tag{9}$$

is obtained from the $SU_2 \supset U_1$ vectors $|jm\rangle$ with the help of a unitary transformation, the matrix elements of which are $\langle jm | ja\Gamma\gamma\rangle$. By applying the same transformation to the $SU_2 \supset U_1$ operators $t_{kq\alpha}$, we obtain $SU_2 \supset G^*$ symmetry adapted unit tensor operators

$$t_{ka\Gamma\gamma\alpha} = \sum_q t_{kq\alpha} \langle kq | ka\Gamma\gamma\rangle . \tag{10}$$

As a word of comment, it should be noted that the transformation from the $\{m\}$ scheme to the $\{a\Gamma\gamma\}$ scheme is chosen in such a way that, for fixed Γ, all sets $\{ | ja\Gamma\gamma\rangle : \gamma \text{ ranging}\}$ and $\{t_{ka\Gamma\gamma\alpha} : \gamma \text{ ranging}\}$ span the same (rather than equivalent) irreducible matrix representation. This standardization turns out to be a necessary requirement for the application of Racah's factorization lemma[34] to the chain $SU_2 \supset G^*$ (cf. Appendices A and B). When applied simultaneously to vectors and operators, the transformation under consideration allows us to transform Eq. (1) in the following $SU_2 \supset G^*$ symmetry adapted form

$$\langle j_1 a_1 \Gamma_1 \gamma_1 | t_{ka\Gamma\gamma\alpha} | j_2 a_2 \Gamma_2 \gamma_2 \rangle = \delta(j_1, j_2 + \alpha) f \begin{pmatrix} j_1 & j_2 & k \\ a_1\Gamma_1\gamma_1 & a_2\Gamma_2\gamma_2 & a\Gamma\gamma \end{pmatrix}, \tag{11}$$

where the f symbol is defined via (cf. Appendix B)[35,36]

$$f \begin{pmatrix} j_1 & j_2 & k \\ a_1\Gamma_1\gamma_1 & a_2\Gamma_2\gamma_2 & a\Gamma\gamma \end{pmatrix} = \sum_{m_1 q m_2} (-1)^{j_1 - m_1} \begin{pmatrix} j_1 & k & j_2 \\ -m_1 & q & m_2 \end{pmatrix} \times \langle j_1 m_1 | j_1 a_1 \Gamma_1 \gamma_1 \rangle^* \langle kq | ka\Gamma\gamma\rangle \times \langle j_2 m_2 | j_2 a_2 \Gamma_2 \gamma_2 \rangle . \tag{12}$$

To end up with our notational preliminaries, we define

$$\mathrm{tr}_{\epsilon(j)} A = \sum_m \langle jm | A | jm\rangle , \qquad \mathrm{tr}_{\epsilon(ja\Gamma)} A = \sum_\gamma \langle ja\Gamma\gamma | A | ja\Gamma\gamma\rangle . \tag{13}$$

Further, it will prove convenient to use the abbreviated form μ for $a\Gamma\gamma$. Note that to pass from G^* arbitrary to $G^* = U_1$, it will be sufficient to apply the following correspondence rules

$$f \begin{pmatrix} j_1 & j_2 & k \\ \mu_1 & \mu_2 & \mu \end{pmatrix} \leftrightarrow (-1)^{j_1 - m_1} \begin{pmatrix} j_1 & k & j_2 \\ -m_1 & q & m_2 \end{pmatrix} \text{ and } \mu \leftrightarrow m \text{ or } q. \tag{14}$$

Then, $\delta(\mu_2\ \mu_1)$ will be understood as $\delta(a_2 a_1)\delta(\Gamma_2\Gamma_1)\delta(\gamma_2\gamma_1)$ or as $\delta(m_2 m_1)$. Finally, all other symbols will have their usual significance.

Phase choices: It is straightforward to verify that

$$t_{k\mu\alpha}{}^{\dagger} = (-1)^{k+\alpha} \sum_{\mu'} \begin{pmatrix} k \\ \mu\mu' \end{pmatrix} t_{k\mu'-\alpha} \tag{15}$$

and

$$K t_{k\mu\alpha} K^{-1} = \sum_{\mu'} \begin{pmatrix} k \\ \mu\mu' \end{pmatrix} t_{k\mu'\alpha}\,, \tag{16}$$

so that

$$t_{k\mu\alpha}{}^{\dagger} = (-1)^{k+\alpha} K t_{k\mu-\alpha} K^{-1}, \tag{17}$$

where K is the Wigner time-reversal operator and the Herring–Wigner metric tensor[37] is defined through (cf. Appendix B)[36]

$$\begin{pmatrix} j \\ \mu\mu' \end{pmatrix} = \sum_{m} (-1)^{j+m} \langle jm \mid j\mu\rangle^* \langle j-m \mid j\mu'\rangle^*. \tag{18}$$

Note that when $G^* = \mathrm{U}_1$, Eqs. (15) and (16) reduce to

$$t_{kq\alpha}{}^{\dagger} = (-1)^{-q+\alpha} t_{k-q-\alpha} \tag{19}$$

and

$$K t_{kq\alpha} K^{-1} = (-1)^{k+q} t_{k-q\alpha}\,, \tag{20}$$

respectively, in agreement with usual phase choices (cf. Ref. 38).

Orthogonality and completeness relations: The relations

$$\sum_{\mu} t_{k\mu\alpha}{}^{\dagger} t_{k\mu\alpha'} = \delta(\alpha'\alpha)\delta(j+\alpha,k,j)(2j+1)^{-1} \tag{21}$$

and

$$\sum_{\alpha} [2(j-\alpha)+1]\, t_{k\mu\alpha} t_{k\mu'\alpha}{}^{\dagger} = \delta(\mu'\mu), \tag{22}$$

are valid when acting on $\epsilon(j)$. They follow from the orthogonality property of the f coefficients (cf. Appendix B). It is to be observed that Eqs. (21) and (22) parallel the orthogonality relation (2.14) and the completeness relation (2.15), respectively, given in Ref. 3 for the unit tensor operators of U_n.

$\epsilon(j)$-orthogonality property: The $t_{k\mu\alpha}$'s are $\epsilon(j)$-mutually orthogonal in the sense that

$$\mathrm{tr}_{\epsilon(j)} t_{k_1\mu_1\alpha_1}{}^{\dagger} t_{k_2\mu_2\alpha_2} = \delta(j+\alpha_1,j,k_1)\delta(k_2k_1) \times \delta(\mu_2\mu_1)\delta(\alpha_2\alpha_1)(2k_1+1)^{-1}. \tag{23}$$

The proof of Eq. (23) directly follows from the orthogonality property of the f coefficients. In the particular case $\alpha_1 = \alpha_2 = 0$, Eq. (23) shows that the $(2j+1)^2$ operators u^k_{μ}, with μ ranging and $k = 0(1)2j$, acting on $\epsilon(j)$ span a pre-Hilbert space τ with respect to the Hilbert–Schmidt scalar product $(A.B) = \mathrm{tr}_{\epsilon(j)} A^{\dagger} B$. The vectors $u^{k_1}_{\mu_1}$ and $u^{k_2}_{\mu_2}$ are thus mutually orthogonal on τ (see also Ref. 19). Equation (23) generalizes the trace relation known for the $\mathrm{SU}_2 \supset \mathrm{U}_1$ diagonal unit tensor operators (cf. for example Ref. 39). This latter relation is very popular in the density matrix theory and its application to various fields as, e.g., the optical pumping theory. It also specializes to the so-called barycenter rule which depicts the spectroscopic stability principle.

Fourier analysis: The coefficients $c(k\mu\alpha)$ of the development

$$H = \sum_{k\mu\alpha} c(k\mu\alpha) t_{k\mu\alpha} \tag{24}$$

of an arbitrary operator H acting on $\epsilon(j)$ are given by

$$c(k\mu\alpha)\delta(j+\alpha,j,k) = (2k+1)\,\mathrm{tr}_{\epsilon(j)} t_{k\mu\alpha}{}^{\dagger} H. \tag{25}$$

Equation (25) is an immediate consequence of Eq. (23). As a concrete example, the $\mathrm{SU}_2 \supset G^*$ Wigner projection-transfer operator (known when $G^* = \mathrm{U}_1$ as the Hill–Wheeler integral in nuclear physics[40])

$$P_{j\mu\mu'} = \frac{2j+1}{|\mathrm{SU}_2|} \int_{\mathrm{SU}_2} dR\, D^j(R)_{\mu\mu'}{}^* P_R \tag{26}$$

can be developed as

$$P_{j\mu\mu'} = \sum_{k\mu''} (2k+1) f \begin{pmatrix} j & j & k \\ \mu & \mu' & \mu'' \end{pmatrix} u^k_{\mu''}\,. \tag{27}$$

As another example, the development

$$Y_{\kappa\mu} = (-1)^l (2l+1)^{1/2} ((2k+1)/4\pi)^{1/2} \times \sum_{\alpha} (-1)^{\alpha} [2(l+\alpha)+1]^{1/2} \begin{pmatrix} l+\alpha & k & l \\ 0 & 0 & 0 \end{pmatrix} t_{k\mu\alpha} \tag{28}$$

holds as far as $l'\mu' - l\mu$ matrix elements are concerned.

$\epsilon(ja\Gamma)$-orthogonality property: The $t_{ka\Gamma\gamma}$'s are $\epsilon(ja\Gamma)$-mutually orthogonal in the sense that

$$\mathrm{tr}_{\epsilon(ja\Gamma)} t_{k_1a_1\Gamma_1\gamma_1\alpha_1}{}^{\dagger} t_{k_2a_2\Gamma_2\gamma_2\alpha_2} = \delta(\Gamma_2\Gamma_1)\delta(\gamma_2\gamma_1)\delta(\alpha_2\alpha_1) c(jak_1a_1k_2a_2\Gamma\Gamma_1\alpha_1), \tag{29}$$

where the coefficient c does not depend on γ_2 and γ_1. The proof of Eq. (29) follows by applying Racah's lemma to the chain $\mathrm{SU}_2 \supset G^*$ and by using the so-called orthogonality-completeness property for the Wigner coefficients of G^* (cf. Appendices A and B).

Commutator relations: The commutator ($[\ ,\ ] \equiv [\ ,\ ]_-$) and anticommutator ($[\ ,\ ]_+$) of two $\mathrm{SU}_2 \supset G^*$ unit tensor operators are given by

$$[t_{k_1\mu_1\alpha_1}, t_{k_2\mu_2\alpha_2}]_{\pm} = \sum_{k_3\mu_3\alpha_3} \delta(\alpha_3,\alpha_1+\alpha_2)(-1)^{2j+k_3+\alpha_3} \times (2k_3+1) f \begin{pmatrix} k_3 & k_2 & k_1 \\ \mu_3 & \mu_2 & \mu_1 \end{pmatrix} \times \left[(-1)^{k_1+k_2+k_3+2\alpha_2} \begin{Bmatrix} k_1 & k_2 & k_3 \\ j & j+\alpha_3 & j+\alpha_2 \end{Bmatrix} \pm (-1)^{2\alpha_1} \begin{Bmatrix} k_1 & k_2 & k_3 \\ j+\alpha_3 & j & j+\alpha_1 \end{Bmatrix} \right] t_{k_3\mu_3\alpha_3}\,. \tag{30}$$

The proof of Eq. (30) proceeds along the same lines as the one described by Judd,[9] on the basis of angular momentum recoupling techniques, in the case of $G^* = \mathrm{U}_1$ and $\alpha_1 = \alpha_2 = 0$. (See also Appendix A where the proof is outlined in the case of an arbitrary compact group.)

As a trivial by-product, Eq. (30) provides an expression of the product of the two operators $t_{k_1\mu_1\alpha_1}$ and $t_{k_2\mu_2\alpha_2}$ in terms of operators $t_{k_3\mu_3\alpha_3}$. Products of this type prove useful in various physical problems. In the case $\alpha_1 = \alpha_2 = 0$, they can be used in the calculation of time derivatives of tensor operators and in the evaluation of thermodynamic averages for systems described by a spin Hamiltonian (cf. Refs. 10–13 and more specifically the work by Lindgård in Ref. 11).

By way of illustration, let us consider the case $\Gamma_1 = \Gamma_2 = \Gamma_0$. Equation (30) then yields

$$\begin{aligned}
&[t_{k_1 a_1 \Gamma_0 \gamma_0 \alpha_1}, t_{k_2 a_2 \Gamma_0 \gamma_0 \alpha_2}]_{\pm} \\
&= \sum_{k_3 a_3 \alpha_3} \delta(\alpha_3, \alpha_1 + \alpha_2)(-1)^{2j + k_3 + \alpha_3} \\
&\times (2k_3 + 1) f \begin{pmatrix} k_3 & k_2 & k_1 \\ a_3 \Gamma_0 \gamma_0 & a_2 \Gamma_0 \gamma_0 & a_1 \Gamma_0 \gamma_0 \end{pmatrix} \\
&\times \left[(-1)^{k_1 + k_2 + k_3 + 2\alpha_2} \begin{Bmatrix} k_1 & k_2 & k_3 \\ j & j + \alpha_3 & j + \alpha_2 \end{Bmatrix} \right. \\
&\left. \pm (-1)^{2\alpha_1} \begin{Bmatrix} k_1 & k_2 & k_3 \\ j + \alpha_3 & j & j + \alpha_1 \end{Bmatrix} \right] t_{k_3 a_3 \Gamma_0 \gamma_0 \alpha_3},
\end{aligned} \tag{31}$$

which provides a development of the product of two G-invariant operators of type $t_{k a \Gamma_0 \gamma_0 \alpha}$ as a sum of G-invariant operators of the same type.

By specializing Eq. (30) to the case $\alpha_1 = \alpha_2 = 0$, it appears that, when acting on the space $\epsilon(j)$, the set $\{u_\mu^k : \mu \text{ and } k \text{ ranging}\}$ can be used to generate some subalgebra of $gl_{2j+1;\mathbb{R}}$ with respect to the Lie product $[\ ,\ ]$. This result, of considerable importance for physical applications, was discussed by Racah[34] and Judd[9] in the case $G^* = U_1$. [The passage from $G^* = U_1$ to G^* arbitrary is achieved owing to the correspondence rules (14) and corresponds to a simple change of the structure constants in the Lie algebra of the group $GL_{2j+1;\mathbb{R}}$.[36]] As an illustration, su_2 and su_3 are generated by the sets $\{u_\mu^1 : \mu \text{ ranging}\}$ and $\{u_\mu^1, u_\mu^2 : \mu \text{ and } \mu' \text{ ranging}\}$ acting on $\epsilon(\frac{1}{2})$ and $\epsilon(1)$, respectively. More precisely, su_3 can be described in an $O_3 \supset G$ basis by

$$\begin{aligned}
[u_\mu^1, u_{\mu'}^1] &= \sum_{\mu''} f \begin{pmatrix} 1 & 1 & 1 \\ \mu'' & \mu' & \mu \end{pmatrix} u_{\mu''}^1, \\
[u_\mu^1, u_{\mu'}^2] &= \sqrt{5} \sum_{\mu''} f \begin{pmatrix} 2 & 2 & 1 \\ \mu'' & \mu' & \mu \end{pmatrix} u_{\mu''}^2, \\
[u_\mu^2, u_{\mu'}^2] &= \frac{-3}{\sqrt{5}} \sum_{\mu''} f \begin{pmatrix} 1 & 2 & 2 \\ \mu'' & \mu' & \mu \end{pmatrix} u_{\mu''}^1.
\end{aligned} \tag{32}$$

Similar expressions could be obtained for su_5 in an $O_3 \supset O$ basis, where O is the octahedral group, a basis of interest in those molecular problems where the chain $SU_5 \supset O_5 \supset O_3 \supset O$ is relevant as, e.g., in the Jahn–Teller problem.[41]

3. THE SCHWINGER ALGEBRA

Let $a_\pm$ and $a_\pm^+$ be the operators defined on $\mathscr{E}$ by

$$a_\pm |jm\rangle = (j \pm m)^{1/2} |j - \tfrac{1}{2}, m \mp \tfrac{1}{2}\rangle \tag{33}$$

and

$$a_\pm^+ |jm\rangle = (j \pm m + 1)^{1/2} |j + \tfrac{1}{2}, m \pm \tfrac{1}{2}\rangle, \tag{34}$$

respectively. We easily deduce from Eqs. (33) and (34) the commutation relations

$$\begin{aligned}
&[a_+, a_+^+] = [a_-, a_-^+] = 1, \\
&[a_+, a_-] = [a_+, a_-^+] = [a_+^+, a_-] = [a_+^+, a_-^+] = 0.
\end{aligned} \tag{35}$$

The sets $\{a_+, a_+^+\}$ and $\{a_-, a_-^+\}$ are closed under Hermitian conjugation:

$$a_\pm^+ = (a_\pm)^\dagger. \tag{36}$$

They can be considered as two sets of boson operators corresponding to a two-dimensional isotropic harmonic oscillator (cf. Ref. 5). These sets are interchanged under time-reversal operation:

$$K a_\pm K^{-1} = \mp a_\mp, \quad K a_\pm^+ K^{-1} = \mp a_\mp^+. \tag{37}$$

By induction we get the useful formulas

$$\begin{aligned}
&\langle j'm' | (a_\pm)^p | jm \rangle \\
&\quad = \delta(j', j - p/2) \delta(m', m \mp p/2) \left(\frac{(j \pm m)!}{(j \pm m - p)!} \right)^{1/2}, \\
&\langle j'm' | (a_\pm^+)^p | jm \rangle \\
&\quad = \delta(j', j + p/2) \delta(m', m \pm p/2) \left(\frac{(j \pm m + p)!}{(j \pm m)!} \right)^{1/2}.
\end{aligned} \tag{38}$$

From the four boson creation and annihilation operators $a_\pm$ and $a_\pm^+$, we can form ten linearly independent bilinear expressions. Let us put

$$\begin{aligned}
&k_0^0 = \tfrac{1}{2}(a_+^+ a_+ - a_-^+ a_-), \quad k_\pm^0 = a_\mp^+ a_\pm^+, \\
&k_0^+ = a_+^+ a_-^+, \quad k_\pm^+ = \mp a_\pm^+ a_\pm^+, \\
&k_0^- = a_+ a_-, \quad k_\pm^- = \pm a_\mp a_\mp, \\
&J = a_+^+ a_+ + a_-^+ a_- + 1.
\end{aligned} \tag{39}$$

The notation k_σ^ρ follows the one by Atkins and Seymour[17] except for the operator J which is connected to their $\bar{J}$ by $J = 2\bar{J} + 1$. (J is the simplest operator to be introduced in order to close the set $\{k_\sigma^\rho : \rho \text{ and } \sigma \text{ ranging}\}$ under commutation.) Note that $-k_+^+$, k_-^+, k_+^-, and $-k_-^-$ can be identified with the operators $\hat{M}_+$, $\hat{N}_+$, $\hat{N}_-$, and $\hat{M}_-$, respectively, defined by Witschel.[42]

The operators $J_3 \equiv k_0^0$ and $J_\pm \equiv k_\pm^0$ satisfy the commutation relations

$$[J_3, J_\pm] = \pm J_\pm, \quad [J_+, J_-] = 2J_3 \tag{40}$$

of the *spherical* angular momentum while $K_3 \equiv \frac{1}{2}J$ and $K_\pm \equiv k_0^\pm$ satisfy the commutation relations

$$[K_3, K_\pm] = \pm K_\pm, \quad [K_+, K_-] = -2K_3 \tag{41}$$

of the *hyperbolic* angular momentum introduced by Schwinger.[5] The spherical and hyperbolic angular momentum operators can thus be regarded as the generators of the group SU_2 and $SU_{1,1}$ (two-to-one homomorphic to the group SO_3 and $SO_{2,1}$), respectively.

The properties of the k's and J can be readily deduced from those of the a's. By repeated application of Eqs. (33) and (34), we get (cf. Ref. 17)

$$\begin{aligned}
&k_\pm^0 |jm\rangle = [(j \mp m)(j \pm m + 1)]^{1/2} |j, m \pm 1\rangle, \\
&k_0^0 |jm\rangle = m |jm\rangle, \\
&k_0^+ |jm\rangle = [(j - m + 1)(j + m + 1)]^{1/2} |j + 1, m\rangle, \\
&k_0^- |jm\rangle = [(j - m)(j + m)]^{1/2} |j - 1, m\rangle, \\
&J |jm\rangle = (2j + 1) |jm\rangle, \\
&k_\pm^+ |jm\rangle = \mp [(j \pm m + 1)(j \pm m + 2)]^{1/2} |j + 1, m \pm 1\rangle, \\
&k_\pm^- |jm\rangle = \pm [(j \mp m - 1)(j \mp m)]^{1/2} |j - 1, m \pm 1\rangle.
\end{aligned} \tag{42}$$

TABLE I. Lie brackets of the Schwinger algebra s. For instance $[k_0^-, k_0^+] = J$.

	k_0^0	k_+^0	k_-^0	k_0^+	k_+^+	k_-^+	k_0^-	k_+^-	k_-^-	J
k_0^0	O	k_+^0	$-k_-^0$	O	k_+^+	$-k_-^+$	O	k_+^-	$-k_-^-$	O
k_+^0	$-k_+^0$	O	$2k_0^0$	$-k_+^+$	O	$2k_0^+$	$-k_+^-$	O	$2k_0^-$	O
k_-^0	k_-^0	$-2k_0^0$	O	k_-^+	$-2k_0^+$	O	k_-^-	$-2k_0^-$	O	O
k_0^+	O	k_+^+	$-k_-^+$	O	O	O	$-J$	$-2k_+^0$	$2k_-^0$	$-2k_0^+$
k_+^+	$-k_+^+$	O	$2k_0^+$	O	O	O	$2k_+^0$	O	$-2(J+2k_0^0)$	$-2k_+^+$
k_-^+	k_-^+	$-2k_0^+$	O	O	O	O	$-2k_-^0$	$-2(J-2k_0^0)$	O	$-2k_-^+$
k_0^-	O	k_+^-	$-k_-^-$	J	$-2k_+^0$	$2k_-^0$	O	O	O	$2k_0^-$
k_+^-	$-k_+^-$	O	$2k_0^-$	$2k_+^0$	O	$2(J-2k_0^0)$	O	O	O	$2k_+^-$
k_-^-	k_-^-	$-2k_0^-$	O	$-2k_-^0$	$2(J+2k_0^0)$	O	O	O	O	$2k_-^-$
J	O	O	O	$2k_0^+$	$2k_+^+$	$2k_-^+$	$-2k_0^-$	$-2k_+^-$	$-2k_-^-$	O

TABLE II. Lie brackets of the de Sitter algebra $so_{3,2}$. For instance $[A', B'] = -C'$.

	A'	B'	C'	D'	E'	F'	G'	H'	J'	K'
A'	O	$-C'$	B'	O	$-G'$	$-H'$	E'	F'	O	O
B'	C'	O	$-A'$	O	$-J'$	$-K'$	O	O	E'	F'
C'	$-B'$	A'	O	O	O	O	$-J'$	$-K'$	G'	H'
D'	O	O	O	O	$-F'$	E'	$-H'$	G'	$-K'$	J'
E'	G'	J'	O	F'	O	D'	A'	O	B'	O
F'	H'	K'	O	$-E'$	$-D'$	O	O	A'	O	B'
G'	$-E'$	O	J'	H'	$-A'$	O	O	D'	C'	O
H'	$-F'$	O	K'	$-G'$	O	$-A'$	$-D'$	O	O	C'
J'	O	$-E'$	$-G'$	K'	$-B'$	O	$-C'$	O	O	D'
K'	O	$-F'$	$-H'$	$-J'$	O	$-B'$	O	$-C'$	$-D'$	O

(Observe there is a misprint in the expression for $k^{0}_{+}\,|jm\rangle$ given in Ref. 17.) The set $\{J, k^{\rho}_{\sigma} : \rho \text{ and } \sigma \text{ ranging}\}$ is closed with respect to taking Hermitian conjugate since

$$k^{-\rho}_{\ \sigma} = (k^{\rho}_{\sigma})^{\dagger}, \quad J = J^{\dagger}. \tag{43}$$

In addition, time-reversal operation leads to

$$Kk^{\rho}_{\sigma}K^{-1} = -k^{\rho}_{\ -\sigma}, \quad KJK^{-1} = J. \tag{44}$$

Finally, by finite induction we obtain the useful formulas

$$\begin{aligned}
&\langle j'm' | (k^{0}_{\pm})^{p} | jm\rangle \\
&\quad = \delta(j',j)\delta(m',m \pm p)\left(\frac{(j \mp m)!(j \pm m + p)!}{(j \pm m)!(j \mp m - p)!}\right)^{1/2}, \\
&\langle j'm' | (k^{+}_{0})^{p} | jm\rangle \\
&\quad = \delta(j',j + p)\delta(m',m)\left(\frac{(j - m + p)!(j + m + p)!}{(j - m)!(j + m)!}\right)^{1/2}, \\
&\langle j'm' | (k^{-}_{0})^{p} | jm\rangle \\
&\quad = \delta(j',j - p)\delta(m',m)\left(\frac{(j - m)!(j + m)!}{(j - m - p)!(j + m - p)!}\right)^{1/2}, \\
&\langle j'm' | (k^{+}_{\pm})^{p} | jm\rangle \\
&\quad = \delta(j',j + p)\delta(m',m \pm p)(\mp 1)^{p}\left(\frac{(j \pm m + 2p)!}{(j \pm m)!}\right)^{1/2}, \\
&\langle j'm' | (k^{-}_{\pm})^{p} | jm\rangle \\
&\quad = \delta(j',j - p)\delta(m',m \pm p)\left(\frac{(j \mp m)!}{(j \mp m - 2p)!}\right)^{1/2}.
\end{aligned} \tag{45}$$

The set $\{J, k^{\rho}_{\sigma} : \rho \text{ and } \sigma \text{ ranging}\}$ is closed under commutation and thus describes a Lie algebra we shall note s and refer to as the Schwinger algebra. The relevant Lie brackets are listed in Table I. It is clear from Table I that $(k^{+}_{0}, k^{+}_{+}, k^{+}_{-})$ and $(k^{-}_{0}, k^{-}_{+}, k^{-}_{-})$ behave as vectors under the generators J_3 and $J_{\pm}$ of SU_2, a fact which is at the root of the prescriptions used in Ref. 17 for obtaining diagonal and off-diagonal operator equivalents as polynomials in $\bar{J}$ and k^{ρ}_{σ}. Note that the second order invariant of s,

$$\begin{aligned}
I_2 = {} & \tfrac{1}{6}J^2 - \tfrac{1}{6}K^2 - \tfrac{1}{24}(k^{+}_{+}k^{-}_{-} + k^{-}_{-}k^{+}_{+} \\
& + k^{+}_{-}k^{-}_{+} + k^{-}_{+}k^{+}_{-}),
\end{aligned} \tag{46}$$

assumes a simple form in terms of the Casimir operators J^2 and K^2 of su_2 and $su_{1,1}$, respectively.

It is not difficult to convert s to a semisimple noncompact Lie algebra. Let us introduce the anti-Hermitian (on the space $\mathscr{E}$) operators

$$\begin{aligned}
A' &= (i/2)(k^{0}_{+} + k^{0}_{-}), \\
B' &= \tfrac{1}{2}(k^{0}_{+} - k^{0}_{-}), \\
C' &= ik^{0}_{0}, \\
D' &= (i/2)J, \\
E' &= (i/2)(k^{+}_{0} + k^{-}_{0}), \\
F' &= \tfrac{1}{2}(k^{+}_{0} - k^{-}_{0}), \\
G' &= \tfrac{1}{4}(-k^{+}_{+} + k^{+}_{-} - k^{-}_{+} + k^{-}_{-}), \\
H' &= (i/4)(k^{+}_{+} - k^{+}_{-} - k^{-}_{+} + k^{-}_{-}), \\
J' &= (i/4)(k^{+}_{+} + k^{+}_{-} + k^{-}_{+} + k^{-}_{-}), \\
K' &= \tfrac{1}{4}(k^{+}_{+} + k^{+}_{-} - k^{-}_{+} - k^{-}_{-}).
\end{aligned} \tag{47}$$

It is then a straightforward but tedious piece of work to obtain the commutation relations reported in Table II. Comparison between Table II and the catalog set up by Patera *et al.*[31] for the de Sitter algebras of low dimensions shows that the algebra spanned by the set $\{A', B', \ldots, K'\}$ is isomorphic to $so_{3,2}$. The Schwinger algebra s thus turns out to be connected with the Lie algebra of the (noncompact and semisimple) de Sitter group $SO_{3,2}$.

A two-complex variables realization of s and $\{A', B', \ldots, K'\}$ in the Bargmann[43] space $\mathscr{F}_2$ can be obtained with the help of

$$\begin{aligned}
&k^{+}_{+} = -\xi^2, \quad k^{+}_{0} = \xi\eta, \quad k^{+}_{-} = \eta^2, \\
&k^{0}_{+} = \xi\frac{\partial}{\partial\eta}, \quad k^{0}_{0} = \frac{1}{2}\left(\xi\frac{\partial}{\partial\xi} - \eta\frac{\partial}{\partial\eta}\right), \\
&k^{0}_{-} = \eta\frac{\partial}{\partial\xi}, \\
&k^{-}_{+} = \frac{\partial^2}{\partial\eta^2}, \quad k^{-}_{0} = \frac{\partial^2}{\partial\xi\partial\eta}, \quad k^{-}_{-} = -\frac{\partial^2}{\partial\xi^2}, \\
&J = \xi\frac{\partial}{\partial\xi} + \eta\frac{\partial}{\partial\eta} + 1.
\end{aligned} \tag{48}$$

Note that I_2 can be written in the preceding realization as

$$\begin{aligned}
I_2 = {} & -\frac{1}{12}\left[\left(\xi\frac{\partial}{\partial\xi} + \eta\frac{\partial}{\partial\eta}\right)^2\right. \\
& \left. + 2\xi\frac{\partial}{\partial\xi} + 2\eta\frac{\partial}{\partial\eta} + \frac{7}{2}\right].
\end{aligned} \tag{49}$$

By combining Eqs. (47) and (48), the algebra $\{A', B', \ldots, K'\}$ can be realized in $\mathscr{F}_2$ by

$$\begin{aligned}
A' &= \frac{i}{2}\left(\xi\frac{\partial}{\partial\eta} + \eta\frac{\partial}{\partial\xi}\right), \\
B' &= \frac{1}{2}\left(\xi\frac{\partial}{\partial\eta} - \eta\frac{\partial}{\partial\xi}\right), \\
C' &= \frac{i}{2}\left(\xi\frac{\partial}{\partial\xi} - \eta\frac{\partial}{\partial\eta}\right), \\
D' &= \frac{i}{2}\left(\xi\frac{\partial}{\partial\xi} + \eta\frac{\partial}{\partial\eta} + 1\right), \\
E' &= \frac{i}{2}\left(\xi\eta + \frac{\partial^2}{\partial\xi\,\partial\eta}\right), \\
F' &= \frac{1}{2}\left(\xi\eta - \frac{\partial^2}{\partial\xi\,\partial\eta}\right), \\
G' &= \frac{1}{4}\left(\xi^2 + \eta^2 - \frac{\partial^2}{\partial\xi^2} - \frac{\partial^2}{\partial\eta^2}\right), \\
H' &= \frac{i}{4}\left(-\xi^2 - \eta^2 - \frac{\partial^2}{\partial\xi^2} - \frac{\partial^2}{\partial\eta^2}\right), \\
J' &= \frac{i}{4}\left(-\xi^2 + \eta^2 - \frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2}\right), \\
K' &= \frac{1}{4}\left(-\xi^2 + \eta^2 + \frac{\partial^2}{\partial\xi^2} - \frac{\partial^2}{\partial\eta^2}\right).
\end{aligned} \tag{50}$$

Some other (physical) realizations of both $\{A', B', \ldots, K'\}$ and s could be obtained in terms of space and time coordinates by noticing the relevance of $SO_{3,2}$ as a spectrum generating group of the rigid rotator[27,28] and the symmetric top.[30]

4. BOSON-LIKE APPROACH TO RECURSION RELATIONS

We now go to the coupling of two angular momenta, the boson representations of which are described by the sets $\{a_\pm, a_\pm^+\}$ and $\{b_\pm, b_\pm^+\}$. The coupled-boson operators $\mathscr{J}_3, \mathscr{J}_\pm, \mathscr{K}_3$, and $\mathscr{K}_\pm$ introduced by Schwinger[5] can be rewritten as

$$\begin{aligned}
\mathscr{J}_3 &= \tfrac{1}{2}(a_+^+ a_+ + a_-^+ a_- - b_+^+ b_+ - b_-^+ b_-),\\
\mathscr{J}_+ &= a_+^+ b_+ + a_-^+ b_-,\\
\mathscr{J}_- &= b_+^+ a_+ + b_-^+ a_-,\\
\mathscr{K}_3 &= \tfrac{1}{2}(a_+^+ a_+ + a_-^+ a_- + b_+^+ b_+ + b_-^+ b_-) + 1,\\
\mathscr{K}_+ &= a_+^+ b_-^+ - a_-^+ b_+^+,\\
\mathscr{K}_- &= a_+ b_- - a_- b_+.
\end{aligned} \tag{51}$$

The operators $\mathscr{J}_3$ and $\mathscr{J}_\pm$ satisfy the SO_3 commutation relations (40) while $\mathscr{K}_3$ and $\mathscr{K}_\pm$ satisfy the $SO_{2,1}$ commutation relations (41) and, further, the set $\{\mathscr{J}_3, \mathscr{J}_\pm\}$ commutes with $\{\mathscr{K}_3, \mathscr{K}_\pm\}$.[5] Therefore, $\mathscr{J}_3, \mathscr{J}_\pm, \mathscr{K}_3$, and $\mathscr{K}_\pm$ can be regarded as the generators of a $SO_{3,1}$ Lie group. As a consequence of the SO_3 and $SO_{2,1}$ commutation relations, we can derive the following ladder relationships

$$\begin{aligned}
\mathscr{J}_3 |j_1 j_2 jm\rangle &= (j_1 - j_2)|j_1 j_2 jm\rangle,\\
\mathscr{J}_+ |j_1 j_2 jm\rangle &= [(j - j_1 + j_2)(j + j_1 - j_2 + 1)]^{1/2} |j_1 + \tfrac{1}{2}, j_2 - \tfrac{1}{2}, j, m\rangle,\\
\mathscr{J}_- |j_1 j_2 jm\rangle &= [(j + j_1 - j_2)(j - j_1 + j_2 + 1)]^{1/2} |j_1 - \tfrac{1}{2}, j_2 + \tfrac{1}{2}, j, m\rangle,\\
\mathscr{K}_3 |j_1 j_2 jm\rangle &= (j_1 + j_2 + 1)|j_1 j_2 jm\rangle,\\
\mathscr{K}_+ |j_1 j_2 jm\rangle &= [(j_1 + j_2 - j + 1)(j_1 + j_2 + j + 2)]^{1/2} |j_1 + \tfrac{1}{2}, j_2 + \tfrac{1}{2}, j, m\rangle,\\
\mathscr{K}_- |j_1 j_2 jm\rangle &= [(j_1 + j_2 + j + 1)(j_1 + j_2 - j)]^{1/2} |j_1 - \tfrac{1}{2}, j_2 - \tfrac{1}{2}, j, m\rangle.
\end{aligned} \tag{52}$$

By applying the Schwinger operators $\mathscr{J}_\pm$ on $|j_1 \mp \tfrac{1}{2}, j_2 \pm \tfrac{1}{2}, j, m\rangle$, we directly obtain by means of simple ladder techniques the three-term recursion relations

$$\begin{aligned}
&[(j - j_1 + j_2 + 1)(j + j_1 - j_2)]^{1/2} \langle j_1 j_2 m_1 m_2 | jm\rangle\\
&\quad = [(j_1 + m_1)(j_2 + m_2 + 1)]^{1/2} \langle j_1 - \tfrac{1}{2}, j_2 + \tfrac{1}{2}, m_1 - \tfrac{1}{2}, m_2 + \tfrac{1}{2} | jm\rangle\\
&\qquad + [(j_1 - m_1)(j_2 - m_2 + 1)]^{1/2} \langle j_1 - \tfrac{1}{2}, j_2 + \tfrac{1}{2}, m_1 + \tfrac{1}{2}, m_2 - \tfrac{1}{2} | jm\rangle,
\end{aligned} \tag{53}$$

$$\begin{aligned}
&[(j + j_1 - j_2 + 1)(j - j_1 + j_2)]^{1/2} \langle j_1 j_2 m_1 m_2 | jm\rangle\\
&\quad = [(j_1 + m_1 + 1)(j_2 + m_2)]^{1/2} \langle j_1 + \tfrac{1}{2}, j_2 - \tfrac{1}{2}, m_1 + \tfrac{1}{2}, m_2 - \tfrac{1}{2} | jm\rangle\\
&\qquad + [(j_1 - m_1 + 1)(j_2 - m_2)]^{1/2} \langle j_1 + \tfrac{1}{2}, j_2 - \tfrac{1}{2}, m_1 - \tfrac{1}{2}, m_2 + \tfrac{1}{2} | jm\rangle.
\end{aligned}$$

[What we have just done parallels a well-known procedure to get (trivial) recursion relations by taking convenient matrix elements of the composite shift operator $J_\pm(a,b) = J_\pm(a) + J_\pm(b) \equiv k^0_\pm + l^0_\pm$, where the k's and l's are given by Eqs. (39) in terms of the a's and b's, respectively.] In a similar way, the action of $\mathscr{K}_\pm$ on $|j_1 \mp \tfrac{1}{2}, j_2 \mp \tfrac{1}{2}, j, m\rangle$ produces

$$\begin{aligned}
&[(j_1 + j_2 - j)(j_1 + j_2 + j + 1)]^{1/2} \langle j_1 j_2 m_1 m_2 | jm\rangle\\
&\quad = [(j_1 + m_1)(j_2 - m_2)]^{1/2} \langle j_1 - \tfrac{1}{2}, j_2 - \tfrac{1}{2}, m_1 - \tfrac{1}{2}, m_2 + \tfrac{1}{2} | jm\rangle\\
&\qquad - [(j_1 - m_1)(j_2 + m_2)]^{1/2} \langle j_1 - \tfrac{1}{2}, j_2 - \tfrac{1}{2}, m_1 + \tfrac{1}{2}, m_2 - \tfrac{1}{2} | jm\rangle,
\end{aligned} \tag{54}$$

$$\begin{aligned}
&[(j_1 + j_2 + j + 2)(j_1 + j_2 - j + 1)]^{1/2} \langle j_1 j_2 m_1 m_2 | jm\rangle\\
&\quad = [(j_1 + m_1 + 1)(j_2 - m_2 + 1)]^{1/2} \langle j_1 + \tfrac{1}{2}, j_2 + \tfrac{1}{2}, m_1 + \tfrac{1}{2}, m_2 - \tfrac{1}{2} | jm\rangle\\
&\qquad - [(j_1 - m_1 + 1)(j_2 + m_2 + 1)]^{1/2} \langle j_1 + \tfrac{1}{2}, j_2 + \tfrac{1}{2}, m_1 - \tfrac{1}{2}, m_2 + \tfrac{1}{2} | jm\rangle.
\end{aligned}$$

Equations (53) and (54) have been derived in term of the $3-jm$ symbol by Schulten and Gordon on the basis of recoupling relations and by Barut and Wilson with the aid of Rainville's contiguous relations for generalized hypergeometric functions.[33]

At this point, we foresee a general procedure to obtain recursion relations for $SU_2 \supset U_1$ Wigner coefficients (and possibly SU_2 Racah coefficients). The procedure may be described as follows: (i) construct polynomials in the $\mathscr{J}$, $\mathscr{K}$, k, and l's, (ii) express them in the a and b's (and/or if possible k and l's), and (iii) take convenient matrix elements of the so-obtained expressions.

To illustrate the procedure, let us consider the $\mathscr{K}\mathscr{K}$, $\mathscr{K}\mathscr{J}$, and $\mathscr{J}\mathscr{J}$ bilinear products

$$\mathscr{K}_-\mathscr{K}_- = -k_+^- l_-^- - 2k_0^- l_0^- - k_-^- l_+^-,$$

$$
\begin{aligned}
&\mathcal{J}_-\mathcal{J}_- = k_+^- l_-^+ + 2k_0^- l_0^+ + k_-^- l_+^+ ,\\
&\mathcal{J}_-\mathcal{K}_- = -k_+^- l_-^0 - 2k_0^- l_0^0 - k_-^- l_+^0 ,\\
&\mathcal{J}_+\mathcal{J}_+ = k_+^+ l_-^- + 2k_0^+ l_0^- + k_-^+ l_+^- ,\\
&\mathcal{K}_+\mathcal{K}_+ = -k_+^+ l_-^+ - 2k_0^+ l_0^+ - k_-^+ l_+^+ ,\\
&\mathcal{J}_+\mathcal{K}_+ = -k_+^+ l_-^0 - 2k_0^+ l_0^0 - k_-^+ l_+^0 ,\\
&\mathcal{J}_+\mathcal{K}_- = k_+^0 l_-^- + 2k_0^0 l_0^- + k_-^0 l_+^- ,\\
&\mathcal{J}_-\mathcal{K}_+ = k_+^0 l_-^+ + 2k_0^0 l_0^+ + k_-^0 l_+^+ ,\\
&\tfrac{1}{4}(\mathcal{K}_+\mathcal{K}_- + \mathcal{K}_-\mathcal{K}_+ - \mathcal{J}_+\mathcal{J}_- - \mathcal{J}_-\mathcal{J}_+ - 2) = -k_+^0 l_-^0 - 2k_0^0 l_0^0 - k_-^0 l_+^0 .
\end{aligned}
\tag{55}
$$

[The nine preceding operators are clearly defined in the enveloping algebra of $s(k) \otimes s(l)$ and are indeed proportional to scalar products of the six SU_2 vectors $V^\rho = (V_0^\rho, V_+^\rho, V_-^\rho)$ with $V = k, l$ and $\rho = 0, +, -$; see also Eq. (83) below.] It can be verified that the action of $\mathcal{J}_+\mathcal{K}_+$ on $|j_1+1, j_2, j, m\rangle$ yields the four-term recursion relation

$$
\begin{aligned}
&[(j_1+j_2-j)(j_1+j_2+j+1)(j-j_1+j_2+1)(j+j_1-j_2)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = [(j_1+m_1)(j_1+m_1-1)(j_2-m_2)(j_2+m_2+1)]^{1/2}\langle j_1-1, j_2, m_1-1, m_2+1 \mid jm\rangle\\
&\qquad -2m_2[(j_1-m_1)(j_1+m_1)]^{1/2}\langle j_1-1, j_2, m_1, m_2 \mid jm\rangle\\
&\qquad -[(j_1-m_1)(j_1-m_1-1)(j_2+m_2)(j_2-m_2+1)]^{1/2}\langle j_1-1, j_2, m_1+1, m_2-1 \mid jm\rangle .
\end{aligned}
\tag{56}
$$

Similarly, the eight remaining operators lead to

$$
\begin{aligned}
&[(j_1+j_2+j+3)(j_1+j_2-j+2)(j_1+j_2+j+2)(j_1+j_2-j+3)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = [(j_1-m_1+1)(j_1-m_1+2)(j_2+m_2+1)(j_2+m_2+2)]^{1/2}\langle j_1+1, j_2+1, m_1-1, m_2+1 \mid jm\rangle\\
&\qquad -2[(j_1-m_1+1)(j_1+m_1+1)(j_2-m_2+1)(j_2+m_2+1)]^{1/2}\langle j_1+1, j_2+1, m_1, m_2 \mid jm\rangle\\
&\qquad +[(j_1+m_1+1)(j_1+m_1+2)(j_2-m_2+1)(j_2-m_2+2)]^{1/2}\langle j_1+1, j_2+1, m_1+1, m_2-1 \mid jm\rangle ,\\
&[(j+j_1-j_2+2)(j-j_1+j_2-1)(j+j_1-j_2+1)(j-j_1+j_2)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = [(j_1-m_1+1)(j_1-m_1+2)(j_2-m_2)(j_2-m_2-1)]^{1/2}\langle j_1+1, j_2-1, m_1-1, m_2+1 \mid jm\rangle\\
&\qquad +2[(j_1-m_1+1)(j_1+m_1+1)(j_2-m_2)(j_2+m_2)]^{1/2}\langle j_1+1, j_2-1, m_1, m_2 \mid jm\rangle\\
&\qquad +[(j_1+m_1+1)(j_1+m_1+2)(j_2+m_2)(j_2+m_2-1)]^{1/2}\langle j_1+1, j_2-1, m_1+1, m_2-1 \mid jm\rangle ,\\
&[(j_1+j_2+j+2)(j_1+j_2-j+1)(j+j_1-j_2+1)(j-j_1+j_2)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = -[(j_1-m_1+1)(j_1-m_1+2)(j_2-m_2)(j_2+m_2+1)]^{1/2}\langle j_1+1, j_2, m_1-1, m_2+1 \mid jm\rangle\\
&\qquad -2m_2[(j_1-m_1+1)(j_1+m_1+1)]^{1/2}\langle j_1+1, j_2, m_1, m_2 \mid jm\rangle\\
&\qquad +[(j_1+m_1+1)(j_1+m_1+2)(j_2+m_2)(j_2-m_2+1)]^{1/2}\langle j_1+1, j_2, m_1+1, m_2-1 \mid jm\rangle ,\\
&[(j-j_1+j_2+2)(j+j_1-j_2-1)(j-j_1+j_2+1)(j+j_1-j_2)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = [(j_1+m_1)(j_1+m_1-1)(j_2+m_2+1)(j_2+m_2+2)]^{1/2}\langle j_1-1, j_2+1, m_1-1, m_2+1 \mid jm\rangle\\
&\qquad +2[(j_1-m_1)(j_1+m_1)(j_2-m_2+1)(j_2+m_2+1)]^{1/2}\langle j_1-1, j_2+1, m_1, m_2 \mid jm\rangle\\
&\qquad +[(j_1-m_1)(j_1-m_1-1)(j_2-m_2+1)(j_2+m_2+1)]^{1/2}\langle j_1-1, j_2+1, m_1+1, m_2-1 \mid jm\rangle ,
\end{aligned}
\tag{57}
$$

$$
\begin{aligned}
&[(j_1+j_2-j-1)(j_1+j_2+j)(j_1+j_2-j)(j_1+j_2+j+1)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = [(j_1+m_1)(j_1+m_1-1)(j_2-m_2)(j_2-m_2-1)]^{1/2}\langle j_1-1, j_2-1, m_1-1, m_2+1 \mid jm\rangle\\
&\qquad -2[(j_1-m_1)(j_1+m_1)(j_2-m_2)(j_2+m_2)]^{1/2}\langle j_1-1, j_2-1, m_1, m_2 \mid jm\rangle\\
&\qquad +[(j_1-m_1)(j_1-m_1-1)(j_2+m_2)(j_2+m_2-1)]^{1/2}\langle j_1-1, j_2-1, m_1+1, m_2-1 \mid jm\rangle ,\\
&\quad [(j_1+j_2+j+2)(j_1+j_2-j+1)(j-j_1+j_2+1)(j+j_1-j_2)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = -[(j_1+m_1)(j_1-m_1+1)(j_2+m_2+1)(j_2+m_2+2)]^{1/2}\langle j_1, j_2+1, m_1-1, m_2+1 \mid jm\rangle\\
&\qquad +2m_1[(j_2-m_2+1)(j_2+m_2+1)]^{1/2}\langle j_1, j_2+1, m_1, m_2 \mid jm\rangle\\
&\qquad +[(j_1-m_1)(j_1+m_1+1)(j_2-m_2+1)(j_2-m_2+2)]^{1/2}\langle j_1, j_2+1, m_1+1, m_2-1 \mid jm\rangle ,\\
&[(j_1+j_2-j)(j_1+j_2+j+1)(j+j_1-j_2+1)(j-j_1+j_2)]^{1/2}\langle j_1 j_2 m_1 m_2 \mid jm\rangle\\
&\quad = [(j_1+m_1)(j_1-m_1+1)(j_2-m_2)(j_2-m_2-1)]^{1/2}\langle j_1, j_2-1, m_1-1, m_2+1 \mid jm\rangle
\end{aligned}
$$

$$+2m_1[(j_2-m_2)(j_2+m_2)]^{1/2}\langle j_1,j_2-1,m_1,m_2\,|\,jm\rangle$$
$$-[(j_1-m_1)(j_1+m_1+1)(j_2+m_2)(j_2+m_2-1)]^{1/2}\langle j_1,j_2-1,m_1+1,m_2-1\,|\,jm\rangle,$$

$$[j_1(j_1+1)+j_2(j_2+1)-j(j+1)]\langle j_1j_2m_1m_2\,|\,jm\rangle$$
$$=-[(j_1+m_1)(j_1-m_1+1)(j_2-m_2)(j_2+m_2+1)]^{1/2}\langle j_1,j_2,m_1-1,m_2+1\,|\,jm\rangle-2m_1m_2\langle j_1j_2m_1m_2\,|\,jm\rangle$$
$$-[(j_1-m_1)(j_1+m_1+1)(j_2+m_2)(j_2-m_2+1)]^{1/2}\langle j_1,j_2,m_1+1,m_2-1\,|\,jm\rangle.$$

Needless to say that, all or part of the relations (56) and (57) are probably scattered in the literature. We note for instance that Eq. (56) appears in Ref. 38 [cf. Eq. (3.7.13)] as a function of the $3-jm$ symbol.

As a more sophisticated illustration, we find that the operator

$$j_1k_0^-\mathscr{J}_+\mathscr{K}_++(j_1+1)k_0^+\mathscr{J}_-\mathscr{K}_-$$
$$=j_1(j_1+1)(2j_1+1)[k_0^0+l_0^0-\tfrac{1}{4}k_0^0(\mathscr{K}_+\mathscr{K}_-+\mathscr{K}_-\mathscr{K}_+-\mathscr{J}_+\mathscr{J}_--\mathscr{J}_-\mathscr{J}_+)] \tag{58}$$

produces, after some rearrangement, the relation

$$\left((m_2-m_1)-m\frac{(j_1-j_2)(j_1+j_2+1)}{j(j+1)}\right)\langle j_1j_2m_1m_2\,|\,jm\rangle$$
$$=\frac{[(j+m+1)(j-m+1)(j+j_1+j_2+2)(j_1+j_2-j)(j+j_2-j_1+1)(j+j_1-j_2+1)]^{1/2}}{(j+1)[(2j+1)(2j+3)]^{1/2}}$$
$$\times\langle j_1j_2m_1m_2\,|\,j+1,m\rangle+\frac{[(j+m)(j-m)(j+j_1+j_2+1)(j_1+j_2-j+1)(j+j_2-j_1)(j+j_1-j_2)]^{1/2}}{j[(2j-1)(2j+1)]^{1/2}}$$
$$\times\langle j_1j_2m_1m_2\,|\,j-1,m\rangle, \tag{59}$$

given by Racah[6] in the Condon and Shortley notation and recently rederived by Barut and Wilson and also by Akyeampong and Rashid.[33]

5. BOSON REALIZATIONS OF THE SU_2 UNIT TENSOR

We are now in a position to develop $t_{kq\alpha}$ as a polynomial in the four boson operators $a_\pm$ and $a_\pm^+$.

A. Basic realization

The Majumdar realization: A boson development of $t_{kq\alpha}$, valid as far as $t_{kq\alpha}$ acts on $\epsilon(j)$, is provided by

$$t_{kq\alpha}=\left(\frac{(k+\alpha)!(k-q)!}{(k-\alpha)!(2j+\alpha-k)!(2j+\alpha+k+1)!(k+q)!}\right)^{1/2}(-1)^{k-q}(a_+^+)^{q+\alpha}(a_-)^{q-\alpha}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j+\alpha-q)!(k+q)!}{(q+\alpha)!(k-q)!}+[1-\delta(kq)]\sum_{z=q+\alpha+1}^{k+\alpha}(-1)^z\frac{(2j+2\alpha-z)!(k-\alpha+z)!}{z!(k+\alpha-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j+\alpha+J_z+q-z+t)\right),\quad q\geqslant\alpha\geqslant0, \tag{60}$$

$$t_{k-q\alpha}=\left(\frac{(k+\alpha)!(k-q)!}{(k-\alpha)!(2j+\alpha-k)!(2j+\alpha+k+1)!(k+q)!}\right)^{1/2}(-1)^{2k+q+\alpha}(a_-^+)^{q+\alpha}(a_+)^{q-\alpha}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j+\alpha-q)!(k+q)!}{(q+\alpha)!(k-q)!}+[1-\delta(kq)]\sum_{z=q+\alpha+1}^{k+\alpha}(-1)^z\frac{(2j+2\alpha-z)!(k-\alpha+z)!}{z!(k+\alpha-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j+\alpha-J_z+q-z+t)\right),\quad q\geqslant\alpha\geqslant0, \tag{61}$$

$$t_{kq-\alpha}=\left(\frac{(k+\alpha)!(k-q)!}{(k-\alpha)!(2j-\alpha-k)!(2j-\alpha+k+1)!(k+q)!}\right)^{1/2}(-1)^{2k}(a_-)^{q+\alpha}(a_+^+)^{q-\alpha}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j-\alpha-q)!(k+q)!}{(q+\alpha)!(k-q)!}+[1-\delta(kq)]\sum_{z=q+\alpha+1}^{k+\alpha}(-1)^z\frac{(2j-z)!(k-\alpha+z)!}{z!(k+\alpha-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j-J_z-z+t)\right),\quad q\geqslant\alpha\geqslant0, \tag{62}$$

$$t_{k-q-\alpha}=\left(\frac{(k+\alpha)!(k-q)!}{(k-\alpha)!(2j-\alpha-k)!(2j-\alpha+k+1)!(k+q)!}\right)^{1/2}(-1)^{k-\alpha}(a_+)^{q+\alpha}(a_-^+)^{q-\alpha}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j-\alpha-q)!(k+q)!}{(q+\alpha)!(k-q)!}+[1-\delta(kq)]\sum_{z=q+\alpha+1}^{k+\alpha}(-1)^z\frac{(2j-z)!(k-\alpha+z)!}{z!(k+\alpha-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j+J_z-z+t)\right),\quad q\geqslant\alpha\geqslant0, \tag{63}$$

$$t_{kq\alpha}=\left(\frac{(k+q)!(k-\alpha)!}{(k-q)!(2j+\alpha-k)!(2j+\alpha+k+1)!(k+\alpha)!}\right)^{1/2}(-1)^{k-q}(a_+^+)^{\alpha+q}(a_-^+)^{\alpha-q}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j)!(k+\alpha)!}{(q+\alpha)!(k-\alpha)!}+[1-\delta(k\alpha)]\sum_{z=q+\alpha+1}^{k+q}(-1)^z\frac{(2j+\alpha+q-z)!(k-q+z)!}{z!(k+q-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j+\alpha+J_z+q-z+t)\right),\quad \alpha\geqslant q\geqslant 0, \tag{64}$$

$$t_{k-q\alpha}=\left(\frac{(k+q)!(k-\alpha)!}{(k-q)!(2j+\alpha-k)!(2j+\alpha+k+1)!(k+\alpha)!}\right)^{1/2}(-1)^{2k+q+\alpha}(a_-^+)^{\alpha+q}(a_+^+)^{\alpha-q}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j)!(k+\alpha)!}{(q+\alpha)!(k-\alpha)!}+[1-\delta(k\alpha)]\sum_{z=q+\alpha+1}^{k+q}(-1)^z\frac{(2j+\alpha+q-z)!(k-q+z)!}{z!(k+q-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j+\alpha-J_z+q-z+t)\right),\quad \alpha\geqslant q\geqslant 0, \tag{65}$$

$$t_{kq-\alpha}=\left(\frac{(k+q)!(k-\alpha)!}{(k-q)!(2j-\alpha-k)!(2j-\alpha+k+1)!(k+\alpha)!}\right)^{1/2}(-1)^{2k}(a_-)^{\alpha+q}(a_+)^{\alpha-q}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j-2\alpha)!(k+\alpha)!}{(q+\alpha)!(k-\alpha)!}+[1-\delta(k\alpha)]\sum_{z=q+\alpha+1}^{k+q}(-1)^z\frac{(2j-\alpha+q-z)!(k-q+z)!}{z!(k+q-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j-J_z-z+t)\right),\quad \alpha\geqslant q\geqslant 0, \tag{66}$$

$$t_{k-q-\alpha}=\left(\frac{(k+q)!(k-\alpha)!}{(k-q)!(2j-\alpha-k)!(2j-\alpha+k+1)!(k+\alpha)!}\right)^{1/2}(-1)^{k-\alpha}(a_+)^{\alpha+q}(a_-)^{\alpha-q}$$
$$\times\left((-1)^{q+\alpha}\frac{(2j-2\alpha)!(k+\alpha)!}{(q+\alpha)!(k-\alpha)!}+[1-\delta(k\alpha)]\sum_{z=q+\alpha+1}^{k+q}(-1)^z\frac{(2j-\alpha+q-z)!(k-q+z)!}{z!(k+q-z)!(z-q-\alpha)!}\right.$$
$$\left.\times\prod_{t=1}^{z-q-\alpha}(j+J_z-z+t)\right),\quad \alpha\geqslant q\geqslant 0. \tag{67}$$

Outline of proof: The $j'm'-jm$ matrix element of $t_{kq\alpha}$ as given by Eq. (60) is easily set up by using Eq. (38). The so-obtained relation can be transformed to yield

$$\langle j'm'\,|\,t_{kq\alpha}\,|\,jm\rangle=\delta(j',j+\alpha)\delta(m',m+q)(-1)^{k-q}\left(\frac{(j-m)!}{(j+m)!}\frac{(j+\alpha+m+q)!}{(j+\alpha-m-q)!}\right)^{1/2}$$
$$\times\left(\frac{(k+\alpha)!(k-q)!}{(k-\alpha)!(2j+\alpha-k)!(2j+\alpha+k+1)!(k+q)!}\right)^{1/2}$$
$$\times\sum_z(-1)^z\frac{(2j+2\alpha-z)!(k-\alpha+z)!(j+m)!}{z!(k+\alpha-z)!(z-q-\alpha)!(j+\alpha+m+q-z)!}. \tag{68}$$

Comparison between Eq. (68) and the Majumdar hypergeometric formula[44] specialized to $\langle jkmq|\,j'm'\rangle$ leads us to recognize Eq. (1) in Eq. (68) so that the proof of Eq. (60) is completed. Equations (61)–(67) are then deducible from Eq. (60) owing to appropriate combinations of the following symmetry relations:

$$\langle j_1j_2m_1m_2\,|\,jm\rangle=(-1)^{j_1+j_2-j}\langle j_1j_2-m_1-m_2\,|\,j-m\rangle$$
$$=(-1)^{j_2+m_2}\left(\frac{2j+1}{2j_1+1}\right)^{1/2}\langle jj_2-mm_2\,|\,j_1-m_1\rangle \tag{69}$$

and

$$\langle j_1j_2m_1m_2\,|\,j,m_1+m_2\rangle=(-1)^{j-j_1-m_2}\left(\frac{2j+1}{j+j_1+m_2+1}\right)^{1/2}\langle\tfrac{1}{2}(j_1+j-m_2),j_2,-\tfrac{1}{2}(j-j_1-m_2)+m_1,j-j_1$$
$$\times\,|\,\tfrac{1}{2}(j_1+j+m_2),\tfrac{1}{2}(j-j_1+m_2)+m_1\rangle\,, \tag{70}$$

which describe ordinary[16,38] and Regge[45] symmetry properties of the $SU_2\supset U_1$ Wigner coefficients, respectively.

Numerical examples: As a first (simple) example, we have

$$t_{321}=\left(\frac{4!1!}{2!(2j-2)!(2j+5)!5!}\right)^{1/2}(-1)^5(a_+^+)^3a_-\left((-1)^3\frac{(2j-1)!5!}{3!1!}+(-1)^4\frac{(2j-2)!6!}{4!0!1!}(j+J_z)\right)$$
$$=k_+^+J_+\left(\frac{10(2j-2)!}{(2j+5)!}\right)^{1/2}(3J_z-j+2)\,, \tag{71}$$

in accordance with Ref. 17. As a second (more complicated) example, we get

$$t_{631}=\left(\frac{7!3!}{5!(2j-5)!(2j+8)!9!}\right)^{1/2}(-1)^9(a_+^+)^4(a_-)^2\left((-1)^4\frac{(2j-2)!9!}{4!3!}+(-1)^5\frac{(2j-3)!10!}{5!2!1!}(j+J_z)\right.$$

$$+(-1)^6\frac{(2j-4)!11!}{6!1!2!}(j+J_z-1)(j+J_z)+(-1)^7\frac{(2j-5)!12!}{7!0!3!}(j+J_z-2)(j+J_z-1)(j+J_z)\Big), \tag{72}$$

which can be rearranged as

$$\begin{aligned}t_{631} &= -\left(\frac{9!(2j-5)!}{5!(2j+8)!7!3!}\right)^{1/2}(a_+^+)^4(a_-)^2[(2j-4)(2j-3)(2j-2)\times 210-(2j-4)(2j-3)(j+J_z)\times 1260\\ &\quad +(2j-4)(j+J_z-1)(j+J_z)\times 2310-(j+J_z-2)(j+J_z-1)(j+J_z)\times 1320]\\ &= -6\left(\frac{10(2j-5)!}{(2j+8)!}\right)^{1/2}k_+^+J_+^2(22J_z^3-11jJ_z^2-4j^2J_z+j^3+88J_z^2-41jJ_z-3j^2+142J_z-40j+84),\end{aligned} \tag{73}$$

in diagreement with Ref. 17. The formulas (60)–(67) for $t_{kq\alpha}$ also allow us to handle the case when k is half-an-odd-integer in contradistinction to the known prescriptions[10,17] for constructing operator equivalents. For instance, we obtain

$$t_{17/2\ 11/2\ 7/2}=-10\left(\frac{182(2j-5)!}{5(2j+13)!}\right)^{1/2}(a_+^+)^9(a_-)^2(34J_z^3+12j^2J_z-42jJ_z^2+186J_z^2+33j^2-177jJ_z+374J_z-198j+264). \tag{74}$$

The operator construction: As a particular case, Eq. (64) shows that

$$t_{jmj}=(-1)^{2j}(2j+1)^{-1/2}[(j+m)!(j-m)!]^{-1/2}(a_+^+)^{j+m}(a_-^+)^{j-m},\quad m\geqslant 0, \tag{75}$$

is valid when acting on the boson vacuum space $\epsilon(0)$. On the other hand, Eq. (1) yields

$$|jm\rangle=(-1)^{-2j}(2j+1)^{1/2}t_{jmj}|00\rangle\,. \tag{76}$$

Combination of Eqs. (75) and (76) gives

$$|jm\rangle=[(j+m)!(j-m)]^{-1/2}(a_+^+)^{j+m}(a_-^+)^{j-m}|00\rangle\,, \tag{77}$$

in agreement with Eq. (1.13) of Ref. 5 describing the boson-like operator construction of the angular momentum eigenvectors.

The case $k=\frac{1}{2}$: Equations (60)–(67) show that the four basic boson operators $a_\pm$ and $a_\pm^+$ are connected to the unit tensor operators $t_{1/2\ q\alpha}$ acting on $\epsilon(j)$ through

$$t_{1/2\ \pm 1/2\ 1/2}=-[2(j+1)(2j+1)\ ]^{-1/2}a_\pm^+,\quad t_{1/2\ \pm 1/2\ -1/2}=-[2j(2j+1)\ ]^{-1/2}a_\mp\,. \tag{78}$$

Let $W^k(j,\alpha)$ be the SU_2 tensor defined by its components

$$W_q^k(j,\alpha)=(-1)^{-2k}[2(j+\alpha)+1\ ]^{1/2}t_{kq\alpha} \tag{79}$$

acting on $\epsilon(j)$. By making use of Eqs. (79) and (78), it can be verified that the Schwinger operators defined by Eq. (51) can be nicely rewritten in terms of zeroth-order tensor products:

$$\begin{aligned}\mathscr{J}_3 &= -\frac{1}{\sqrt{2}}[(2j_1+1)\{W^{1/2}(j_1,\tfrac{1}{2})W^{1/2}(j_1,-\tfrac{1}{2})\}_0^0-(2j_2+1)\{W^{1/2}(j_2,\tfrac{1}{2})W^{1/2}(j_2,-\tfrac{1}{2})\}_0^0],\\ \mathscr{J}_\pm &= \mp[2(2j_1+1)(2j_2+1)\ ]^{1/2}\{W^{1/2}(j_1,\pm\tfrac{1}{2})W^{1/2}(j_2,\mp\tfrac{1}{2})\}_0^0,\\ \mathscr{K}_3 &= -\frac{1}{\sqrt{2}}[(2j_1+1)\{W^{1/2}(j_1,\tfrac{1}{2})W^{1/2}(j,-\tfrac{1}{2})\}_0^0+(2j_2+1)\{W^{1/2}(j_2,\tfrac{1}{2})W^{1/2}(j_2,-\tfrac{1}{2})\}_0^0]+1,\\ \mathscr{K}_\pm &= -[2(2j_1+1)(2j_2+1)\ ]^{1/2}\{W^{1/2}(j_1,\pm\tfrac{1}{2})W^{1/2}(j_2,\pm\tfrac{1}{2})\}_0^0.\end{aligned} \tag{80}$$

The case k integer: As a particular case, when k is integer, Eqs. (60)–(67) can be rewritten in terms of the k's by remarking that

$$\begin{aligned}&(a_+^+)^{q+\alpha}(a_-)^{q-\alpha}=(-1)^\alpha(k_+^+)^\alpha(k_+^0)^{q-\alpha},\\ &(a_-^+)^{q+\alpha}(a_+)^{q-\alpha}=(k_-^+)^\alpha(k_-^0)^{q-\alpha},\\ &(a_-)^{q+\alpha}(a_+^+)^{q-\alpha}=(k_+^-)^\alpha(k_+^0)^{q-\alpha},\\ &(a_+)^{q+\alpha}(a_-^+)^{q-\alpha}=(-1)^\alpha(k_-^-)^\alpha(k_-^0)^{q-\alpha},\\ &(a_+^+)^{\alpha+q}(a_-^+)^{\alpha-q}=(-1)^q(k_+^+)^q(k_0^+)^{\alpha-q},\\ &(a_-^+)^{\alpha+q}(a_+^+)^{\alpha-q}=(k_-^+)^q(k_0^+)^{\alpha-q},\\ &(a_-)^{\alpha+q}(a_+)^{\alpha-q}=(k_+^-)^q(k_0^-)^{\alpha-q},\\ &(a_+)^{\alpha+q}(a_-)^{\alpha-q}=(-1)^q(k_-^-)^q(k_0^-)^{\alpha-q}\end{aligned} \tag{81}$$

hold once k, and therefore q and α, are integers.

The case $k=1$: Equations (60)–(67) show that the nine basic operators k_σ^ρ are connected to the unit tensor operators $t_{1q\alpha}$ acting on $\epsilon(j)$ through

$$t_{100} = [\,j(j+1)(2j+1)\,]^{-1/2} k_0^0, \quad t_{1\pm 10} = \mp\,[2j(j+1)(2j+1)\,]^{-1/2} k_\pm^0,$$
$$t_{101} = [(j+1)(2j+1)(2j+3)\,]^{-1/2} k_0^+, \quad t_{1\pm 11} = \mp\,[(2j+1)(2j+2)(2j+3)\,]^{-1/2} k_\pm^+, \tag{82}$$
$$t_{10-1} = -\,[\,j(2j-1)(2j+1)\,]^{-1/2} k_0^-, \quad t_{1\pm 1-1} = \pm\,[2j(2j-1)(2j+1)\,]^{-1/2} k_\pm^-.$$

By using Eqs. (79) and (82), the operators defined by Eq. (55) can be rewritten as

$$\begin{aligned}
&\mathscr{K}_-\mathscr{K}_- = 2[3j_1(2j_1+1)j_2(2j_2+1)\,]^{1/2}\{W^1(j_1,-1)W^1(j_2,-1)\,\}_0^0,\\
&\mathscr{J}_-\mathscr{J}_- = 2[3j_1(2j_1+1)\,(j_2+1)\,(2j_2+1)\,]^{1/2}\{W^1(j_1,-1)W^1(j_2,1)\,\}_0^0,\\
&\mathscr{J}_-\mathscr{K}_- = -\,2[3j_1(2j_1+1)j_2(j_2+1)\,]^{1/2}\{W^1(j_1,-1)W^1(j_2,0)\}_0^0,\\
&\mathscr{J}_+\mathscr{J}_+ = 2[3(j_1+1)\,(2j_1+1)j_2(2j_2+1)\,]^{1/2}\{W^1(j_1,1)W^1(j_2,-1)\,\}_0^0,\\
&\mathscr{K}_+\mathscr{K}_+ = 2[3(j_1+1)(2j_1+1)(j_2+1)(2j_2+1)\,]^{1/2}\{W^1(j_1,1)W^1(j_2,1)\}_0^0,\\
&\mathscr{J}_+\mathscr{K}_+ = 2[3(j_1+1)(2j_1+1)j_2(j_2+1)\,]^{1/2}\{W^1(j_1,1)W^1(j_2,0)\}_0^0,\\
&\mathscr{J}_+\mathscr{K}_- = 2[3j_1(j_1+1)j_2(2j_2+1)\,]^{1/2}\{W^1(j_1,0)W^1(j_2,-1)\,\}_0^0,\\
&\mathscr{J}_-\mathscr{K}_+ = -\,2[3j_1(j_1+1)(j_2+1)(2j_2+1)\,]^{1/2}\{W^1(j_1,0)W^1(j_2,1)\}_0^0,\\
&\tfrac{1}{4}(\mathscr{K}_+\mathscr{K}_- + \mathscr{K}_-\mathscr{K}_+ - \mathscr{J}_+\mathscr{J}_- - \mathscr{J}_-\mathscr{J}_+ - 2) = 2\,[3j_1(j_1+1)j_2(j_2+1)\,]^{1/2}\{W^1(j_1,0)W^1(j_2,0)\}_0^0.
\end{aligned} \tag{83}$$

The diagonal unit tensor: As a limiting case, when $\alpha = 0$, Eqs. (60)–(67) reduce to

$$u^k_{\pm q} = \left(\frac{(k-q)!}{(k+q)!(2j-k)!(2j+k+1)!}\right)^{1/2}(-1)^{(k\pm k+2q)/2} J_\pm{}^q\left((-1)^q\frac{(2j-q)!(k+q)!}{q!(k-q)!}\right.$$
$$\left. + [1-\delta(kq)]\sum_{z=q+1}^{k}(-1)^z\frac{(2j-z)!(k+z)!}{z!(k-z)!(z-q)!}\prod_{t=1}^{z-q}(j\pm J_z+q-z+t)\right), \quad q\geqslant 0. \tag{84}$$

Specialization of Eq. (84) leads to the well-known result

$$u^k_{\pm k} = (-1)^{(k\pm k)/2}\frac{1}{k!}\left(\frac{(2k)!(2j-k)!}{(2j+k+1)!}\right)^{1/2} J_\pm{}^k. \tag{85}$$

We note that Eqs. (84) and (85) are in agreement with the lemma[19] according to which

$$\{J_\pm{}^r J_z{}^s : s = 0(1)(2j-r);\ r = 0(1)2j\}$$

constitutes a basis of τ.

As a by-product, it is to be noted that Eq. (84) supplies a trick for obtaining a closed form expression of any harmonic polynomial $r^k Y_{kq}$ in Cartesian coordinates. It is sufficient in Eq. (84): (i) to ignore all terms of order lower than k which are of quantum-mechanical origin, (ii) to make the replacements $J_\pm \to x \pm iy, J_z \to z$, and $J^2 \to r^2 = x^2+y^2+z^2$, and (iii) to multiply the so-obtained expression by the factor

$$\frac{1}{2^k}\left(\frac{2k+1}{4\pi}\right)^{1/2}\left(\frac{(2j+k+1)!}{(2j-k)!}\right)^{1/2}.$$

Such a trick (cf. also Ref. 16) constitutes the reciprocal part of the Stevens prescription described in Sec. 1. As an illustration, Eq. (84) yields

$$u_4^7 = 4\sqrt{462}\left(\frac{(2j-7)!}{(2j+8)!}\right)^{1/2} J_+{}^4(13J_z^3+78J_z^2+179J_z-3J_zJ^2-6J^2+150) \tag{86}$$

and application of the trick leads to

$$r^7 Y_{74} = \tfrac{3}{64}(770/\pi)^{1/2}(x+iy)^4(13z^3-3zr^2). \tag{87}$$

A trick corresponding to the reciprocal part of the Atkins and Seymour[17] prescription could also be stated. Nevertheless, its interest would be quasinull except for checking purposes.

The algebra su_3: Equations (60)–(67) show that the set $\{u_q^k : q$ ranging and $k = 0(1)2\}$ acting on $\epsilon(1)$ can be realized by

$$\begin{aligned}
&u_0^0 = \frac{1}{\sqrt{3}}, \quad u_0^1 = \frac{1}{\sqrt{6}}J_z, \quad u^1_{\pm 1} = \mp\frac{1}{2\sqrt{3}}J_\pm,\\
&u_0^2 = \frac{1}{\sqrt{30}}(3J_z^2-2), \quad u^2_{\pm 1} = \frac{1}{\sqrt{20}}J_\pm(-1\mp 2J_z), \quad u^2_{\pm 2} = \frac{1}{\sqrt{20}}J^2{}_\pm.
\end{aligned} \tag{88}$$

It can be verified that Eq. (88) satisfies the commutation relations (32) specialized to the case $G^* = U_1$.

B. Other realizations

We now turn to list some alternative boson realizations of $t_{kq\alpha}$.

The van der Waerden realization:

$$t_{kq\alpha} = (-1)^{k+\alpha}\left(\frac{(k+q)!(k-q)!(k+\alpha)!(k-\alpha)!(2j+\alpha-k)!}{(2j+\alpha+k+1)!}\right)^{1/2} \times \sum_z (-1)^z \frac{(a_-^+)^{k-q-z}(a_-)^{k-\alpha-z}(a_+^+)^{q+\alpha+z}(a_+)^z}{(k-q-z)!(k-\alpha-z)!(q+\alpha+z)!z!}. \tag{89}$$

The Zukauskas–Mauza realization:

$$t_{kq\alpha} = (-1)^{q+\alpha}\left(\frac{(k+q)!(k-q)!(k+\alpha)!(k-\alpha)!(2j+\alpha-k)!}{(2j+\alpha+k+1)!}\right)^{1/2} \times \sum_z (-1)^z \frac{(a_+)^{k-q-z}(a_+^+)^{k+\alpha-z}(a_-)^{q-\alpha+z}(a_-^+)^z}{(k-q-z)!(k+\alpha-z)!(q-\alpha+z)!z!}. \tag{90}$$

The Wigner realization:

$$t_{kq\alpha} = (-1)^{k-q}\left(\frac{(k+\alpha)!(k-\alpha)!(2j+\alpha-k)!}{(k+q)!(k-q)!(2j+\alpha+k+1)!}\right)^{1/2} \times \sum_z (-1)^z \frac{(a_+^+)^z(a_+)^{k-q}(a_+^+)^{k+\alpha-z}(a_-^+)^{k+\alpha-z}(a_-)^{k+q}(a_-^+)^z}{(k+\alpha-z)!z!}. \tag{91}$$

The Racah realization:

$$t_{kq\alpha} = (-1)^{2k}\left(\frac{(k+q)!(k-q)!(2j+\alpha-k)!}{(k+\alpha)!(k-\alpha)!(2j+\alpha+k+1)!}\right)^{1/2} \times \sum_z (-1)^z \frac{(a_+^+)^{k+q-z}(a_+)^{k-\alpha}(a_+^+)^z(a_-)^{k+q-z}(a_-^+)^{k+\alpha}(a_-)^z}{(k+q-z)!z!}. \tag{92}$$

Outline of proof: It is sufficient to verify that the $j'm' - jm$ matrix element of $t_{kq\alpha}$ as given by Eqs. (89)–(92) verifies Eq. (1) with the $SU_2 \supset U_1$ Clebsch–Gordan coefficient responding to the formulas of van der Waerden,[46] Zukauskas and Mauza,[47] Wigner,[48] and Racah.[6]

In all honesty, it is to be mentioned that the van der Waerden realization (89) turns out to be a rewriting in our terminology of the boson representation of the $SU_2 \supset U_1$ Wigner–Eckart theorem recently derived by Yamamura *et al.*[23]

C. General comments

At this point, a brief comparison between the four latter realizations and the Majumdar realization is in order. First, in the case when k is integer, inspection of Eqs. (60)–(67) and (89)–(92) shows that only the Majumdar realization of $t_{kq\alpha}$ can be easily rewritten in terms of the operators k_σ^ρ defined by Eq. (39). Second, this realization allows one to get, with merely simple algebraic manipulations, diagonal and off-diagonal operator equivalents in a form very close to the one underlying the already published tables.[10–13,17,19] Such a form could be also obtainable in principle from the other realizations. However, this would require long winded manipulations of commutation relations. Third, Eqs. (60)–(67) are particularly appropriate to hand and machine calculations. For instance, the determination of the J_z polynomial part of the Majumdar realization of $t_{kq\alpha}$ amounts in last analysis to solving a Cramer system. In this connection, Eqs. (60)–(67) have been coded on a computer to produce tables of diagonal ($\alpha = 0$) and off-diagonal ($\alpha \neq 0$) operator equivalents $t_{k\alpha\Gamma\gamma\alpha}$ adapted to the cubical, tetragonal, and trigonal groups for use in molecular and solid state physics.[49]

To close Sec. 5, it is perhaps worth stressing that the operator $W_q^k(j,\alpha)$ defined by Eq. (79) is actually a $SU_2 \supset U_1$ Wigner operator in the sense that (cf. Appendix A)

$$\langle j'm' | W_q^k(j,\alpha) | jm \rangle = \delta(j'j+\alpha)\langle jkmq | j'm' \rangle . \tag{93}$$

In the same vein, a SU_2 Racah operator $R(j_1\alpha_1 j_2\alpha_2 jk)$ may be defined via (cf. Appendix A)

$$\langle j_1'j_2'j'm' | R(j_1\alpha_1 j_2\alpha_2 jk) | j_1j_2jm \rangle = \delta(j_1'j_1+\alpha_1)\delta(j_2'j_2+\alpha_2)\delta(j'j)\delta(m'm)W(j_1j_2j_1+\alpha_1 j_2+\alpha_2 jk). \tag{94}$$

It is immediate to prove that (cf. Appendix A)

$$R(j_1\alpha_1 j_2\alpha_2 jk) = (-1)^{j_1+\alpha_1+j_2-j-k}[2(j_1+\alpha_1)+1]^{-1/2}[2(j_2+\alpha_2)+1]^{-1/2}(2k+1)^{1/2}\{W^k(j_1,\alpha_1)W^k(j_2,\alpha_2)\}_0^0 \tag{95}$$

indeed verifies Eq. (94). Boson realizations of $R(j_1\alpha_1 j_2\alpha_2 jk)$ can be obtained by combining Eqs. (79) and (95) with the various boson realizations of $t_{kq\alpha}$ given in this section. In this respect, the Schwinger operators $\mathscr{J}_\pm$ and $\mathscr{K}_\pm$ are of type $R(j_1\alpha_1 j_2\alpha_2 j\frac{1}{2})$ while the operators $\mathscr{J}_\rho\mathscr{J}_\sigma$, $\mathscr{J}_\rho\mathscr{K}_\sigma$, and $\mathscr{K}_\rho\mathscr{K}_\sigma$ are of type $R(j_1\alpha_1 j_2\alpha_2 j1)$.

6. CONCLUDING REMARKS

We started from the general properties of the unit tensor operators $t_{k\mu\alpha}$ adapted to a chain $SU_2 \supset G^*$ which can be obtained without specifying the realization and ended with boson realizations of the operators $t_{kq\alpha}$ adapted to the chain $SU_2 \supset U_1$.

Among the $t_{k\mu\alpha}$'s, we have paid special attention to the set $\{t_{1q\alpha} : q \text{ and } \alpha \text{ ranging}\}$. Such a nine-dimensional set is easily completed with a tenth operator to produce the ten-dimensional Schwinger algebra. We have converted this algebra to the Lie algebra of the de Sitter group $SO_{3,2}$. It is to be noted that the Lie group $SO_{3,2}$ has already been exhibited in connection with Schwinger's calculus. Indeed, Györghyi and Kövesi-Domokos[27] have proved that $O_{3,2}$ enters the spherical rotor problem as a dynamical group: The $O_{3,2}$ group can be realized in terms of Dirac brackets of the Cartesian coordinates and the canonically conjugate momenta for the spherical rotor. It is also to be noted that $SO_{3,2}$ has been exhibited in an apparently distant field, namely, the Coulomb problem. In fact, Englefield[29] has obtained an $O_{3,2}$ Lie algebra in terms of ladder operators acting on the hydrogen-like eigenfunctions and therefore similar to the operators k^{ρ}_{σ} although they are constructed without any relation to the Jordan–Schwinger representation of orbital angular momentum. The de Sitter group $SO_{3,2}$ thus appears as the most general structure of relevance for the Schwinger two-dimensional oscillator approach to the angular momentum theory.

The coupled-boson representation describing the coupling of two angular momenta then appears to be depicted by $SO_{3,2} \otimes SO_{3,2}$. We have shown how the simple manipulation of ladder operators of type: (i) coupled-boson operators $\mathscr{J}$ and $\mathscr{K}$ (related to $SO_{3,1}$) and (ii) bilinear expressions in $\mathscr{J}$ and $\mathscr{K}$ (related to the enveloping algebra of $SO_{3,2} \otimes SO_{3,2}$) allow to obtain, without any intermediate step, basic families of recursion relations for $SU_2 \supset U_1$ Wigner coefficients. We conjecture that any recursion relation for the $SU_2 \supset U_1$ Wigner coefficients (and possibly the SU_2 Racah coefficients) could be reached on the same pattern through a detailed study of the chain $SO_{3,2} \otimes SO_{3,2} \supset SO_{3,2} \supset SO_{3,1}$. We are inclined to think our procedure to get recursion relations resembles (although we have been unable to correctly connect the two procedures) Shaw's procedure[33] based on the use of a Lorentz group $\mathscr{L} \uparrow (C_2)$. The real reason for this is to be found in the fact that both procedures are very close, in a certain way, to the Bargmann[43] treatment of the three-dimensional rotation group.

The various boson realizations of the $SU_2 \supset U_1$ unit tensor operators $t_{kq\alpha}$ settle an important result of this study. We limited ourself to five realizations which mimick five basic formulas for the $SU_2 \supset U_1$ Wigner coefficient. Although we conjecture it is possible to get a boson realization of $t_{kq\alpha}$ from any factorial expression of the $SU_2 \supset U_1$ Wigner coefficient, actually, the Majumdar realization seems to be the most efficient for producing polarized harmonics or operator equivalents in a form suitable for subsequent application. Of course, it should be possible to pass from one realization to another by using either the commutation relations of the four basic boson operators or symmetries of the $SU_2 \supset U_1$ Wigner coefficient. For example, the Zukauskas–Mauza and the Wigner realizations are linked by an ordinary symmetry while the two Majumdar realizations (60) and (65) are connected by a Regge symmetry. This suggests that the 72 ordinary and Regge symmetries of the $3-jm$ symbol could be handled by looking at the symmetries of the $t_{kq\alpha}$'s.

The properties of the $t_{k\mu\alpha}$'s can be classified in two parts. Some of them clearly show that the $t_{k\mu\alpha}$'s serve as a basis for expanding physical interactions. Some others emphasize the interest of the $t_{k\mu\alpha}$'s for Lie group theoretical analyses of the spectrum of physical interactions. We close by mentioning some avenues of future outlook.

Chain of groups: In the case k integer, Eqs. (60)–(67) can be considered as a solution to the problem of constructing $so_3 \supset so_2$ operators in the enveloping algebra of $so_{3,2}$. A related problem, of interest for producing recursion relations, would be the state labelling problem associated with the chain $SO_{3,2} \otimes SO_{3,2} \supset SO_{3,2} \supset SO_{3,1}$. The two mentioned problems require for the part $SO_{3,2} \supset SO_3 \otimes SO_2 \supset SO_3 \supset SO_2$ the consideration of a set of four commuting operators taken from the enveloping algebra of $so_{3,2}$ since, in addition to the two (second and fourth order) invariants of $SO_{3,2}$, two extra labelling operators are needed to completely label the $SO_{3,2} \supset SO_3 \supset SO_2$ states. The method of the generating function recently developed by the Montréal group[50] could be a good starting point for tackling these problems.

Half-integer rank tensor: It is generally more or less stated in physics textbooks that only integer rank irreducible tensors are suitable for physical applications. We ascertain, however, that calculation of intensities in x-ray and ultraviolet photoelectron spectroscopy[51] as well as in the recently developed bremsstrahlung isochromat spectroscopy[52] for transition metal and rare earth ions in crystalline matrices requires the introduction of half-an-odd-integer irreducible tensors. We hope that part of the material reported here shall be useful in those fields where half-integer rank tensors are of interest.

Theory of complex spectra: Specialization to the case $G^* = U_1$ and inversion of Eq. (27) yields

$$\delta(j,k,j)u^k_q = \sum_{mm'} (-1)^{j-m} \begin{pmatrix} j & k & j \\ -m & q & m' \end{pmatrix} P^j_{mm'} . \qquad (96)$$

Equation (96) curiously bears the same form that the intrashell operators introduced by Harter and Patterson[53] in their alternative (Gel'fand) basis for the theory of complex spectra. This can be easily understood by realizing that the set $\{P^j_{\mu\mu'} : \mu \text{ and } \mu' \text{ ranging}\}$ furnishes a basis of the Lie algebra of the unitary group U_{2j+1}. We felt it would be worthwhile to explore these matters in more detail.

Projection operator: The projection operator[40] $P_{jm} = P^j_{mm}$ [cf. Eq. (26)] can be realized in terms of boson operators by introducing the boson realizations of u^k_q obtainable from Sec. 5 into Eq. (27). This leads to expressions for P_{jm} that can be proven to be equivalent to the von Neumann (product) form and the Löwdin (sum and product) forms of P_{jm}. We are presently examining the respective merits, with special reference to computer coding, of the various possible realizations of P_{jm}.

Lie superalgebra: In recent years, several works have been devoted to mathematical studies of low dimensional Z_2-graded Lie algebras [cf. Ref. 54]. The concept of the Z_2-graded Lie algebra is particularly appropriate for describing fermion–bose supersymmetry in various contexts as, e.g., in the one of space–time symmetries for particle physics. We note the (infinite) set $\{t_{k\mu\alpha} : \alpha, \mu, \text{ and } k \text{ ranging}\}$ is closed

under anticommutation and commutation so that the $t_{k\mu\alpha}$'s could be used for generating Lie superalgebras.

Factorization method: It has been already noticed by Witschel[42] that there exists a connection between the Schrödinger–Infeld–Hull factorization method[55] and the Schwinger boson calculus. A Lie-like aspect of this connection arises in this paper. In fact, the Schwinger algebra admits several subalgebras of type $\mathcal{G}(a,b)$, where $\mathcal{G}(a,b)$ denotes the complex four-dimensional Lie algebra discussed by Miller[55] in relation with the factorization method. In the same vein, let us mention that the realizations of $t_{kq\alpha}$ given here enable to obtain useful expressions for the O_3 shift operators recently investigated by Hughes and Yadegar[56] and which turn out to be relevant to the factorization method and to the classification and analysis of representations of various Lie groups.

Symplecton calculus: The Jordan–Schwinger boson representation of angular momentum has been generalized by Biedenharn and Louck to give the so-called symplecton representation.[57] Such a generalization leads us to replace the boson calculus by a symplecton calculus. It would be (at least curious and perhaps) interesting, especially in view of the rotational SU_3 nuclear model which motivated the introduction of the symplecton representation, to translate some of the results of this paper in the language of symplecton operators.

ACKNOWLEDGMENTS

One of us (M.K.) would like to thank J. Paldus and J. Čízek for their kind and stimulating hospitality at the Quantum Theory Group of the University of Waterloo in the fall of 1978, during which time part of this work was completed. He would also like to thank J. Patera and J. Harnad (Université de Montréal), R.T. Sharp (McGill University), and K.B. Wolf (Universidad Nacional Autonoma de Mexico) for helpful conversations which provided the starting point of the $so_{3,2}$ analysis of the Schwinger algebra. The authors are grateful to J. Sau (Université Lyon-I) for calling their attention to Refs. 23 and 25.

APPENDIX A: ON THE WIGNER–RACAH ALGEBRA OF A COMPACT GROUP

Notation: Let J be a compact topological (finite or continuous) group. We use j or k to denote an IRC of J. Then, j_0 stands for the identity IRC of J. Let $\epsilon(j) = \{|jm\rangle : m \text{ ranging}\}$ be an irreducible subspace, associated with j, of the representation space $\mathcal{E}$ of J. We note by $[j]$ the dimension of $\epsilon(j)$. A definite choice of basis is assumed for each space $\epsilon(j)$, thus defining the $\{m\}$ scheme, and the various J irreducible tensorial sets of vectors and operators of $\mathcal{E}$ are required to transform according to this scheme. Further, $\langle j_1 j_2 m_1 m_2 | b\, jm\rangle$ is a Wigner coefficient of J compatible with the $\{m\}$ scheme, b being a Kronecker multiplicity label to be used when j occurs several times in $j_1 \otimes j_2$. Finally, $\Delta(j|j_1 \otimes j_2) = 0$ or 1 according to whether the frequency of j in $j_1 \otimes j_2$ is equal to 0 or different from 0. The notation for a subgroup G of J can be deduced from the preceding ones thanks to the correspondences: $j \to \Gamma$, $m \to \gamma$, and $b \to \beta$. The subduction $J \downarrow G$ amounts in many respects to replacing m by $\mu \equiv a\Gamma\gamma$, where a is a branching multiplicity label, thus defining the $\{a\Gamma\gamma\}$ scheme. The Wigner coefficients of J in the $\{a\Gamma\gamma\}$ scheme are referred to as $J \supset G$ Wigner coefficients.

The orthogonality-completeness property: The $J \supset G$ Wigner coefficients satisfy a well-known unitary property. In view of this property and the completeness property for the representation matrices of J in the $\{a\Gamma\gamma\}$ scheme, we can derive the following orthogonality-completeness relation[36]

$$\sum_{j_2 \mu_2 b} [j_2] \langle j_1 j_2 \mu_1 \mu_2 | b\, j\mu\rangle \langle j_1 j_2 \mu_1' \mu_2 | b\, j\mu'\rangle^* = \delta(\mu_1' \mu_1)\delta(\mu'\mu)[j]$$

and its dual relation

$$\sum_{\mu_1 \mu} \langle j_1 j_2 \mu_1 \mu_2 | b\, j\mu\rangle^* \langle j_1 j_2' \mu_1 \mu_2' | b'\, j\mu\rangle = \Delta(j|j_1 \otimes j_2)\delta(j_2' j_2)\delta(\mu_2' \mu_2)\delta(b'b)[j_2]^{-1}[j].$$

Similar relations hold when replacing $\Sigma_{j_2 \mu_2 b}$ and $\Sigma_{\mu_1 \mu}$ by $\Sigma_{j_1 \mu_1 b}$ and $\Sigma_{\mu_2 \mu}$, respectively.

The Wigner operator: A unit tensor operator $\hat{t}_{j_2 \mu_2 bj}$ may be defined through

$$\langle j'\mu | \hat{t}_{j_2 \mu_2 bj} | j_1 \mu_1\rangle = \delta(j'j) \exp[i\varphi(b\, j j_1 j_2)][j]^{-1/2} \langle j_1 j_2 \mu_1 \mu_2 | b\, j\mu\rangle^*,$$

where φ is a J-dependent phase factor. We call the operator

$$\hat{W}^{j_2}_{\mu_2}(bj) = \exp[-i\varphi(bjj_1 j_2)][j]^{1/2} t_{j_2 \mu_2 bj}$$

a Wigner operator since its $j\mu - j_1 \mu_1$ matrix element is nothing but than $\langle j_1 j_2 \mu_1 \mu_2 | b\, j\mu\rangle^*$. Application to $\hat{W}^{j_2}_{a_2 \Gamma_2 \gamma_2}(bj)$ of the Wigner–Eckart theorem for G in the $\{a\Gamma\gamma\}$ scheme directly leads to Racah's lemma:

$$\langle j_1 j_2 a_1 \Gamma_1 \gamma_1 a_2 \Gamma_2 \gamma_2 | b\, ja\Gamma\gamma\rangle = \sum_{\beta} (j_1 a_1 \Gamma_1 + j_2 a_2 \Gamma_2 | b\, ja\beta\Gamma) \langle \Gamma_1 \Gamma_2 \gamma_1 \gamma_2 | \beta\Gamma\gamma\rangle,$$

where $(j_1 a_1 \Gamma_1 + j_2 a_2 \Gamma_2 | b\, ja\beta\Gamma)$ is, in modern parlance, a $J \supset G$ isoscalar factor. This connection between the Racah lemma and the Wigner–Eckart theorem takes its origin in the fact that they are both more or less direct corollaries of Schur's lemma.

The Racah operator: As a consequence of the great orthogonality theorem (again a corollary of Schur's lemma!) for the representation matrices of J in the $\{a\Gamma\gamma\}$ scheme, the recoupling coefficients

$$\langle j_1 (j_2 j_3) b_{23} j_{23} b'j'\mu' | (j_1 j_2) b_{12} j_{12} b_3 b\, j\mu\rangle = \sum_{\substack{\mu_1 \mu_2 \mu_3 \\ \mu_{12} \mu_{23}}} \langle j_1 j_2 \mu_1 \mu_2 | b_{12} j_{12} \mu_{12}\rangle \langle j_{12} j_3 \mu_{12} \mu_3 | b\, j\mu\rangle \times \langle j_2 j_3 \mu_2 \mu_3 | b_{23} j_{23} \mu_{23}\rangle^* \langle j_1 j_{23} \mu_1 \mu_{23} | b'j'\mu'\rangle^*$$

enjoy the property

$$\langle j_1 (j_2 j_3) b_{23} j_{23} b'j'\mu' | (j_1 j_2) b_{12} j_{12} j_3 b\, j\mu\rangle = \delta(j'j)\delta(\mu'\mu) \langle j_1 (j_2 j_3) b_{23} j_{23} b'j | (j_1 j_2) b_{12} j_{12} j_3 b\, j\rangle,$$

where

$$\langle j_1 (j_2 j_3) b_{23} j_{23} b'j | (j_1 j_2) b_{12} j_{12} j_3 b\, j\rangle$$

$$= [j]^{-1}\sum_{\mu} \langle j_1(j_2 j_3)b_{23}j_{23}b'j\mu|(j_1 j_2)b_{12}j_{12}j_3 b j\mu\rangle$$

is a J Racah coefficient.

A Racah operator $\hat{R}(j_1 j_1' j_2 j_2' jkbb'b_1' b_2')$ may be defined through

$$\langle (j_1'' j_2'')b'j'\mu'|\hat{R}(j_1 j_1' j_2 j_2' jkbb'b_1' b_2')|(j_1 j_2)bj\mu\rangle$$
$$= \delta(j_1'' j_1')\delta(j_2'' j_2')\delta(j'j)\delta(\mu'\mu)$$
$$\times [j]^{-1/2}[k]^{-1/2}\langle j_1(j_2 j_2')b'kb_2' j_1'|(j_1 j_2)bj j_2' b_1' j_1'\rangle.$$

The Racah operator $\hat{R}(j_1 j_1' j_2 j_2' jkbb'b_1' b_2')$ is connected to the Wigner operators $\hat{W}^k(b_1' j_1')$ and $\hat{W}^k(b_2' j_2')$ acting on two distinct spaces $\mathscr{E}_1$ and $\mathscr{E}_2$ by the tensor product relation

$$\hat{R}(j_1 j_1' j_2 j_2' jkbb'b_1' b_2') \sim \{\hat{W}^k(b_1' j_1')\hat{W}^k(b_2' j_2')\}^{j_0}$$

in the case where k is orthogonal or symplectic.

Commutation relations: The commutator acting on $\epsilon(j)$ of two Wigner operators is given by[36]

$$\left[\hat{W}^{k_1}_{\mu_1}(b_1 j_1), \hat{W}^{k_2}_{\mu_2}(b_2\mu_2)\right]_{\pm}$$
$$= \sum_{bk_3\mu_3 b_3 j_3} \hat{W}^{k_3}_{\mu_3}(b_3 j_3)$$
$$\times [\delta(j_3 j_1)\langle j(k_2 k_1)bk_3 b_3 j_3|(jk_2)b_2 j_2 k_1 b_1 j_1\rangle$$
$$\times \langle k_2 k_1 \mu_2\mu_1|bk_3\mu_3\rangle$$
$$\pm \delta(j_3 j_2)\langle j(k_1 k_2)bk_3 b_3 j_3|(jk_1)b_1 j_1 k_2 b_2 j_2\rangle$$
$$\times \langle k_1 k_2 \mu_1\mu_2|bk_3\mu_3\rangle]^*.$$

The proof is based on recoupling techniques. From the definition of the Racah coefficient as a function of the Wigner coefficient, we may derive, by making use of the unitarity and orthogonality-completeness properties of the Wigner coefficients, a relation that corresponds to the matrix elements of the desired commutation relation, thus completing the proof.

The latter relation generalizes the ones obtained in the diagonal case ($j_1 = j_2 = j_3 = j$) for: (i) the chain $SO_3 \supset SO_2$ in connection with atomic and nuclear spectroscopy[9,34] and (ii) a finite group in connection with the Jahn–Teller effect.[41]

Lie algebra and Wigner–Racah algebra: In the diagonal case ($j_1 = j_2 = j_3 = j$), it can be seen that the general commutation relation of $\hat{W}^{k_1}_{\mu_1}(b_1 j)$ and $\hat{W}^{k_2}_{\mu_2}(b_2 j)$ defines the structure of the Lie group $GL_{[j];\mathbb{R}}$ and of its subgroups. This result gives a precise meaning to the word algebra in what is generally referred to as Wigner–Racah algebra. The Lie algebra $\oplus_j gl_{[j];\mathbb{R}}$ can thus be associated with the Wigner–Racah algebra of the compact group J. It should be observed that the basic ingredients, viz, (i) the Wigner operators of J, (ii) the (basis- or G-dependent) Wigner coefficients of J, and (iii) the (basis- or G-independent) Racah coefficients of J, all appear in the defining Lie law of the Wigner-Racah algebra of J.

We establish now a parallel between this Appendix and the main body of the present paper. In the case $J = SU_2$, there is no multiplicity label b. Assuming the phase choice $\varphi(j_1 j_2 k) = 2k\pi$, the operator $\hat{t}_{k,q,j+\alpha}$ identifies to $t_{kq\alpha}$ as defined by Eq. (1). On the other hand, the operators $\hat{W}^k_q(j+\alpha)$ and $\hat{R}(j_1, j_1+\alpha_1, j_2, j_2+\alpha_2, j, k)$ identify with $W^k_q(j,\alpha)$ and $R(j_1\alpha_1 j_2\alpha_2 jk)$ as defined by Eqs. (79) and (95), respectively. Lastly, specialization of the general commutation relation to the chain $SU_2 \supset G^*$ leads to Eq. (30).

APPENDIX B: THE f SYMBOL

The f symbol (f for fractional) defined via Eq. (12) can be equivalently rewritten as

$$f\begin{pmatrix} j_1 & j_2 & k \\ \mu_1 & \mu_2 & \mu \end{pmatrix}$$
$$= (-1)^{2k}(2j_1+1)^{-1/2}\langle j_2 k\mu_2\mu|\, j_1\mu_1\rangle^*,$$

where

$$\langle j_2 k\mu_2\mu|j_1\mu_1\rangle = \sum_{m_2 q m_1} \langle j_2 m_2|j_2\mu_2\rangle^*\langle kq|k\mu\rangle^*$$
$$\times \langle j_2 k m_2 q|j_1 m_1\rangle\ \langle j_1 m_1|j_1\mu_1\rangle$$

is a $SU_2 \supset G^*$ symmetry adapted Clebsch–Gordan coefficient.[36] It constitutes an extension of the f coefficients relative to the cubical, tetragonal, and trigonal double groups used by Racah, Low, and some of their students[35] in crystal-field analysis for iron group ions in solids. The main interest of the f symbol lies in the concise form

$$\langle u_1 j_1\mu_1|T^k_\mu|u_2 j_2\mu_2\rangle$$
$$= (u_1 j_1\|T^k\|u_2 j_2) f\begin{pmatrix} j_1 & j_2 & k \\ \mu_1 & \mu_2 & \mu \end{pmatrix}$$

the Wigner–Eckart theorem assumes when transcribed in a $SU_2 \supset G^*$ basis.

The $1-j\mu$ symbol defined through Eq. (18) derives from the f symbol since

$$f\begin{pmatrix} 0 & j & j' \\ \Gamma_0 & \mu & \mu' \end{pmatrix} = \delta(j'j)(2j+1)^{-1/2}\begin{pmatrix} j & \\ \mu & \mu' \end{pmatrix}^*.$$

We may think of the f symbol as a third-rank tensor (second-rank covariant and first-rank contravariant). A third-rank covariant, and therefore more symmetrical, tensor arises with the introduction of the 3-$j\mu$ symbol

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1 & \mu_2 & \mu_3 \end{pmatrix} = \sum_{\mu_3'} \begin{pmatrix} j_3 & \\ \mu_3 & \mu_3' \end{pmatrix}^* f\begin{pmatrix} j_3 & j_2 & j_1 \\ \mu_3' & \mu_2 & \mu_1 \end{pmatrix}$$

which identifies to the Wigner 3-jm symbol when $G^* = U_1$.

By using the unitarity property of both the $SU_2 \supset U_1$ Wigner coefficients and the transformation from the $\{m\}$ scheme to the $\{\mu \equiv a\Gamma\gamma\}$ scheme, the f (or 3-$j\mu$) symbol may be seen to satisfy orthogonality relations[36] similar to the ones of the 3-jm symbol. In addition, it may be also fractioned[36] according to the Racah lemma sketched in Appendix A. Finally, the SU_2 recoupling coefficients can be redefined with the help of the f (or 3-$j\mu$) symbol. In this regard, we have

$$\sum_{\mu_3} f\begin{pmatrix} j_1 & j_3 & k_2 \\ \mu_1 & \mu_3 & \mu_2' \end{pmatrix} f\begin{pmatrix} j_3 & j_2 & k_1 \\ \mu_3 & \mu_2 & \mu_1' \end{pmatrix}$$
$$= \sum_{k_3\mu_3'} (-1)^{j_1+j_2+k_1-k_2}(2k_3+1)\begin{Bmatrix} k_1 & k_2 & k_3 \\ j_1 & j_2 & j_3 \end{Bmatrix}$$
$$\times f\begin{pmatrix} k_3 & k_1 & k_2 \\ \mu_3' & \mu_1' & \mu_2' \end{pmatrix} f\begin{pmatrix} j_1 & j_2 & k_3 \\ \mu_1 & \mu_2 & \mu_3' \end{pmatrix},$$

a relation that is central in the derivation of Eq. (30).

[1] L.C. Biedenharn, Phys. Lett. **3**, 254 (1963); L.C. Biedenharn, J. Math. Phys. **4**, 436 (1963); G.E. Baird and L.C. Biedenharn, J. Math. Phys. **4**, 1449 (1963); *ibid.* **5**, 1723, 1730 (1964).

[2]L.C. Biedenharn and H. van Dam, *Quantum Theory of Angular Momentum* (Academic, New York, 1965).
[3]J.D. Louck and L.C. Biedenharn, J. Math. Phys. **11**, 2368 (1970). See also L.C. Biedenharn, J.D. Louck, E. Chacón, and M. Ciftan, J. Math. Phys. **13**, 1957 (1972); L.C. Biedenharn and J.D. Louck, J. Math. Phys. **13**, 1985 (1972); J.D. Louck and L.C. Biedenharn, J. Math. Phys. **14**, 1336 (1973).
[4]M.K.F. Wong, J. Math. Phys. **12**, 1530 (1971).
[5]J. Schwinger, Report U.S. Atom. Energy Comm. NYO-3071, 1952, published in Ref. 2. See also P. Jordan, Z. Phys. **94**, 531 (1935).
[6]G. Racah, Phys. Rev. **62**, 438 (1942). See also J.P. Elliott, Proc. R. Soc. London A **245**, 128 (1958); G. Racah and J. Stein, Phys. Rev. **156**, 58 (1967).
[7]G. Grenet and M. Kibler, J. Solid State Chem. **23**, 1 (1978).
[8]L.L. Hirst, Adv. Phys. **27**, 231 (1978).
[9]B.R. Judd, *Operator Techniques in Atomic Spectroscopy* (McGraw-Hill, New York, 1963).
[10]K.W.H. Stevens, Proc. Phys. Soc. **65**, 209, 311 (1952). See also R.J. Elliott and K.W.H. Stevens, Proc. R. Soc. A **218**, 553 (1953); *ibid.* **219**, 387 (1953); B.R. Judd, Proc. R. Soc. A **227**, 552 (1955); *ibid.* Mol. Phys. **2**, 407 (1959); R.J. Elliott, B.R. Judd, and W.A. Runciman, Proc. R. Soc. A **240**, 509 (1957); J.M. Baker, B. Bleaney, and W. Hayes, Proc. R. Soc. A **247**, 141 (1958); D.A. Jones, J.M. Baker, and D.F. D. Pope, Proc. Phys. Soc. **74**, 249 (1959); M.T. Hutchings, Sol. State Phys. **16**, 227 (1964); *ibid.* Proc. Phys. Soc. **85**, 397 (1965).
[11]H.A. Buckmaster, Can. J. Phys. **40**, 1670 (1962). See also D. Smith and J.H.M. Thornley, Proc. Phys. Soc. **89**, 779 (1966); R.J. Birgeneau, Can. J. Phys. **45**, 3761 (1967); P.-A. Lindgård, J. Phys. C **8**, 3401 (1975).
[12]H.A. Buckmaster, R. Chatterjee, and Y.H. Shing, Phys. Stat. Sol. (a) **13**, 9 (1972).
[13]W. Low, *Paramagnetic Resonance in Solids* (Academic, New York, 1960); A. Abragam and B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions* (Clarendon, Oxford, 1970).
[14]T. Holstein and H. Primakoff, Phys. Rev. **58**, 1098 (1940); F.J. Dyson, Phys. Rev. **102**, 1217 (1956); S.V. Maleev, Zh. Eksp. Teor. Fiz. **33**, 1010 (1957); *ibid.* Sov. Phys. -JETP **7**, 1048 (1958); G.T. Trammell, J. Appl. Phys. **31**, 362 S (1960); *ibid.* Phys. Rev. **131**, 932 (1963); B. Grover, Phys. Rev. A **140**, 1944 (1965).
[15]P.-A. Lindgård and O. Danielsen, J. Phys. C **7**, 1523 (1974); P.-A. Lindgård and A. Kowalska, J. Phys. C **9**, 2081 (1976); A. Kowalska and P.-A. Lindgård, Physica **86**, 1111 (1977).
[16]M.E. Rose, *Elementary Theory of Angular Momentum* (Wiley, New York, 1957).
[17]P.W. Atkins and J.C. Dobson, Proc. R. Soc. A **321**, 321 (1971); P.W. Atkins and P.A. Seymour, Molec. Phys. **25**, 113 (1973). See also W. Witschel, Molec. Phys. **26**, 419 (1973).
[18]B. Meyer, Can. J. Math. **6**, 135 (1954). See also J. Patera and P. Winternitz, J. Chem. Phys. **65**, 2725 (1976); L. Michel, in *Group Theoretical Methods in Physics*, edited by R.T. Sharp and B. Kolman (Academic, New York, 1977).
[19]P.L. Corio, J. Math. Phys. **9**, 1067 (1968). See also P.L. Corio, SIAM J. Appl. Math. **29**, 661 (1975).
[20]J.M. Caola, Phys. Lett. A **47**, 357 (1974).
[21]A.K. Bose, Phys. Lett. A **50**, 425 (1975); A.K. Bose and R.A.B. Devine, Phys. Lett. A **55**, 267 (1975).
[22]G. Grenet and M. Kibler, Phys. Lett. A **68**, 147 (1978); *ibid.* **71**, 323 (1979).
[23]M. Yamamura, T. Suzuki, and H. Ichihashi, Progr. Theor. Phys. (Kyoto) **60**, 197 (1978). See also H. Ichihashi and M. Yamamura, Progr. Theor. Phys. (Kyoto) **60**, 753 (1978).
[24]D.C. Mattis, *The Theory of Magnetism* (Harper and Row, New York, 1965).
[25]T. Kishimoto and T. Tamura, Nucl. Phys. A **192**, 246 (1972); *ibid.* **270**, 317 (1976); O. Scholten, F. Iachello, and A. Arima, Ann. Phys. (N.Y.) **115**, 325 (1978); T. Otsuka, A. Arima, and F. Iachello, Nucl. Phys. A **309**, 1 (1978); O. Castaños, E. Chacón, A. Frank, and M. Moshinsky, J. Math. Phys. **20**, 35 (1979).
[26]O. Hara and T. Gotō, Progr. Theor. Phys. (Kyoto) Suppl. **41**, 56 (1968).
[27]G. Györgyi and S. Kövesi-Domokos, Report IPNO/TH123 (Paris University, 1968).
[28]R.L. Anderson, S. Kumei, and C.E. Wulfman, J. Math. Phys. **14**, 1527 (1973).
[29]M.J. Englefield, *Group Theory and the Coulomb Problem* (Wiley–Interscience, New York, 1972). See also W. Witschel, J. Phys. A: Math. Nucl. Gen. **7**, 1847 (1974).
[30]A. Bohm and R.B. Teese, J. Math. Phys. **17**, 94 (1976).
[31]J. Patera, R.T. Sharp, P. Winternitz, and H. Zassenhaus, J. Math. Phys. **18**, 2259 (1977).
[32]M. Flato and C. Fronsdal, Lett. Math. Phys. **2**, 421 (1978).
[33]K. Schulten and R.G. Gordon, J. Math. Phys. **16**, 1961, 1971 (1975); A.O. Barut and R. Wilson, J. Math. Phys. **17**, 900 (1976); D.A. Akyeampong and M.A. Rashid, Report IC/75/145 (Miramare–Trieste, 1975); R. Shaw, Preprint (Hull University, 1975).
[34]G. Racah, Phys. Rev. **76**, 1352 (1949); Ergeb. Exakt. Naturw. **37**, 28 (1965).
[35]G. Schönfeld, M. Sc. thesis (Jerusalem University, 1960); M. Flato, M. Sc. thesis (Jerusalem University, 1961); G. Racah and M. Flato, unpublished; G. Rosengarten, M. Sc. thesis (Jerusalem University, 1961); W. Low, in *Quantum Electronics*, edited by C.H. Townes (Columbia University, New York, 1960); W. Low, in *Advances in Quantum Electronics*, edited by J.R. Singer (Columbia University, New York, 1961); W. Low and G. Rosengarten, J. Molec. Spectrosc. **12**, 319 (1964); M. Flato, J. Molec. Spectrosc. **17**, 300 (1965); W. Low, in *Spectroscopic and Group Theoretical Methods in Physics*, edited by F. Bloch, S.G. Cohen, A. de-Shalit, S. Sambursky, and I. Talmi (North–Holland, Amsterdam, 1968).
[36]M. Kibler, J. Molec. Spectrosc. **26**, 111 (1968); *ibid.* C.R. Acad. Sci. (Paris) B **268**, 1221 (1969); *ibid.* Int. J. Quant. Chem. **3**, 795 (1969); *ibid.* Report LYCEN 6946 (Lyon University, 1969); *ibid.* Report LYCEN 7843 (Lyon University, 1978); *ibid.* in *Recent Advances in Group Theory and Their Application to Spectroscopy*, edited by J.C. Donini (Plenum, New York, 1979).
[37]E.P. Wigner, 1940 manuscript, published in Ref. 2
[38]A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University, New Jersey, 1960).
[39]J. Killingbeck, J. Phys. A **8**, L121 (1975).
[40]P.-O. Löwdin, Rev. Mod. Phys. **36**, 966 (1964); J. Shapiro, J. Math. Phys. **6**, 1680 (1965); Yu.F. Smirnov, in *Clustering Phenomena in Nuclei* (IAEA, Vienna, 1969). See also J.-L. Calais, Tech. Note No. 25 (Uppsala University, 1959) and Int. J. Quantum Chem. **2**, 715 (1968).
[41]D.R. Pooler, J. Phys. A **11**, 1045 (1978).
[42]W. Witschel, J. Phys. A **7**, 1847 (1974).
[43]V. Bargmann, Rev. Mod. Phys. **34**, 829 (1962), reprinted in Ref. 2.
[44]S.D. Majumdar, Progr. Theor. Phys. (Kyoto) **20**, 798 (1958).
[45]K. Venkatesh, J. Math. Phys. **19**, 2060 (1978).
[46]B.L. van der Waerden, *Die Gruppenteoretische Methode in der Quantenmechanik* (Springer, Berlin, 1932).
[47]K. Zukauskas and E. Mauza, Vilniaus universiteto Moksło darbai **7**, 55 (1957). See also A.P. Yutsis, I.B. Levinson, and V.V. Vanagas, *Matematicheskii Apparat Teorii Momenta Kolichestva Dvizheniya* (Vilnius, U.S.S.R., 1960). Translated by A. Sen and R.N. Sen: *Mathematical Apparatus of the Theory of Angular Momentum* (Israel Program for Scientific Translations, Jerusalem, 1962).
[48]E.P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Vieweg, Braunschweig, 1931). Translated by J.J. Griffin, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959).
[49]G. Grenet and M. Kibler, unpublished.
[50]J. Patera, R.T. Sharp, and P. Winternitz, J. Math. Phys. **19**, 2362 (1978); P.E. Desmier and R.T. Sharp, J. Math. Phys. **20**, 74 (1979); J. Patera and R.T. Sharp, in *Recent Advances in Group Theory and Their Application to Spectroscopy*, edited by J.C. Donini (Plenum, New York, 1979); *ibid.* in *Proceedings of the VII International Colloquium on Group Theoretical Methods in Physics* (Springer, Berlin, 1979).
[51]P.A. Cox, Struct. Bonding **24**, 59 (1975); G. Grenet, Y. Jugnet, Tran Minh Duc, and M. Kibler, Chem. Phys. Lett. **62**, 125 (1979).
[52]J.K. Lang, Y. Baer, and P.A. Cox, Phys. Rev. Lett. **42**, 74 (1979).
[53]W.G. Harter, Phys. Rev. A **8**, 2819 (1973); W.G. Harter and C.W. Patterson, Phys. Rev. A **13**, 1067 (1975); C.W. Patterson and W.G. Harter, Phys. Rev. A **15**, 2372 (1977).
[54]N.B. Backhouse, J. Phys. A **12**, 21 (1979); M. Bednář and V.F. Šachl, J. Math. Phys. **20**, 367 (1979); P.D. Jarvis and H.S. Green, Preprint THEP 78/9-7 (Southampton University, 1978).
[55]E. Schrödinger, Proc. R. Irish Acad. A **46**, 9 (1940); L. Infeld and T.E. Hull, Rev. Mod. Phys. **23**, 21 (1951); W. Miller, Jr., *Lie Theory and Special Functions* (Academic, New York, 1968).
[56]J.W.B. Hughes and J. Yadegar, J. Math. Phys. **19**, 2068 (1978).
[57]L.C. Biedenharn and J.D. Louck, Ann. Phys. (N.Y.) **63**, 459 (1971); *ibid.* in *Group Theory and its Applications*, Vol. II, edited by E.M. Loebl (Academic, New York, 1971).

Projective covering group versus representation groups[a),b)]

José F. Cariñena
Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, Spain

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

(Received 30 March 1979; accepted for publication 2 November 1979)

The relation between two methods of finding the unitary projective representations of a connected Lie group G is studied. It is shown that the one of the projective covering group $\tilde{G}$ is singled out as the simplest one, but one could also get a connected and simply connected Lie splitting group of minimal dimension by using a similar procedure.

1. INTRODUCTION

The character of rays rather than vectors of the states in quantum physics was pointed out the first time by H. Weyl (at least to the best of our knowledge) in his pioneer book on group theory and quantum mechanics.[1] When symmetries of a quantum system are considered, one is faced with the problem of how to determine the projective representations of the symmetry group. This problem was solved by Schur[2] (for finite groups) long time before the birth of the quantum mechanics.

As Wigner has shown[3] only semiunitary (unitary/antiunitary) projective representations (hereafter SUPR's) are pertinent in quantum theory. The classification of "relativistic" quantum elementary systems after the irreducible UPR's of the Poincaré group was given by Wigner.[4] Bargmann[5] has shown how to reduce the problem of finding the unitary projective representations of a connected Lie group G to the one of determining the unitary representations of a family of groups G_σ (extensions of G by T).

The study of the SUPR's of a finite group has been carried out by Janssen,[6] and also by the authors in a very recent paper.[7] Essentially we have to determine an associated group G' and a epimorphism $p':G'\to G$ such that any SUPR of G can be lifted to a SUR of G'. Such a pair (G',p') is called a splitting group for G; they always exist, and moreover there is (at least) one of minimal order which is called a representation group for G. A more detailed study can be found in Ref. 7. We will follow the notations and conventions of that paper.

Bargmann's method has been also studied by Mackey[8] (locally compact groups) and Cattaneo[9] (Polish groups). The method of finding a representation group has been given by Cattaneo[10] (Polish groups) and also by one of us[11] (Lie groups) by using similar methods to those employed in Ref. 7, for finite groups.

For the case of connected Lie groups (where only UPR's have to be considered) another method is advisable such as we have pointed out in an earlier paper[12]: to every connected Lie group G there correspond a well determined (and very easy to find) connected and simply connected (c.s.c.) Lie group $\tilde{G}$ and a (continuous) epimorphism $\tilde{p}:\tilde{G}\to G$ such that each UPR of G can be lifted to an UR of $\tilde{G}$. This group $\tilde{G}$ has been called the "projective covering group of G", although the manifold $\tilde{G}$ being not a topological covering of G in the general case, this name is somewhat misleading and thus the use of quotation marks. A study of the usefulness of this associated group in classical mechanics can be found in Ref. 13.

It is worthy of remark that splitting and representation groups for G have to be looked for in a more general class than that of G. For instance, in a paper by Moore[14] one can find an example of a second countable locally compact (s.c.l.c.) group which does not have any s.c.l.c. splitting group. But the existence of the "projective covering group" $\tilde{G}$ for any connected Lie group shows that any connected Lie group has (at least) one c.s.c. splitting Lie group. This class of groups is in a one-to-one correspondence with the class of the Lie algebras and the handling of the latter is in the background of almost all theoretical physicists. Therefore it may be worthwhile to spend a little time in the comparison of the two methods of finding the UPR's of a connected Lie group G: that of projective covering group $\tilde{G}$ (which always exists) and the one of a "minimal" splitting group, the representation group (if it exists). This is the aim of this paper.

2. THE STRUCTURE OF THE REPRESENTATION GROUPS

This section is devoted to clarify the following question: If G is a s.c.l.c., does a s.c.l.c. representation group $\tilde{G}$ exist? The answer is given by Cattaneo in a recent paper,[10] in terms of the existence of a suitable topology over the group $H^2(G,T)$.

We give here a slightly different but equivalent formulation: There is such a group $\tilde{G}$ iff there exist an homomorphic section $s:H^2(G,T)\to Z^2(G,T)$ (Borel cohomology with respect to the trivial action of G on T) and a locally compact topology τ on $H^2(G,T)$ in such a way that the maps $W_s(g,h):H^2(G,T)\to T$, defined by $W_s(g,h)\bar{\omega}=s(\bar{\omega})(g,h)$ are continuous for every pair $(g,h)\in G\times G$. Moreover, if these conditions hold, the corresponding representation group can be described as the "right-hand part" of the extension of G by $[H^2(G,T)]^\wedge$ characterized by the factor system W_s: $G\times G\to[H^2(G,T)]^\wedge$ given by the former expression. The dual group $[H^2(G,T)]^\wedge$ is the one relative to the topology τ defined in $H^2(G,T)$ and it is endowed with the compact-open

[a)]This work has been supported by Instituto de Estudios Nucleares.
[b)]Partly from the Ph.D. Thesis quoted in Ref. 11.

 0022-2488/80/030440-04$1.00

topology of τ. The topology in the middle group is the only one making topological such extension. For any given section s, if there exists a topology making continuous all the maps $W_s(g,h)$, such topology is unique; similarly, if s_1 and s_2 are two different sections such that there exist topologies τ_1 and τ_2 making continuous the maps $W_s(g,h)$ and $W_{s_2}(g,h)$ respectively, then τ_1 and τ_2 coincide. This property may be easily deduced from Prop. 8, ii) in Ref. 10.

Therefore, if $\overline{G}$ is a s.c.l.c. representation group for G (with respect to the epimorphism $\bar{p}:\overline{G}\to G$), it is topologically isomorphic to a topological extension of G by a well-determined topological group $[H^2(G,T)]\hat{}$.

Let us suppose now that G is not only a s.c.l.c. group but also a real finite dimensional (r.f.d.) Lie group. When is the representation group $\overline{G}$ a r.f.d. Lie group? The answer is given by the following theorem:

Theorem 1:Let G be a r.f.d. Lie group. There exists a Lie representation group for G iff the group $H^2(G,T)$, endowed with the unique (if it exists) locally compact topology such that all the maps $W_s(g,h)$ are continuous (for some section s that is itself a Lie group).

Proof: If A is an Abelian Lie group, the dual group $\hat{A}$ (with the usual compact-open topology) is also a Lie group. If $(\overline{G},\bar{p})$ is a Lie representation group, then $[H^2(G,T)]\hat{} = \ker\bar{p}$ is canonically imbedded in $\overline{G}$ as a closed subgroup. Therefore, $[H^2(G,T)]\hat{}$ is a Lie group. This proves the necessity. The sufficiency follows because in all (Borel) topological extensions of a Lie group the middle group is a Lie group.[15]

It is worthy of remark that a Lie group (even if connected) does not necessarily have an associated Lie representation group. Fortunately this does not occur for the Lie groups usually arising in quantum physics. On the other hand a connected Lie group has always associated one splitting group, the "projective covering group", whose structure we study next.

3. THE STRUCTURE OF THE "PROJECTIVE COVERING GROUP"

From now on we shall restrict ourselves to r.f.d. connected Lie groups. For such a group G, there exists a connected and simply connected (c.s.c.) Lie splitting group $\widetilde{G}$, called "projective covering group of G." This was shown in an earlier paper[12] by the authors, and it can be constructed, via its Lie algebra, in a well-defined way. Let LG be the Lie algebra of G and let us denote by G^* the universal covering of G. The second cohomology group $H^2(G^*,T)$ is canonically identified with $H^2(LG;\mathbb{R})$ in the usual way. Let n be the dimension of the linear space $H^2(LG;\mathbb{R})$. With these notations the recipe is as follows: We take a basis $[\Xi_i]$ $(i=1,\cdots,n)$ in $H^2(LG;\mathbb{R})$ and we pick out a cocycle $\Xi_i\in Z^2(LG;\mathbb{R})$ in each class $[\Xi_i]$. Now, in the linear space $\mathbb{R}^n\oplus LG$ we define a Lie algebra structure by means of

$$[(a_1,a_2,\ldots,a_n;\alpha),(b_1,b_2,\ldots,b_n;\beta)]$$
$$=\{\Xi_1(\alpha,\beta),\Xi_2(\alpha,\beta),\ldots,\Xi_n(\alpha,\beta);[\alpha,\beta]\}.$$

The Lie algebra which has been obtained with the former recipe is independent (up to isomorphisms) of the choices of the basis in $H^2(LG;\mathbb{R})$ and the representatives in $Z^2(LG;\mathbb{R})$ of these basis elements. The group $\widetilde{G}$ is the (unique) connected and simply connected Lie group corresponding to this Lie algebra $L\widetilde{G}$. The epimorphism $\pi:L\widetilde{G}\to LG$ defined by $\pi(a_1,a_2,\ldots,a_n;\alpha)=\alpha$ induces a continuous epimorphism $p:\widetilde{G}\to G^*$ which is independent of all arbitrary choices which have been made before. If $p^*:G^*\to G$ is the canonical projection and $\tilde{p}=p^*\circ p$, then $(\tilde{p},\widetilde{G})$ is a c.s.c. Lie splitting group, the "projective covering group" of G. A more detailed study may be found in Ref. 12.

In order to compare the "projective covering group" $\widetilde{G}$ with a representation group $\overline{G}$, a description of $\widetilde{G}$ in terms of G would be advisable without any reference to the Lie algebra LG. This is the aim of the next paragraphs.

The kernel of p is the subgroup of $\widetilde{G}$ whose Lie algebra is the ideal $\{a_1,a_2,\ldots,a_n;0),\ a_i\in\mathbb{R}\}$ in $L\widetilde{G}$. Therefore $\ker p$ is topologically isomorphic with $\mathbb{R}^n$ because $\widetilde{G}$ is a c.s.c Lie group. Then $\widetilde{G}$ is the middle group of an extension $1\to\mathbb{R}^n\to\widetilde{G}\xrightarrow{p}G^*\to 1$.

On the other hand G^* is the middle group of an extension $1\to\pi_1(G)\xrightarrow{i^*}G^*\xrightarrow{p^*}G\to 1$, and the morphism $i^*:\pi_1(G)\to G^*$ leads to an exact commutative diagram.

$$\begin{array}{ccccccc} 1 & \to & \mathbb{R}^n & \to & E & \xrightarrow{p'} & \pi_1(G)\to 1 \\ & & \| & & \downarrow\phi & & \downarrow i^* \\ 1 & \to & \mathbb{R}^n & \to & \widetilde{G} & \xrightarrow{p} & G^*\to 1 \end{array}$$

where the upper row is the restriction to $\pi_1(G)$ of the lower one, that is to say, ϕ and i^* are the canonical injections of E into $\widetilde{G}$ and of $\pi_1(G)$ into G^*, respectively. Then $g\in E$ iff $p[\phi(g)]=i^*[p'(g)]$. Now, it is easy to show that $\ker\tilde{p}=\phi(E)$. In fact $\phi(E)\subseteq\ker\tilde{p}$ because of the relation $\tilde{p}\circ\phi=p^*\circ p\circ\phi=p^*\circ i^*p'$. Furthermore, if $g\in\ker\tilde{p}$ then $p(g)\in\ker p^*$ and therefore there exists $g'\in E$ such that $p(g)=i^*[p'(g')]=p\circ\phi(g')$, whence $\phi(g')$ and g may only differ in some element $r\in\mathbb{R}^n$, $g=r\phi(g')=\phi(rg')$, that is, $\ker\tilde{p}\subseteq\phi(E)$.

In the extension $1\to\mathbb{R}^n\to\ker\tilde{p}\to\pi_1(G)\to 1$ (with the identification $\phi:E\leftrightarrow\ker\tilde{p}$) the group $\pi_1(G)$ acts on $\mathbb{R}^n$ via G^*, that is, its action is trivial; furthermore this extension is inessential because of the divisibility of the Abelian group $\mathbb{R}^n$, and hence $E\simeq\ker\tilde{p}$ is isomorphic with the direct product $\mathbb{R}^n\otimes\pi_1(G)$.

This result we have found is the following one: The projective covering group $(\widetilde{G},\tilde{p})$ of G is "the right-hand part of" an extension of G by $\mathbb{R}^n\otimes\pi_1(G)$.

One would like to make clear what this extension is, that is to say, what the action of G on $\mathbb{R}^n\otimes\pi_1(G)$ and the factor system are. If so, one could construct $\widetilde{G}$ directly from G. However, this is not a useful procedure for building up the group $\widetilde{G}$, because of the simpler method of determining $\widetilde{G}$ through its Lie algebra, and we will not go deeper into this point.

4. REPRESENTATION GROUPS VERSUS PROJECTIVE COVERING GROUP

Now we are able to discuss the relation between the

projective covering group and the representation groups. The "projective group" $\widetilde{G}$ is at least a splitting group for G, but, is it a representation group? A very simple counterexample shows that it is not. If G is a compact and connected Lie group, $H^2(G,T) \simeq [t(\pi_1(G))]\hat{}$, where $t(A)$ denotes the torsion subgroup of the Abelian group A.[14] Moreover, the second cohomology group $H^2(G,T)$ is a finite group and the only topology endowing it with a s.c.l.c. group structure is the discrete topology. Therefore a representation group for such a group G is the middle group of an extension of G by $t(\pi_1(G))$. On the other hand if G^* is such that $H^2(G^*,T) = 0$ the projective covering group of G is just G^*. Hence, any compact and connected Lie group G satisfying $H^2(G,T) = 0$ and $\pi_1(G) \neq t(\pi_1(G))$ furnishes a negative answer to our question. The simplest example is the one used by Cattaneo,[10] $G = T$, because then $H^2(G^*,T) = 0$, $\pi_1(G) = Z$ and $t[\pi_1(G)] = 0$, and therefore $\overline{G} = T$ but $\widetilde{G} = \mathbb{R}$.

This example suggests a possible (but wrong) relation: If $\overline{G}$ is a Lie representation group for G, then its universal covering group $(\overline{G})^*$ is (topologically) isomorphic to $\widetilde{G}$. But this statement is also false as one can show by using $G = T \otimes T$. Then $\pi_1(T \otimes T) = Z \otimes Z$ and $H^2(G,T) = t(Z \otimes Z) = 0$ that is $\overline{G} = G$, while $G^* = \mathbb{R} \otimes \mathbb{R}$, and $H^2(\mathbb{R} \otimes \mathbb{R}, T) = \mathbb{R}$ shows that $\widetilde{G}$ is a three-dimensional Lie group which can not be topologically isomorphic to $G^* = \mathbb{R} \otimes \mathbb{R}$.

Another interesting example is the Galilei group G_2 in two space dimensions.[16] Then $H^2(G_2^*,T) = \mathbb{R}^3$ and $H^2(G_2,T) = \mathbb{R}^2$. The canonical topology in $H^2(G_2,T)$ can be seen to be convenient for our purposes and while $\overline{G}_2$ is some extension of G_2 by $\mathbb{R}^2$ (hence an eight-parameter group), the projective covering group $\widetilde{G}_2$ is an extension of G_2^* by $\mathbb{R}^3$ (hence a nine-parameter group) and the relation $(\overline{G})^* = \widetilde{G}$ does not hold.

However in most cases the conjecture $(\overline{G})^* = \widetilde{G}$ is true and one can ask what are the necessary and sufficient conditions for the conjecture to be true. The statement of such conditions is the main part of this paper. Before going into this question we recall that if $\tau: G' \to G$ is a Lie group homomorphism there is an induced homomorphism[17] $\Lambda_\tau: H^2(G,T) \to H^2(G',T)$, called inflation map. In particuar we will consider the canonical epimorphism $p^*: G^* \to G$. Moreover there is a bijection between $H^2(G^*,T)$ and $H^2(LG,\mathbb{R})$ because of G^* being a simply connected group.

Theorem 2: Let G be a connected real Lie group such that it has an almost connected Lie representation group $\overline{G}$. The projective covering group $\widetilde{G}$ of G is topologically isomorphic with $(\overline{G})^*$ iff the linear space spanned by $\Lambda_{p^*}(H^2(G,T))$ is all the linear space $H^2(LG;\mathbb{R})$.

Proof: The representation group $\overline{G}$ is connected even if $H^2(G,T)$ is not.[14] This result has been shown by Moore for compact connected groups (II, Chapter III, Prop. 3,3), but the proof given by Moore remains valid if $\overline{G}$ is an almost connected group (i.e., the factor group of $\overline{G}$ by the connected component of the identity is compact). Therefore we only have to show that $L\overline{G} = L\widetilde{G}$. The Abelian group $[H^2(G,T)]\hat{}$ will be isomorphic to $\mathbb{R}^r \otimes T^s \otimes D$ with r and s nonnegative integers and D an Abelian discrete group. Let m be $m = r + s$ and $n = \dim H^2(G^*,T)$. In order to prove the isomorphism $L\overline{G} \simeq L\widetilde{G}$, a simple analysis of the method of construction of $\widetilde{G}$ shows that it is sufficient to demonstrate: (i) the equality $m = n$ and, (ii) the set of the cohomology classes of the components of a cocycle Ξ defining the extension $1 \to \mathbb{R}^m \to L\overline{G} \to LG \to 1$ is a basis of $H^2(LG,\mathbb{R})$. Although the point (ii) can be seen as an immediate consequence of $\overline{G}$ being a representation group, we give an explicit proof because it will be relevant in some sense for the construction of a "minimal" simply connected splitting group such as indicated in the next section.

(i) The inequality $m \leqslant n$ arises in a natural way. One considers the exact restriction-inflation sequence $\widehat{G^*} \to [\pi_1(G)]\hat{} \to H^2(G,T) \xrightarrow{\inf} H^2(G^*,T)$ which is associated with the exact sequence[14] $1 \to \pi_1(G) \to G^* \to G \to 1$. Here inf denotes Λ_{p^*}. If the natural topologies are considered in each set of the sequence all the maps are continuous.

The Abelian group $\mathbb{R}^n$ has not proper compact Lie subgroups and the continuity of the inflation map $\mathbb{R}^n \otimes Z^s \otimes D \xrightarrow{\inf} \mathbb{R}^n$ implies that $D \subset \ker(\inf)$. The image of $[\pi_1(G)]\hat{}$ must be in the maximal tori of $H^2(G,T)$ and hence $\ker(\inf) = \widehat{D}$. But, if $\mathbb{R}^n$ contains a subgroup isomorphic to $\mathbb{R}^r \otimes Z^s$, then[18] $r + s = m \leqslant n$.

On the other hand by hypothesis the $\Lambda_{p^*}(H^2(G,T))$ (isomorphic to $\mathbb{R}^r \otimes Z^s$) spans $H^2(LG,\mathbb{R})$ and therefore $r + s = m = n$. It is worthy of remark that this last hypothesis is necessary; as an example if $G = T \otimes T$, $m = 0$ and $n = 1$, so, $(\overline{G})^* \neq \widetilde{G}$.

(ii) Let us suppose that $m = n$ but only $r < n$ among the (cohomology classes of) components of the cocycle Ξ are linearly independent. If $q = n - r$, the Lie algebra $L\overline{G}$ is the direct sum of the Abelian ideal $\mathbb{R}^q$ and another ideal which is a nontrivial extension of LG by $\mathbb{R}^r$. As a consequence, the group $\overline{G}$ is a direct product of a subgroup (isomorphic to) $\mathbb{R}^u \otimes T^v \otimes F$ (where $u,v \geqslant 0$, $u + v = q$, and $F \subset D$), and a subgroup $\overline{\overline{G}}$ which is a nontrivial extension of G by a subgroup B {the complementary of $\mathbb{R}^u \otimes T^v \otimes F$ in $[H^2(G,T)]$}. But in this case $\overline{\overline{G}}$ will also be a splitting group for G and therefore must be an homomorphism $\psi: \widehat{B} \to H^2(G,T)$; let A be the kernel of such homomorphism: The group $\overline{\overline{G}}$ is a splitting group and there exists an epimorphism $B \simeq H^2(G,T)/A \to H^2(G,T)$.

But the group $H^2(G,T)$ being isomorphic to $\mathbb{R}^r \otimes Z^s \otimes \widehat{D}$, the subgroup A must lie in $\widehat{D}$ in such a way that $\widehat{D}/A \simeq \widehat{D}$ (which is possible even if A is not trivial) and this is only possible when $r = n$.

Conversely, it is quite clear that if $L\overline{G}$ and $L\widetilde{G}$ are isomorphic, then $\Lambda_{p^*}(H^2(G,T))$ is a subgroup $\mathbb{R}^r \otimes Z^s$ in $\mathbb{R}^n$ such that $r + s = n$.

5. CONCLUDING REMARKS

The results we have obtained can be summed up as follows: If one intends to find the UPR's of a connected Lie group G by means of the unitary representations of a new Lie group G' demanding also "absolute minimality" (in the sense of Ref. 10) in the choice of such a group $\overline{G}$, it can occur that such a group $\overline{G}$ does not exist. By relaxing this condition of minimality, such a group G' always exists, even if we re-

strict G' to be a simply connected group. The "projective covering group" $\widetilde{G}$ is such a group which is singled out because of the simplicity of its construction through the corresponding Lie algebra structure. But, is $\widetilde{G}$ the "minimal" simply connected splitting group for G? Evidently $\widetilde{G}$ is not.

For the construction of such a minimal group G' via its algebra following the pattern of the construction of $\widetilde{G}$ already indicated a whole basis of $H^2(G^*,T)$ is not needed but only a basis of the subspace $\Lambda_p.(H^2(G,T))$. The new group which we would obtain will be a c.s.c. Lie group whose dimension is minimal among all such splitting groups; it will be essentially unique and it will coincide with the universal covering group of any representation group of G (cf. proof of Theorem 2). However, in order to determine such a group G' one has to calculate the second cohomology group $H^2(G,T)$ as well as $\Lambda_p.(H^2(G,T))$ and this may be a hard task. This is the reason why the use of $\widetilde{G}$ is recommended: The determination of $\widetilde{G}$ is straightforward.

[1]H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York 1950).
[2]J. Schur, J. reine u ang. Math. **127**, 20 (1904), **132**, 85 (1907).
[3]E.P. Wigner, *Group Theory* ... (Academic, New York, 1959).
[4]E.P. Wigner, Ann. Math. **40**, 149 (1939).
[5]V. Bargmann, Ann. Math. **59**, 1 (1954).
[6]T. Janssen, J. Math. Phys. **13**, 342 (1972).
[7]J.F. Cariñena and M. Santander, J. Math. Phys. **20**, 2168 (1979).
[8]G.W. Mackey, Acta Math. **99**, 265 (1958).
[9]U. Cattaneo, Rep. Math. Phys. **9**, 31 (1976).
[10]U. Cattaneo, J. Math. Phys. **19**, 452 (1978).
[11]M. Santander, *Linealización de representaciones proyectivas semiunitarias de grupos*, Ph.D. tesis, Universidad de Valladolid, 1977.
[12]J.F. Cariñena and M. Santander, J. Math. Phys. **16**, 1416 (1975).
[13]L. Martinez Alonso, J. Math. Phys. **17**, 1177 (1976), **20**, 219 (1979).
[14]C.C. Moore, Trans. Amer. Math. Soc. **113**, 40, **113**, 64 (1964).
[15]L. Calabi, Ann. Math. Pura Appl. **32**, 295 (1951).
[16]J.M. Lévy-Leblond, "Galilei group and Galilean invariance", in *Group Theory and Its Applications, Vol., II*, edited by E.M. Loebl (Acadmic, New York, 1971).
[17]U. Cattaneo and A. Janner, J. Math. Phys. **15**, 1155 (1974).
[18]N. Bourbaki, *General Topology* (Hermann, Paris, 1966).

On an infinitesimal approach to semisimple Lie groups and raising and lowering operators of O(n) and U(n)

M. D. Gould
Department of Mathematical Physics, University of Adelaide, G.P.O. Box 498, Adelaide, South Australia, 5001

(Received 7 February 1978; revised manuscript received 11 January 1979)

A purely algebraic approach to the evaluation of the fundamental Wigner coefficients and reduced matrix elements of O(n) and U(n) is given. The method employs the explicit use of projection operators which may be constructed using the polynomial identities satisfied by the infinitesimal generators of the group. As an application of this technique, a certain set of raising and lowering operators for O(n) and U(n) are constructed. They are simpler in appearance than those previously constructed since they may be written in a compact product form. They are, moreover, Hermitian conjugates of one another, and therefore are easily normalized.

1. INTRODUCTION

It has been shown[1-5] that the infinitesimal generators of semisimple Lie groups may be assembled into a matrix which satisfies a certain polynomial identity (herein called the characteristic identity) over the center of the universal enveloping algebra. On a representation admitting an infinitesimal character such an identity reduces to a polynomial identity over the underlying field (usually the real or complex field). In such a case the polynomial identity may be written in a convenient factored form.[2,3] In particular when acting on a finite dimensional irreducible representation of the group the polynomial identity reduces to the identities encountered recently by several authors for the various classical groups.[5-7]

The idea of assembling the infinitesimal generators of the classical groups into a matrix is not new and has proved in the past to be a useful technique.[8-10] By taking traces of powers of such a matrix one obtains a full set of invariants which serve to label the representations of the group completely, a fact which was recognized early.[4,10] The eigenvalues of such invariants (and invariants of a related nature) have been studied and computed by several authors.[11-13] Recently, however, a simple and systematic procedure for evaluating such invariants has been developed which employs the characteristic identities (see Ref. 11 and also compare with Ref. 13, Appendix B). In fact one may simply construct a full set of invariants for an arbitrary semisimple Lie group by the use of such methods. A general formula for the eigenvalues of these invariants, which applies to infinite as well as finite dimensional representations, has been developed in Ref. 3. This fact alone illustrates the potential importance of the characteristic identities in applications to group theory.

This paper is the first in a series which deals with the applications of the characteristic identities to the theory of groups. One of our principle aims is to show how one may evaluate the multiplicity free Wigner coefficients of a semisimple Lie group. It was noted early by Fano[14] that the characteristic identities (for the unitary groups of low order) were valuable for the explicit construction of projection operators. This idea was incorporated into subsequent work of Baird and Biedenharn[4] who noted that this algebraic technique would combine nicely with their evaluation of the fundamental Wigner coefficients for the general unitary groups. However, the idea was not considered further and it is our aim to pursue this matter in detail. The principle motivation for this paper, however, is that since the characteristic identities for arbitrary semisimple Lie groups are known the technique is generalizable to more general groups other than the relatively well known unitary groups.

The present paper deals solely with the orthogonal and unitary groups although possible extensions to more general groups are discussed in the concluding section. It is our principle aim to show how certain fundamental Wigner coefficients for O(n) and U(n) may be evaluated in a straightforward and simple manner by applying the use of projection operators which are constructed by means of the characteristic identities. At the same time we shall make an effort to relate our results to those obtained by Biedenharn, Louck, and Baird[15,16] who have given the evaluation of all multiplicity free Wigner coefficients for the unitary groups. Although our approach is intimately related to the approach employed by Baird and Biedenharn[16] there is one essential difference. The Wigner coefficients of the group are obtainable using only the properties of the projectors for which we have an explicit expression in terms of polynomials in the group generators. Calculations may then be carried out using only the Lie algebra commutation relations. Since the identities and associated projectors have been constructed explicitly for arbitrary semisimple Lie groups[17] this method is generalizable, in principle, to the general case.

As an application of this technique we shall construct a certain set of raising and lowering operators for U(n). These operators are different to those constructed by Nagel and Moshinsky[18] and may be written in a convenient product form. They are moreover Hermitian conjugates of one another which makes their normalization simple.

We shall also consider an extension of these results to the orthogonal subgroup of U(n). In particular the raising and lowering operators of the orthogonal groups are obtained which are different from those obtained by Wong[19] and Pang and Hecht.[20] Our approach to the orthogonal group in particular is considerably simpler than previous treatments and is no more difficult than the U(n) case. The raising and lowering operators for O(n) may also be written

 0022-2488/80/030444-10$1.00

in a compact product form and are moreover Hermitian conjugates of one another.

2. WIGNER COEFFICIENTS OF U(n)

The generators a_j^i of the Lie group U(n) satisfy the commutation relations

$$[a_j^i, a_l^k] = \delta_j^k a_l^i - \delta_l^i a_j^k$$

and the hermiticity property

$$(a_j^i)^\dagger = a_i^j.$$

These generators may be assembled into a square matrix a whose (i,j) entry is the generator a_j^i. Polynomials in a may then be defined recursively by the formula

$$(a^m)_j^i = (a^{m-1})_k^i a_j^k = a_k^i (a^{m-1})_j^k.$$

By a simple induction argument one may show that if $p(x)$ is any polynomial then the entries of the matrix $p(a)$ satisfy the Hermiticity property

$$[p(a)_j^i]^\dagger = p(a)_i^j. \tag{1}$$

It has been shown[2,3] that the matrix a satisfies a polynomial identity over the center of the enveloping algebra which may be written in its factorized form as

$$\prod_{r=1}^{n} (a - \alpha_r) = 0, \tag{2}$$

where the α_r are invariants of the group whose eigenvalues on a representation of U(n) with representation label $(\lambda_1,\ldots,\lambda_n)$ are given by

$$\alpha_r = \lambda_r + n - r.$$

Associated with the matrix a is its adjoint $\bar{a}$ defined by $\bar{a}_j^i = -a_j^i$. As for the matrix a one may define polynomials in the matrix $\bar{a}$ recursively by the formula

$$(\bar{a}^m)_j^i = (\bar{a}^{m-1})_j^k \bar{a}_k^i = \bar{a}_j^k (\bar{a}^{m-1})_k^i.$$

The adjoint matrix $\bar{a}$ satisfies the polynomial identity

$$\prod_{r=1}^{n} (\bar{a} - \bar{\alpha}_r) = 0, \tag{3}$$

where the roots $\bar{\alpha}_r$ are related to the α_r by

$$\bar{\alpha}_r = n - 1 - \alpha_r.$$

By virtue of the identities (2) and (3) projection operators $P[r]$ and $\bar{P}[r]$ may be constructed by setting

$$P[r] = \prod_{l \neq r} \left(\frac{a - \alpha_l}{\alpha_r - \alpha_l} \right), \quad \bar{P}[r] = \prod_{l \neq r} \left(\frac{\bar{a} - \bar{\alpha}_l}{\bar{\alpha}_r - \bar{\alpha}_l} \right).$$

Such projection operators are useful since they may be used to define rather general functions of the matrix a by setting

$$p(a) = \sum_{r=1}^{n} p(\alpha_r) P[r], \quad p(\bar{a}) = \sum_{r=1}^{n} p(\bar{\alpha}_r) \bar{P}[r]. \tag{4}$$

Suppose now that π^* denotes the fundamental contragredient vector representation of U(n) and let V^* be the representation space of π^*. Then on V^* the generators a_j^i are represented by elementary matrices $\pi^*(a_j^i) = -E_i^j$, where E_i^j, has 1 in the (j,i) position and zeros elsewhere. It follows then that the matrix a may be written

$$a = \sum_{i,j=1}^{n} E_j^i a_j^i = -\sum_{i,j=1}^{n} \pi^*(a_i^j) a_j^i.$$

In passing it is interesting to note that in the case of SU(2) where π^* is the spin representation one obtians the matrix originally considered by Dirac in connection with wave equations for particles of higher spin. This same matrix also occurs in the quantum mechanics of a spinning electron in three dimensions (Landés interval rule) and appears again as one of the ladder operators considered by Louck[21] and others.

Now let $V(\lambda)$ denote a finite dimensional irreducible module over U(n), with highest weight λ, and let π_λ be the representation afforded by $V(\lambda)$. On $V(\lambda)$ the entries of the matrix a are represented by endormorphisms $a_j^i \equiv \pi_\lambda(a_j^i)$. In such a case a may be written in the form

$$a = -\sum_{i,j=1}^{n} \pi^*(a_i^j) \pi_\lambda(a_j^i)$$

and hence a may be regarded as an operator on the product space $V^* \otimes V(\lambda)$.

Following Hannabuss[6] the matrix a may be written in invariant form, when acting on $V(\lambda)$, as

$$a = -\tfrac{1}{2}[\pi^* \otimes \pi_\lambda(I_2) - \pi^*(I_2) \otimes 1 - 1 \otimes \pi_\lambda(I_2)],$$

where I_2 is the universal Casimir element

$$I_2 = \sum_{i,j=1}^{n} a_j^i a_i^j.$$

To see this we note that a is an operator on the tensor product space $V^* \otimes V(\lambda)$. On V^* the generators are represented by elementary matrices $-E_i^j$ and on $V(\lambda)$ by $\pi_\lambda(a_j^i)$. On $V^* \otimes V(\lambda)$ the generators therefore are

$$\pi^* \otimes \pi_\lambda(a_j^i) = -E_i^j \otimes 1 + 1 \otimes \pi_\lambda(a_j^i),$$

so one may write

$$\pi^* \otimes \pi_\lambda(I_2) = \sum_{i,j=1}^{n} [-E_i^j \otimes 1 + 1 \otimes \pi_\lambda(a_j^i)] \times [-E_j^i \otimes 1 + 1 \otimes \pi_\lambda(a_i^j)].$$

Rearranging this expression we obtain

$$\sum_{i,j=1}^{n} E_i^j \pi_\lambda(a_i^j) = -\tfrac{1}{2}[\pi^* \otimes \pi_\lambda(I_2) - \pi^*(I_2) \otimes 1 - 1 \otimes \pi_\lambda(I_2)],$$

where the left-hand side is the matrix a of U(n) as required.

Decomposing the tensor product representation $V^* \otimes V(\lambda)$ we obtain the Clebsch–Gordon reduction

$$V^* \otimes V(\lambda) = \sum_{r=1}^{n} V(\lambda - \Delta_r),$$

where Δ_r is the weight with 1 in position r and zeros elsewhere. On each space $V(\lambda - \Delta_r)$ the operator a takes the constant value

$$-\tfrac{1}{2}[\pi_{\lambda - \Delta_r}(I_2) - \pi^*(I_2) - \pi_\lambda(I_2)] = \lambda_r + n - r,$$

which are the roots α_r of the characteristic identity. More generally if $p(x)$ is any polynomial then on the space $V(\lambda - \Delta_r)$, $p(a)$ takes the constant $p(\alpha_r)$. In particular the U(n) projector $P[r]$ takes the constant value 1 on the space $V(\lambda - \Delta_r)$ and zero on the remaining $V(\lambda - \Delta_k)$ $(k \neq r)$. Thus $P[r]$ projects $V^* \otimes V(\lambda)$ onto $V(\lambda - \Delta_r)$. We may express this in terms of Gel'fand patterns as

$$P[r]\left|\begin{matrix}\lambda-\Delta_s\\(\mu)\end{matrix}\right\rangle=\delta_{rs}\left|\begin{matrix}\lambda-\Delta_s\\(\mu)\end{matrix}\right\rangle, \tag{5}$$

where

$$\left|\begin{matrix}\lambda-\Delta_s\\(\mu)\end{matrix}\right\rangle$$

denotes a Gel'fand basis state for the space $V(\lambda-\Delta_s)$ and (μ) denotes a Gel'fand pattern for the subgroup $U(n-1)$.

Now let

$$\left|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle,\quad\left|\begin{matrix}\lambda\\(\nu')\end{matrix}\right\rangle$$

be two Gel'fand basis states in the space $V(\lambda)$. Then the matrix elements of the (i,j) entry of the projector $P[r]$ between these states are given by

$$\left\langle\begin{matrix}\lambda\\(\nu')\end{matrix}\right|P[r]^i_j\left|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle=\left\langle\begin{matrix}\lambda\ \overline{1\dot{0}}\\(\nu');\ i\end{matrix}\right|P[r]\left|\begin{matrix}\overline{1\dot{0}}\ \lambda\\j\ ;(\nu)\end{matrix}\right\rangle,$$

where

$$\left|\begin{matrix}\overline{1\dot{0}}\\i\end{matrix}\right\rangle$$

forms the usual basis for the contragredient vector representation and where

$$\left|\begin{matrix}\overline{1\dot{0}}\ \lambda\\j\ ;(\nu)\end{matrix}\right\rangle$$

denotes the product state

$$\left|\begin{matrix}\overline{1\dot{0}}\\j\end{matrix}\right\rangle\otimes\left|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle.$$

Introducing a complete set of states for the product representation $V^*\otimes V(\lambda)$ these matrix elements may in turn be written

$$\begin{aligned}\left\langle\begin{matrix}\lambda\\(\nu')\end{matrix}\right|P[r]^i_j\left|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle=\sum_{\substack{(\mu),(\mu')\\s,t}}&\left\langle\begin{matrix}\lambda\ \overline{1\dot{0}}\\(\nu');\ i\end{matrix}\right|\left.\begin{matrix}\lambda-\Delta_s\\(\mu')\end{matrix}\right\rangle\\&\times\left\langle\begin{matrix}\lambda-\Delta_s\\(\mu')\end{matrix}\right|P[r]\left|\begin{matrix}\lambda-\Delta_t\\(\mu)\end{matrix}\right\rangle\\&\times\left\langle\begin{matrix}\lambda-\Delta_t\\(\mu)\end{matrix}\right|\left.\begin{matrix}\overline{1\dot{0}}\ \lambda\\j\ ;(\nu)\end{matrix}\right\rangle,\end{aligned}$$

where the sum on (μ) and (μ') is over all patterns in the spaces $V(\lambda-\Delta_t)$ and $V(\lambda-\Delta_s)$, respectively. Using property (5) of the projector $P[r]$ the right hand side equals

$$\sum_{(\mu)}\left\langle\begin{matrix}\lambda\ \overline{1\dot{0}}\\(\nu');\ i\end{matrix}\right|\left.\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right\rangle\left\langle\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right|\left.\begin{matrix}\overline{1\dot{0}}\ \lambda\\j\ ;(\nu)\end{matrix}\right\rangle \tag{6}$$

The numbers

$$\left\langle\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right|\left.\begin{matrix}\overline{1\dot{0}}\ \lambda\\j\ ;(\nu)\end{matrix}\right\rangle$$

are of fundamental importance since they are Wigner coefficients.

Note that Eq. (6) implies, in view of Eq. (4), that the matrix elements of the group generators are given by

$$\begin{aligned}\left\langle\begin{matrix}\lambda\\(\nu')\end{matrix}\right|a^i_j\left|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle=\sum_{r=1}^{n}(\lambda_r+n-r)\sum_{(\mu)}&\left\langle\begin{matrix}\lambda\ \overline{1\dot{0}}\\(\nu');\ i\end{matrix}\right|\left.\begin{matrix}\lambda-\Delta_s\\(\mu)\end{matrix}\right\rangle\\&\times\left\langle\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right|\left.\begin{matrix}\overline{1\dot{0}}\ \lambda\\j\ ;(\nu)\end{matrix}\right\rangle.\end{aligned}$$

This equation is quite easily generalized to more general groups and indicates that in determining the matrix elements of the group generators only the associated Wigner coefficients are required.

In certain special cases the sum (6) reduces to a single term enabling an evaluation of certain Wigner coefficients by an independent evaluation of the left hand side. The case of primary interest to us is the matrix element of the (r,r) entry of the projector $P[r]$ between the maximal state $|\lambda\rangle$ of $V(\lambda)$. In this case one obtains

$$\langle\lambda|P[r]^r_r|\lambda\rangle=\sum_{(\mu)}\left|\left\langle\lambda;\begin{matrix}\overline{1\dot{0}}\\r\end{matrix}\right|\left.\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right\rangle\right|^2.$$

However, the state

$$\left|\lambda;\begin{matrix}\overline{1\dot{0}}\\r\end{matrix}\right\rangle$$

has weight $\lambda-\Delta_r$ which is the highest weight occurring in $V(\lambda-\Delta_r)$. Hence, the term

$$\left\langle\lambda;\begin{matrix}\overline{1\dot{0}}\\r\end{matrix}\right|\left.\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right\rangle$$

vanishes unless

$$\left|\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right\rangle$$

coincides with the unique maximal state $|\lambda-\Delta_r\rangle$ of $V(\lambda-\Delta_r)$. One therefore obtains

$$\langle\lambda|P[r]^r_r|\lambda\rangle=\left|\left\langle\lambda;\begin{matrix}\overline{1\dot{0}}\\r\end{matrix}\right|\left.\lambda-\Delta_r\right\rangle\right|^2. \tag{7}$$

Another case of interest, which we shall not treat here, is the case where $i=j=n$ in Eq. (6). In this case one obtains the important result

$$\left\langle\begin{matrix}\lambda\\(\nu')\end{matrix}\right|P[r]^n_n\left|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle=\delta_{(\nu')(\nu)}\left|\left\langle\begin{matrix}\lambda\ \overline{1\dot{0}}\\(\nu');\ n\end{matrix}\right|\left.\begin{matrix}\lambda-\Delta_r\\(\nu)\end{matrix}\right\rangle\right|^2.$$

The matrix elements of $P[r]^n_n$ are quite easily evaluated as demonstrated in Ref. 22. This then enables a complete determination of all fundamental Wigner coefficients and ultimately the matrix elements of the group generators. We shall illustrate this procedure in a forthcoming publication. It is also interesting to note that the technique is capable of generalization to infinite dimensions enabling a treatment of the noncompact groups.

The matrix element (7) is quite easily determined as demonstrated in Ref. 22, according to which we may write

$$\left|\left\langle\lambda;\begin{matrix}1\dot{0}\\r\end{matrix}\right|\left.\lambda-\Delta_r\right\rangle\right|^2=\prod_{l>r}\left(\frac{\lambda_r-\lambda_l+l-r-1}{\lambda_r-\lambda_l+l-r}\right). \tag{8}$$

Now if $|\lambda\rangle$ is the maximal state in $V(\lambda)$ then one obtains

$$P[r]^i_j|\lambda\rangle=0,\quad\text{for } j>r \text{ and } i=1,\ldots,n. \tag{9}$$

To see this suppose

$$\left|\begin{matrix}\lambda\\(\nu')\end{matrix}\right\rangle$$

is an arbitrary state in $V(\lambda)$. Application of Eq. (6) immediately gives

$$\left\langle\begin{matrix}\lambda\\(\nu')\end{matrix}\right|P[r]^i_j\left|\lambda\right\rangle=\sum_{(\mu)}\left\langle\begin{matrix}\lambda\ \overline{1\dot{0}}\\(\nu');\ i\end{matrix}\right|\left.\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right\rangle\left\langle\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right|\left.\begin{matrix}\overline{1\dot{0}}\\j\ ;\lambda\end{matrix}\right\rangle. \tag{10}$$

However, if $j > r$ then the state

$$\left|\begin{matrix}1\dot{0}\\ j\end{matrix};\lambda\right\rangle$$

has weight $\lambda - \Delta_j$, which is greater than the highest weight occurring in $V(\lambda - \Delta_r)$. It follows immediately that the sum (10) vanishes. Since

$$\left|\begin{matrix}\lambda\\ (\nu)\end{matrix}\right\rangle$$

was chosen arbitrarily one obtains Eq. (9).

By applying similar considerations to the matrix $\bar{a}$ one obtains the equation

$$\left\langle\begin{matrix}\lambda\\ (\nu')\end{matrix}\middle|\bar{P}[r]^i_j\middle|\begin{matrix}\lambda\\ (\nu)\end{matrix}\right\rangle = \sum_{(\mu)}\left\langle\begin{matrix}\lambda & 1\dot{0}\\ (\nu')\,; & j\end{matrix}\middle|\begin{matrix}\lambda+\Delta_r\\ (\mu)\end{matrix}\right\rangle\left\langle\begin{matrix}\lambda+\Delta_r\\ (\mu)\end{matrix}\middle|\begin{matrix}1\dot{0} & \lambda\\ j\,; & (\nu)\end{matrix}\right\rangle$$

where

$$\left|\begin{matrix}1\dot{0}\\ i\end{matrix}\right\rangle$$

forms the usual basis for the fundamental vector representation. One may then deduce the results

$$\langle\lambda\,|\bar{P}[r]^r_r|\lambda\rangle = \left|\left\langle\lambda;\begin{matrix}1\dot{0}\\ r\end{matrix}\middle|\lambda+\Delta_r\right\rangle\right|^2, \tag{11}$$

$$\bar{P}[r]^i_j|\lambda\rangle = 0, \quad \text{for } i<r, \quad j=1,\ldots,n.$$

According to Ref. 22 the Wigner coefficient (11) may be evaluated using the formula

$$\left|\left\langle\lambda;\begin{matrix}1\dot{0}\\ r\end{matrix}\middle|\lambda+\Delta_r\right\rangle\right|^2 = \prod_{l<r}\left(\frac{\lambda_r-\lambda_l+l-r+1}{\lambda_r-\lambda_l+l-r}\right). \tag{12}$$

3. REDUCED MATRIX ELEMENTS OF U(n)

Recall that a U(n) vector operator ψ is defined as a collection of components $\psi^i (i = 1,\ldots n)$ which satisfy

$$[a^i_j,\psi^k] = \delta^k_j\psi^i. \tag{13}$$

By taking the Hermitian conjugate of this relation one obtains the transformation law of contragredient vector operators

$$[a^i_j,\psi^\dagger_k] = -\delta^i_k\psi^\dagger_j. \tag{14}$$

It is known[5,16] that a vector operator ψ may be resolved into a sum of shift vectors $\psi[r]$:

$$\psi = \sum_{r=1}^{n}\psi[r],$$

where $\psi[r]$ increases the eigenvalue of the representation label λ_r by one unit leaving the other λ_k unchanged;

$$\lambda_k\psi[r] = \psi[r](\lambda_k + \delta_{kr}).$$

Similarly, a contragredient vector operator $\psi^\dagger$ may be resolved into shift components $\psi^\dagger[r] = (\psi[r])^\dagger$ which decrease the representation labels by one unit:

$$\lambda_k\psi^\dagger[r] = \psi^\dagger[r](\lambda_k - \delta_{kr}).$$

Such shift vectors may be constructed by application of the projectors $P[r]$ and $\bar{P}[r]$ as follows:

$$\psi[r] = P[r]\psi = \psi\bar{P}[r], \quad \psi^\dagger[r] = \psi^\dagger P[r] = \bar{P}[r]\psi^\dagger. \tag{15}$$

One may show (see Ref. 22 for details) that the following relations hold:

$$\psi[r]\psi^\dagger[r] = \bar{m}_r P[r], \quad \psi^\dagger[r]\psi[r] = m_r\bar{P}[r]. \tag{16}$$

The invariants m_r and $\bar{m}_r$ are of particular interest since their eigenvalues on finite-dimensional irreducible representations are the squares of the reduced matrix elements of ψ and $\psi^\dagger$, respectively. Equation (16) may then be regarded as an operator generalization of the Wigner–Eckart theorem.

The operator $\tilde{\psi}[r] = (\bar{m}_r)^{-1/2}\psi[r]$ is therefore a fundamental Wigner operator which have been treated in detail by Biedenharn, Giovannini, Louck, and Baird.[4,15,23] Equation (16) may then be rewritten in the form

$$\tilde{\psi}[r]\tilde{\psi}^\dagger[r] = P[r],$$

an observation previously made by Louck and Biedenharn.[13] The important thing from our point of view is that we have an explicit expression for $P[r]$ as a polynomial in the matrix a. Equation (16) was arrived at using only purely algebraic techniques.

Note that by taking the trace of Eq. (16) one obtains the relations

$$\bar{m}_r = \frac{\psi[i]^i\psi^\dagger[r]_i}{t_r(P[r])}, \quad m_r = \frac{\psi^\dagger[r]_i\psi[r]^i}{t_r(\bar{P}[r])}, \tag{17}$$

which enables a systematic method for determining the reduced matrix elements. (The traces of the projectors $P[r]$ and $\bar{P}[r]$ have been evaluated in Refs. 11 and 22.)

We now note that if $|\lambda\rangle$ is a maximal weight state of weight λ then one obtains, in view of Eqs. (9) and (11),

$$\psi^\dagger[r]_i|\lambda\rangle = \psi^\dagger[r]_jP[r]^j_i|\lambda\rangle = 0, \quad \text{for } i>r,$$

$$\psi[r]^i|\lambda\rangle = \psi^j\bar{P}[r]^i_j|\lambda\rangle = 0, \quad \text{for } i<r.$$

From this it follows that $\psi[r]^r|\lambda\rangle$ and $\psi^\dagger[r]_r|\lambda\rangle$ are maximal weight states of U(n) of weight $(\lambda + \Delta_r)$ and $(\lambda - \Delta_r)$, respectively. One may secure the normalization of these states from the equations

$$\psi[r]^r\psi^\dagger[r]_r = \bar{m}_rP[r]^r_r, \quad \psi^\dagger[r]_r\psi[r]^r = m_r\bar{P}[r]^r_r \tag{18}$$

which may be evaluated using Eqs. (8) and (12).

Suppose now we look at the subgroup embedding U(n)$\subset$U($n+1$). The generators of U($n+1$) may also be assembled into a matrix $\hat{a}$ as we did for U(n). This matrix satisfies an $(n+1)$ degree polynomial identity analogous to Eq. (2). We denote the U($n+1$) characteristic identity by

$$\prod_{r=1}^{n+1}(\hat{a}-\beta_r) = 0,$$

where the β_r are invariants of the group which take constant values $\beta_r = \lambda_r + n + 1 - r$ on a representation of U($n+1$) with highest weight $(\lambda_1,\ldots,\lambda_{n+1})$.

In our previous notation let ψ denote the U(n) vector operator with components $\psi^i = a^i_{n+1}$ and let $\psi^\dagger_i = a^{n+1}_i$ $(i = 1,\ldots n)$ denote its contragredient. In this case the reduced matrix elements of ψ and $\psi^\dagger$ may be evaluated as a function of the β_k and α_r by applying formula (17) (see Ref. 22 for details). One obtains the formulas

$$m_r = (-1)^n\prod_{p=1}^{n+1}(\beta_p-\alpha_r-1)\prod_{\substack{l=1\\ \neq r}}^{n}(\alpha_r-\alpha_l+1)^{-1}, \tag{19}$$

$$\bar{m}_r = (-1)^n \prod_{p=1}^{n+1} (\beta_p - \alpha_r) \prod_{\substack{l=1 \\ \neq r}}^{n} (\alpha_r - \alpha_l - 1)^{-1}.$$

The raising and lowering operators $\psi[r]^r$ and $\psi^\dagger[r]_r$ will shift between maximal weight states of U(n) in a given irreducible representation of U($n+1$). The normalization of these operators may be obtained directly from Eq. (18) since the quantities $m_r \bar{m}_r P[r]^r_r$ and $\bar{P}[r]^r_r$ may all be evaluated using Eqs. (8), (12), and (19).

4. RAISING AND LOWERING OPERATORS OF U(n)

The U(n) generators a^i_j, where i and j are restricted to values 1,...,m (for some positive integer m less than n), form the generators of the unitary subgroup U(m) of U(n). We see therefore that U(n) admits the canonical chain of subgroups

$$\mathrm{U}(n) \supset \mathrm{U}(n-1) \supset \cdots \supset \mathrm{U}(1). \tag{20}$$

The irreducible representations of the groups U(m), $1 \leqslant m \leqslant n$, may be characterized by partitions $(\lambda_{1m}, \lambda_{2m}, \ldots, \lambda_{mm})$ where the λ_{im} are integers satisfying

$$\lambda_{1m} \geqslant \lambda_{2m} \geqslant \cdots \geqslant \lambda_{mm} \geqslant 0.$$

The partitions of two groups U($m+1$) and U(m) in the chain (20) are related by the inequalities

$$\lambda_{1\,m+1} \geqslant \lambda_{1m} \geqslant \lambda_{2\,m+1} \geqslant \lambda_{2m} \geqslant \cdots \geqslant \lambda_{mm} \geqslant \lambda_{m+1\,m+1}.$$

The set of partitions for the chain (20) is most conveniently arranged into a Gel'fand pattern which has been described by Gel'fand and Zetlin[24] and appears in the paper by Nagel and Moshinsky.[18]

If U($m+1$) is a subgroup in the chain (20) we shall denote, for convenience, a maximal weight vector of U(m) [i.e., a semimaximal state of U($m+1$)] simply by the pattern

$$\left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle.$$

This pattern denotes a maximal weight state of U(m) of weight (λ_{im}) contained in an irreducible representation of U($m+1$) with highest weight (λ_{im+1}).

When acting on the above state the shift components $\psi_m[r]$ of the U(m) vector operator

$$\psi^i_m = a^i_{m+1}$$

are given by [see Eq. (15)]

$$\psi_m[r]^i = a^j{}_{m+1} \prod_{\substack{l=1 \\ \neq r}}^{m} \left[\frac{\bar{a} + \lambda_{lm} - l + 1}{\lambda_{lm} - \lambda_{rm} + r - l} \right]^i_j,$$

where $\bar{a}$ is the U(m) adjoint matrix. Similarly, the shift components of the U(m) contragredient vector operator $\psi^\dagger_m$ are given by

$$\psi^\dagger_m[r]_i = a^{m+1}_j \prod_{\substack{l=1 \\ \neq r}}^{m} \left[\frac{a - \lambda_{lm} - m + l}{\lambda_{rm} - \lambda_{lm} + l - r} \right]^j_i.$$

For convenience we denote the rth component of $\psi_m[r]$ by ψ^r_m, viz., $\psi^r_m = \psi_m[r]^r$. In view of our previous remarks we have

$$\psi^r_m \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle \propto \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} + \delta_{ir} \end{matrix} \right\rangle, \psi^{\dagger r}_m \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle \propto \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} - \delta_{ir} \end{matrix} \right\rangle.$$

We therfore have our required raising and lowering operators and it just remains to obtain the normalization. Using the hermiticity relation

$$\left(\psi^{\dagger r}_m \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle \right)^\dagger = \left\langle \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right| \psi^r_m$$

the normalization constants $N^r_m(\lambda_{im+1}, \lambda_{im})$ of the lowering operators $\psi^{\dagger r}_m$ are given by

$$N^r_m = \left\langle \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right| \psi^r_m \psi^{\dagger r}_m \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle^{1/2},$$

which, using Eq. (18), may in turn be written

$$N^r_m = \left\langle \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right| \bar{m}_r P[r]^r_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle^{1/2},$$

where $P[r]$ is the U(m) projector

$$P[r] = \prod_{\substack{l=1 \\ \neq r}}^{m} \left(\frac{a - \lambda_{lm} - m + l}{\lambda_{rm} - \lambda_{lm} + l - r} \right).$$

Substituting formulas (8) and (19) into the above expression, noting that the operators β_k take constant values $\lambda_{km+1} + m + 1 - k$, while the α_r take constant values $\lambda_{rm} + n - r$, one obtains the formula

$$\begin{aligned} N^r_m = \Bigg[& (-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p + 1) \\ & \times \prod_{l<r} (\lambda_{rm} - \lambda_{lm} + l - r - 1)^{-1} \prod_{l>r} (\lambda_{rm} \\ & - \lambda_{lm} + l - r)^{-1} \Bigg]^{1/2} \end{aligned} \tag{21}$$

Our normalized lowering operators therefore are

$$(N^r_m)^{-1} \psi^{\dagger r}_m.$$

Proceeding in a similar fashion the normalization constants $\bar{N}^r_m$ of the raising operators ψ^r_m are given by

$$\bar{N}^r_m = \left\langle \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right| m_r \bar{P}[r]^r_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle^{1/2},$$

which may be evaluated using Eqs. (12) and (19). One thereby obtains

$$\begin{aligned} \bar{N}^r_m = \Bigg[& (-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p) \\ & \times \prod_{l>r} (\lambda_{rm} - \lambda_{lm} + l - r + 1)^{-1} \prod_{l<r} (\lambda_{rm} \\ & - \lambda_{lm} + l - r)^{-1} \Bigg]^{1/2}. \end{aligned}$$

We may now write down an arbitrary Gel'fand basis state in terms of lowering operators acting on the maximal weight state $|\mathrm{max}\rangle$ of U(n). We have

$$\left| \begin{matrix} \lambda_{1n} & & \lambda_{2n} & \cdots\cdots & \lambda_{nn} \\ & \lambda_{1n-1} & & \cdots & \lambda_{n-1\,n-1} \\ & & \cdots & & \\ & & \lambda_{12} & & \lambda_{22} \\ & & & \lambda_{11} & \end{matrix} \right\rangle = N[\lambda]^{-1} \prod_{m=1}^{n-1} \prod_{r=1}^{m} (\psi^{\dagger r}_m)^{\lambda_{rm+1} - \lambda_{rm}} |\mathrm{max}\rangle. \tag{22}$$

The normalization constant $N[\lambda]$ appearing in this expression is easily computed by repeated application of Eq. (21). We readily obtain

$$N[\lambda]=\Bigg(\prod_{m=1}^{n-1}\prod_{r=1}^{m}\prod_{l<r}\frac{(\lambda_{rm+1}-\lambda_{lm}+l-r-1)!}{(\lambda_{rm}-\lambda_{lm}+l-r-1)!}\times\prod_{l>r}\frac{(\lambda_{rm}-\lambda_{lm+1}+l-r)!}{(\lambda_{rm+1}-\lambda_{lm+1}+l-r)!}\times\prod_{p=1}^{m+1}\frac{(\lambda_{pm+1}-\lambda_{rm}+r-p+1)!}{(\lambda_{pm+1}-\lambda_{rm+1}+r-p+1)!}\Bigg)^{1/2}.$$

It should be noted that the products of lowering operators appearing in Eq. (22) are ordered in such a way that the lowering operators for the group U(m) appear on the right of the lowering operators for the group U($m-1$). The lowering operators $\psi_m^{\dagger r}$ and $\psi_m^{\dagger k}$ for U(m) are ordered so that $\psi_m^{\dagger r}$ appears on the right of $\psi_m^{\dagger k}$ when $r<k$. However, changing the order of two lowering operators for U(m) will only result in a change of normalization constant $N[\lambda]$.

5. EXTENSION TO O(n)

Without loss of generality we may take the generators of the orthogonal subgroup of U(n) to be

$$\alpha_j^i=a_j^i-a^j{}_i, \tag{23}$$

where the a_j^i are the generators of U(n). This corresponds to the choice of O(n) metric $g_{ij}=\delta_{ij}$. The generators (23) satisfy the commutation relations

$$[\alpha^i{}_j,\alpha_l^k]=\delta_j^k\alpha_l^i-\delta_l^i\alpha_j^k-\delta_i^k\alpha^j{}_l+\delta^j{}_l\alpha_i^k \tag{24}$$

and the hermiticity requirement

$$(\alpha_j^i)^\dagger=\alpha^j{}_i.$$

The representations of O(n) may be labeled by the maximum eigenvalues of the operators

$$-i\alpha_{2r}^{2r-1},\quad r=1,...,h,$$

where

$$h=\left[\frac{n}{2}\right]=\begin{cases}\frac{1}{2}n & n \text{ even},\\ \frac{1}{2}(n-1) & n \text{ odd}.\end{cases}$$

As for U(n) one may consider the O(n) matrix α, whose (i,j) entry is the generator α_j^i, and its adjoint $\bar{\alpha}$ with entries $\bar{\alpha}^i{}_j=-\alpha_j^i$. The matrices α and $\bar{\alpha}$ satisfy polynomial identities of the form

$$\prod_{r=1}^{n}(\alpha-\alpha_r)=0,\quad \prod_{r=1}^{n}(\bar{\alpha}-\bar{\alpha}_r)=0,$$

where the α_r are invariants of the group whose eigenvalues on an irreducible representation with highest weight $(\lambda_1,...,\lambda_h)$ are given by

$$\alpha_r=\lambda_r+n-1-r,\quad \alpha_{n+1-r}=r-1-\lambda_r,\quad r=1,...,h,$$

with

$$\alpha_{h+1}=h,\quad \text{for } n=2h+1.$$

Following the notation of Green[5] we may define labels λ_r for $r>h$ by setting

$$\lambda_{n+1-r}=1-\lambda_r,\quad r=1,...h, \tag{25}$$

with

$$\lambda_{h+1}=1,\quad \text{for } n=2h+1.$$

The roots $\bar{\alpha}_r$ appearing in the adjoint identity are related to the roots α_r by $\bar{\alpha}_r=\alpha_{n+1-r}$.

Unlike the U(n) case the O(n) generators defined by Eq. (23) are not in Cartan form. However, it is easily checked,[14] that the matrix a defined by

$$a_j^i=(M^{-1})^i_p\alpha_q^pM^q_j, \tag{26}$$

where M is the numerical unitary matrix with entries

$$M_j^{2j-1}=\frac{1}{\sqrt{2}}=M^{2j-1}_{n+1-j},$$

$$M_j^{2j}=-\frac{i}{\sqrt{2}}=-M^{2j}_{n+1-j},\quad j=1,...,h,$$

all other entries being zero except when $n=2h+1$, where we have an additional nonzero entry $M^n_{h+1}=1$, which has entries consisting of the O(n) generators in their root space forms.

These generators satisfy the commutation relations

$$[a_j^i,a_l^k]=\delta_j^ka_l^i-\delta_l^ia_j^k-\delta^k_{n+1-i}a_l^{n+1-j}+\delta_l^{n+1-j}a^k_{n+1-i}. \tag{27}$$

The diagonal entries of the matrix a are given by

$$a_r^r=-a^{n+1-r}_{n+1-r}=-i\alpha_{2r}^{2r-1},\quad r=1,...h,$$

with

$$a^{h+1}_{h+1}=0,\quad \text{for } n=2h+1.$$

In view of the commutation relations (27) the entries a_j^i are in Cartan form with the positive roots above the diagonal and negative roots below the diagonal of the matrix a in analogy with U(n).

More generally, if $p(x)$ is any polynomial, then the matrices $p(a)$ and $p(\alpha)$ are related by

$$p(a)=M^{-1}p(\alpha)M. \tag{28}$$

Similarly, we have

$$p(\bar{a})=(\bar{M})^{-1}p(\bar{\alpha})\bar{M},$$

where

$$\bar{M}^p_q=M^p_q,\quad (\bar{M})^{-1p}_q=(M^{-1})^p_q.$$

From this we see that the matrices a and α satisfy the same characteristic identity.

As for U(n) one may construct a set of projectors

$$P[r]=\prod_{l\neq r}\left(\frac{a-\alpha_l}{\alpha_r-\alpha_l}\right),\quad \bar{P}[r]=\prod_{l\neq r}\left(\frac{\bar{a}-\bar{\alpha}_l}{\alpha_r-\bar{\alpha}_l}\right),$$

from which one may define arbitrary functions of a and $\bar{a}$ as in Eq. (4).

Now let π_λ denote a finite dimensional irreducible represenation of O(n) with highest weight λ, and let $V(\lambda)$ be the representation space of π_λ. As before the matrix elements of entries of the projectors $P[r]$ and $\bar{P}[r]$ are bilinear combinations of Wigner coefficients

$$\left\langle\begin{matrix}\lambda\\(\nu')\end{matrix}\middle|P[r]^i_j\middle|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle=\sum_{(\mu)}\left\langle\begin{matrix}\lambda & \overrightarrow{10}\\(\nu')' & i\end{matrix}\middle|\begin{matrix}\lambda-\Delta_r\\(\mu)\end{matrix}\right\rangle\left\langle\begin{matrix}\lambda-\Delta_r & \overline{10}\\(\mu) & j\end{matrix}\middle|\begin{matrix}\lambda\\(\nu)\end{matrix}\right\rangle \tag{29}$$

$$r=1,...,n$$

$$\left\langle \begin{matrix} \lambda \\ (\nu') \end{matrix} \middle| \bar{P}[r]^i_j \middle| \begin{matrix} \lambda \\ (\nu) \end{matrix} \right\rangle = \sum_{(\mu)} \left\langle \begin{matrix} \lambda & 1\dot{0} \\ (\nu') & j \end{matrix} \middle| \begin{matrix} \lambda - \Delta_r \\ (\mu) \end{matrix} \right\rangle \left\langle \begin{matrix} \lambda - \Delta_r \\ (\mu) \end{matrix} \middle| \begin{matrix} 1\dot{0} & \lambda \\ i & (\nu) \end{matrix} \right\rangle,$$

where

$$\left| \begin{matrix} 1\dot{0} \\ i \end{matrix} \right\rangle$$

constitutes a basis state of the fundamental vector representation of weight Δ_i where we define labels Δ_i for $i > h$ by $\Delta_i = -\Delta_{n+1-i}$. From Eq. (29) one may deduce the relations

$$\langle \lambda | P[r]^i_r | \lambda \rangle = \left| \left\langle \lambda; \overline{\begin{matrix} 1\dot{0} \\ r \end{matrix}} \middle| \lambda - \Delta_r \right\rangle \right|^2,$$

$$\langle \lambda | \bar{P}[r]^i_r | \lambda \rangle = \left| \left\langle \lambda; \begin{matrix} 1\dot{0} \\ r \end{matrix} \middle| \lambda + \Delta_r \right\rangle \right|^2,$$

which may be evaluated using the formulas (see Ref. 22)

$$\left| \left\langle \lambda; \overline{\begin{matrix} 1\dot{0} \\ r \end{matrix}} \middle| \lambda - \Delta_r \right\rangle \right|^2 = \prod_{l>r} \left(\frac{\lambda_r - \lambda_l - 1 + \delta_{l,h+1} - \delta_{l,n+1-r} + l - r}{\lambda_r - \lambda_l + l - r} \right), \quad n = 2h+1,$$

$$= \prod_{l>r} \left(\frac{\lambda_r - \lambda_l - 1 + \delta_{l,n+1-r} + l - r}{\lambda_r - \lambda_l + l - r} \right), \quad n = 2h, \tag{30}$$

$$\left| \left\langle \lambda; \begin{matrix} 1\dot{0} \\ r \end{matrix} \middle| \lambda + \Delta_r \right\rangle \right|^2 = \prod_{l<r} \left(\frac{\lambda_r - \lambda_l + l - r + 1 - 2\delta_{l,h+1} + \delta_{l,n+1-r}}{\lambda_r - \lambda_l + l - r} \right), \quad n = 2h+1,$$

$$= \prod_{l<r} \left(\frac{\lambda_r - \lambda_l + 1 + l - r - \delta_{l,n+1-r}}{\lambda_r - \lambda_l + l - r} \right), \quad n = 2h, \tag{31}$$

where we define labels λ_r for $r > h$ in accordance with Eq. (25).

Using Eq. (29) one may also deduce the relations

$$P[r]^i_j | \lambda \rangle = 0, \quad \text{for } j > r, \quad i = 1,...,n,$$

$$\bar{P}[r]^i_j | \lambda \rangle = 0, \quad \text{for } i < r, \quad j = 1,...,n. \tag{32}$$

6. REDUCED MATRIX ELEMENTS OF O(n)

With respect to the generators (23) we define an O(n) vector operator as an operator with n components $\tilde{\psi}^i$ which satisfy

$$[\alpha^i_j, \tilde{\psi}^k] = \delta^k_j \tilde{\psi}^i - \delta^k_i \tilde{\psi}^j. \tag{33}$$

By taking the Hermitian conjuate of this relation one obtains the transformation law of contragredient vector operators

$$[\alpha^i_j, \tilde{\psi}^\dagger_k] = -\delta^i_k \tilde{\psi}^\dagger_j + \delta^j_k \tilde{\psi}^\dagger_i. \tag{34}$$

By applying the change of basis matrix M the vector operator $\tilde{\psi}$ gets transformed to

$$\psi^i = (M^{-1})^i_j \tilde{\psi}^j.$$

The vector operator ψ therefore has components

$$\psi^j = \tilde{\psi}^{2j-1} + i\tilde{\psi}^{2j},$$

$$\psi^{n+1-j} = \tilde{\psi}^{2j-1} - i\tilde{\psi}^{2j}, \quad j = 1,...,h,$$

with

$$\psi^{h+1} = \tilde{\psi}^n, \quad \text{for } n = 2h+1,$$

which transform according to

$$[a^i_j, \psi^k] = \delta^k_j \psi^i - \delta^k_{n+1-i} \psi^{n+1-j}. \tag{35}$$

Similarly, the contragredient vector operator $\tilde{\psi}^\dagger$ gets transformed into a contragredient vector operator $\psi^\dagger$ with components

$$\psi^\dagger_i = \tilde{\psi}^\dagger_j M^j_i,$$

which transforms according to

$$[a^i_j, \psi^\dagger_k] = -\delta^i_k \psi^\dagger_j + \delta^{n+1-j}_k \psi^\dagger_{n+1-i}. \tag{36}$$

From now on we refer to a vector (contragredient vector) operator as an operator with n components transforming according to Eq. (35) [Eq. (36)].

The O(n) vector operator ψ may be resolved into a sum of shift vectors

$$\psi^\bullet = \sum_{r=1}^{n} \psi[r]$$

which alter the representation labels according to

$$\lambda_k \psi[r] = \psi[r](\lambda_k + \delta_{kr}),$$

$$\lambda_k \psi[n+1-r] = \psi[n+1-r](\lambda_k - \delta_{kr}), \quad r = 1,...,h,$$

with

$$\lambda_k \psi[h+1] = \psi[h+1]\lambda_k, \quad \text{for } n = 2h+1.$$

The shift components $\psi^\dagger[r]$ of $\psi^\dagger$ therefore alter the representation labels according to

$$\lambda_k \psi^\dagger[r] = \psi^\dagger[r](\lambda_k - \delta_{kr}),$$

$$\lambda_k \psi^\dagger[n+1-r] = \psi^\dagger[n+1-r](\lambda_k + \delta_{kr}), \quad r = 1,...,h,$$

with

$$\lambda_k \psi^\dagger[h+1] = \psi^\dagger[h+1]\lambda_k, \quad \text{for } n = 2h+1.$$

These shifts components may be constructed by application of the projectors $P[r]$ and $\bar{P}[r]$ as in Eq. (15).

One also obtains the relations [cf. Eq. (16)]

$$\psi[r]\psi^\dagger[r] = \bar{m}_r P[r], \quad \psi^\dagger[r]\psi[r] = m_r \bar{P}[r],$$

where $\bar{m}_r = m_{n+1-r}$ are the squares of the reduced matrix elements of the vector operator ψ. These reduced matrix elements may be evaluated using Eq. (17).

The generators α^i_j of O($n+1$) may also be assembled into a matrix $\hat{\alpha}$ as for O(n). The matrix $\hat{\alpha}$ satisfies a polynomial identity

$$\prod_{r=1} (\hat{\alpha} - \beta_r) = 0,$$

where the β_r take constant values on a finite dimensional irreducible representation of O($n+1$) with highest weight λ given by

$$\beta_r = \lambda_r + n - r, \quad r = 1,...,n+1,$$

where we define labels λ_r for $r > [(n+1)/2]$ as in Eq. (25).

In the special case where ψ is the vector operator $\psi^i = (M^{-1})^i_j \alpha^j_{n+1}$ (where we sum on j from 1 to n), with adjoint $\psi^\dagger_i = \alpha^{n+1}_j M^j_i$, the reduced matrix elements m_r and

$\bar{m}_r$ may be evaluated as a function of the β_k and α_r using Eq. (17). We readily obtain (see Ref. 22)

$$\bar{m}_r = m_{n+1-r} = (-1)^n \prod_{p=1}^{n+1} (\beta_p - \alpha_r) \times \prod_{l \neq r} (\alpha_r - \alpha_l - 1 - \delta_{l,n+1-r})^{-1}, \quad n = 2h,$$

$$\bar{m}_r = m_{n+1-r} = (-1)^n \prod_{p=1}^{n+1} (\beta_p - \alpha_r) \times \prod_{l \neq r} (\alpha_r - \alpha_l - l + \delta_{l,h+1} - \delta_{l,n+1-r})^{-1}, \quad n = 2h+1. \quad (37)$$

7. RAISING AND LOWERING OPERATORS OF O(n)

The orthogonal group admits the canonical chain of subgroups

$$\mathrm{O}(n) \supset \mathrm{O}(n-1) \supset \cdots \supset \mathrm{O}(2). \quad (38)$$

The irreducible representations of the groups O(m) ($m = 2,...,n$) are characteristized by partitions $(\lambda_{1m}, \lambda_{2m}, ..., \lambda_{hm})$ ($h = [m/2]$) which satisfy the inequalities

$$\lambda_{1,m} \geqslant \lambda_{2m} \geqslant \cdots \geqslant \lambda_{h-1,m} \geqslant |\lambda_{h,m}|, \quad m = 2h,$$
$$\lambda_{1,m} \geqslant \lambda_{2,m} \geqslant \cdots \geqslant \lambda_{h-1,m} \geqslant \lambda_{h,m} \geqslant 0., \quad m = 2h+1,$$

where the labels are simultaneously all integers or all half odd integers.

The partitions of two groups O($m+1$) and O(m) in the canonical chain (38) are related by inequalities

$$\lambda_{1,2p+1} \geqslant \lambda_{1,2p} \geqslant \lambda_{2,2p+1} \geqslant \lambda_{2,2p} \geqslant \cdots \geqslant \lambda_{p,2p+1} \geqslant \lambda_{p,2p} \geqslant -\lambda_{p,2p+1},$$
$$\lambda_{1,2p} \geqslant \lambda_{1,2p-1} \geqslant \lambda_{2,2p} \geqslant \lambda_{2,2p-1} \geqslant \cdots \geqslant \lambda_{p-1,2p-1} \geqslant |\lambda_{p2p}|.$$

The set of partitions for the chain (38) may be rearranged into a Gel'fand pattern which has been described by Gel'fand and Zetlin[24] and appears in the paper by Pang and Hecht.[20]

Following our U(n) notation we shall denote a maximal weight state of O(m) with representation label (λ_{im}) contained in a representation of O($m+1$) with representation label (λ_{im+1}) by

$$\left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle.$$

Let us also denote the components $\psi_m[r]^r$ of the O(m) vector $\psi^i_m = (M^{-1})^i_j \alpha^j_{m+1}$ simply by ψ^r_m.

In order to incorporate all possible shifts we need only consider the operators ψ^r_m and their adjoints $\psi^{\dagger r}_m$ for values of r in the range $r = 1,...,h$. Of course, in the case when $m = 2h+1$ is odd one may also consider the zero shift operator ψ^0_m and its adjoint $\psi^{\dagger 0}_m$ defined by $\psi^0_m = \psi_m[h+1]^{h+1}$. However, we do not require the zero shift operator for our purposes.

We then have

$$\psi^r_m \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle \propto \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} + \delta_{ir} \end{matrix} \right\rangle,$$
$$\psi^{\dagger r}_m \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle \propto \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} - \delta_{ir} \end{matrix} \right\rangle, \quad r = 1,...,h.$$

The normalization constants N^r_m of the lowering operators $\psi^{\dagger r}_m$ and the normalization constants $\bar{N}^r_m$ of the raising operators ψ^r_m are given by

$$N^r_m = \left\langle \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right| \bar{m}_r P[r]^r_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle^{1/2}, \quad (39)$$

$$\bar{N}^r_m = \left\langle \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right| m_r \bar{P}[r]^r_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle^{1/2}. \quad (40)$$

Due to differences in the normalization associated with O(m) for m odd and even we shall now consider each case separately.

O($m = 2h+1$)

In this case we have, in accordance with Eqs. (30) and (37),

$$\bar{m}_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle = (-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p + 1) \times \prod_{l \neq r} (\lambda_{rm} - \lambda_{lm} + l - r - 1 + \delta_{l,h+1} - \delta_{l,n+1-r})^{-1} \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle,$$

$$P[r]^r_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle = \prod_{l>r} \left(\frac{\lambda_{rm} - \lambda_{lm} - 1 + \delta_{l,h+1} - \delta_{l,n+1-r} + l - r}{\lambda_{rm} - \lambda_{lm} + l - r} \right) \times \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle,$$

which gives

$$N^r_m = \left[(-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p + 1) \times \prod_{l>r} (\lambda_{rm} - \lambda_{lm} + l - r)^{-1} \prod_{l<r} (\lambda_{rm} - \lambda_{lm} + l - r - 1)^{-1} \right]^{1/2}.$$

We also have, in accordance with Eqs. (31) and (37)

$$m_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle = (-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p) \times \prod_{l \neq r} (\lambda_{rm} - \lambda_{lm} + l - r + 1 - 2\,\delta_{l,h+1} + \delta_{l,n+1-r})^{-1} \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle,$$

$$\bar{P}[r]^r_r \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle = \prod_{l<r} \left(\frac{\lambda_{rm} - \lambda_{lm} + l - r + 1 - 2\delta_{l,h+1} + \delta_{l,n+1-r}}{\lambda_{rm} - \lambda_{lm} + l - r} \right) \times \left| \begin{matrix} \lambda_{im+1} \\ \lambda_{im} \end{matrix} \right\rangle,$$

where we have made use of the identities

$$\bar{\beta}_p - \bar{\alpha}_r = \alpha_r - \beta_p + 1,$$
$$\bar{\alpha}_r - \bar{\alpha}_l = \alpha_l - \alpha_r + \delta_{l,h+1}, \quad \text{for } r \neq h+1.$$

Substituting these expressions into Eq. (40) one obtains

$$\bar{N}^r_m = \Big[(-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p) \times \prod_{l<r} (\lambda_{rm} - \lambda_{lm} + l - r)^{-1} \prod_{l>r} (\lambda_{rm} - \lambda_{lm} + l - r + 1 - 2\delta_{l,h+1} + \delta_{l,n+1-r})^{-1}\Big]^{1/2}.$$

O(m = 2h)

As for the case m odd we may substitute the formulas (30), (31), and (37) into Eqs. (39) and (40) to give

$$N^r_m = \Big[(-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p + 1 - \delta_{p,h+1}) \times \prod_{l>r} (\lambda_{rm} - \lambda_{lm} + l - r)^{-1} \prod_{l<r} (\lambda_{rm} - \lambda_{lm} + l - r - 1)^{-1}\Big]^{1/2},$$

$$\bar{N}^r_m = \Big[(-1)^m \prod_{p=1}^{m+1} (\lambda_{pm+1} - \lambda_{rm} + r - p - \delta_{p,h+1}) \prod_{l<r} (\lambda_{rm} - \lambda_{lm} + l - r)^{-1} \prod_{l>r} (\lambda_{rm} - \lambda_{lm} + l - r + 1 + \delta_{l,n+1-r})^{-1}\Big]^{1/2},$$

where we have made use of the identities

$$\bar{\beta}_p - \bar{\alpha}_r = \alpha_r - \beta_p + 1 + \delta_{p,h+1}$$

and

$$\bar{\alpha}_r - \bar{\alpha}_l = \alpha_l - \alpha_r .$$

Our normalized lowering and raising operators are therefore given by $(N^r_m)^{-1}\psi^{\dagger r}_m$ and $(\bar{N}^r_m)^{-1}\psi^r_m$, respectively, for values of r in the range $r = 1,...,h$.

8. GENERALIZATIONS (see Refs. 3 and 17)

Let L be a semisimple Lie algebra with Cartan subalgebra H and let Φ be the set of roots of L relative to H. Let Δ denote a base of L and let Φ^+ denote the corresponding set of positive roots. Finally let Λ denote the set of integral weights and $\Lambda^+ \subset \Lambda$ the set of dominant integral weights.

Fix a basis $\{x_1,...,x_l\}$ $(l = \dim L)$ of L and let $\{x^1,...,x^l\}$ be the corresponding dual basis with respect to the Killing form on L. Let $V(\lambda)$, $\lambda \in \Lambda^+$, be a finite dimensional irreducible module over L with highest weight λ and let π_λ be the representation afforded by $V(\lambda)$. One may then consider the operator

$$A = -\frac{1}{2}\sum_{r=1}^{l} [\pi_\lambda(x^r) \otimes x_r + \pi_\lambda(x_r) \otimes x^r],$$

which may be regarded as a $d \times d$ matrix $[d = \dim V(\lambda)]$ with entries from L:

$$A_{ij} = -\frac{1}{2}\sum_{r=1}^{l} [\pi_\lambda(x^r)_{ij} x_r + \pi_\lambda(x_r)_{ij} x^r],$$

where $\pi_\lambda(x)$ is the matrix [with respect to some suitably chosen basis in $V(\lambda)$] representing the generator under the representation π_λ.

The matrix A is clearly a generalization of the matrices considered for U(n) and O(n). Upon setting π_λ to be the contragredient vector representation of U(n) and choosing the basis $a^i_j (i,j = 1,...,n)$ one obtains the matrix with entries (a^i_j).

If $V(\mu)$, $\mu \in \Lambda^+$, is a finite dimensional representation of L, with highest weight μ, then acting on $V(\mu)$ the matrix A may be written

$$A = -\frac{1}{2}\sum_{r=1}^{l} [\pi_\lambda(x^r) \otimes \pi_\mu(x_r) + \pi_\lambda(x_r) \otimes \pi_\mu(x^r)].$$

A may clearly be regarded as an operator on the product space $V(\lambda) \otimes V(\mu)$. If $\lambda_1,...,\lambda_k$ are the distinct weights occurring in $V(\lambda)$ then the matrix A, when acting on $V(\mu)$, satisfies the polynomial identity[3]

$$\prod_{i=1}^{k} [A - \tfrac{1}{2}(\lambda,\lambda+2\delta) - \tfrac{1}{2}(\lambda_i, 2(\mu+\delta) + \lambda_i)] = 0, \tag{41}$$

where $(\,,\,)$ is the inner product induced on the weights by the Killing form and δ is the half sum of the positive roots (see Humphreys[25]).

One may write the Clebsch–Gordon reduction of the product space $V(\lambda) \otimes V(\mu)$ in the symbolic form[3,25]

$$V(\lambda) \otimes V(\mu) = \bigoplus_{i=1}^{k} n(i) V(\mu + \lambda_i),$$

where $V(\mu + \lambda_i)$ is a finite dimensional irreducible representation, which admits the infinitesimal character $\chi_{\mu+\lambda_i}$ (in the notation of Humphreys[25]), which is unique if it exists. The multiplicities $n(i)$ may be obtained using Klimyk's formula[26] (see also Refs. 3).

Using the identity (41) one may construct the projection operators

$$P[i] = \prod_{\substack{j=1 \\ \neq i}}^{k} \left\{ \frac{A - \tfrac{1}{2}(\lambda,\lambda+2\delta) - \tfrac{1}{2}(\lambda_j, 2(\mu+\delta)+\lambda_j)}{\tfrac{1}{2}[(\lambda_i, 2(\mu+\delta)+\lambda_i) - \tfrac{1}{2}(\lambda_j, 2(\mu+\delta)+\lambda_j)]} \right\}$$

$$i = 1,...,k,$$

which project the space $V(\lambda) \otimes V(\mu)$ onto the subspace $V(\mu + \lambda_i)$. The matrix elements of the entries of the projector between basis states in the space $V(\mu)$ are therefore bilinear combinations of Clebsch–Gordan coefficients of the form

$$\left\langle \begin{matrix} \mu + \lambda_i \\ (\nu) \end{matrix} \middle| \begin{matrix} \lambda & \mu \\ (\rho) & (\tau) \end{matrix} \right\rangle, \tag{42}$$

where

$$\left| \begin{matrix} \lambda & \mu \\ (\rho) & (\tau) \end{matrix} \right\rangle,$$

denotes the product state

$$\left| \begin{matrix} \lambda \\ (\rho) \end{matrix} \right\rangle \otimes \left| \begin{matrix} \mu \\ (\tau) \end{matrix} \right\rangle$$

and where

$$\left| \begin{matrix} \lambda \\ (\rho) \end{matrix} \right\rangle$$

refers to a state in the representation $V(\lambda)$ where (ρ) denotes a set of labels used to distinguish the basis states. This then opens up the possibility of determining the Wigner coefficients (42) by exploiting the properties of the projectors $P[r]$. The nicest case occurs when the weights in $V(\lambda)$ all occur

with multiplicity one. Then the tensor product is multiplicity free, which is precisely the case considered in our treatment of O(n) and U(n).

In particular, one may consider more general Wigner coefficients for the unitary group by choosing our reference representation π_λ to be one of the tensor representations.

Returning to the general case, on unitary representations of the group, the generators behave under Hermitian conjugation like $(x^i)^\dagger = x_i$. In particular, if x_α is an element of the root space L_α, corresponding to the root $\alpha \in \Phi$, then one may deduce $(x_\alpha)^\dagger \in L_{-\alpha}$.

A tensor operator with highest weight $\lambda \in \Lambda^+$ is defined to be a collection of operators T_i, $i = 1,...,d$, which transform under commutation with elements of L like the representation π_λ:

$$[x,T_i] = \pi_\lambda(x)^j_i T_j .$$

One may project out the shift components $T[i]$ of the tensor T by applying the projector $P[i]$[15]:

$$T[i] = TP[i].$$

One may deduce the relation

$$T[i]T^\dagger[i] = M(i)P[i], \tag{43}$$

where $T^\dagger[i] = (T[i])^\dagger$ are the shift components of the contragredient tensor $T^\dagger$. The quantity $M(i)$ appearing in Eq. (43) is the reduced matrix element which may be evaluated using

$$M(i) = \frac{\sum_{j=1}^{d} T[i]_j T^\dagger[i]^j}{t_r(P[i])},$$

where $t_r(P[i])$ (the trace of the projector $P[i]$) may be evaluated from the formula[3]

$$t_r(P[i]) = c(i) \prod_{\alpha \in \Phi^+} \frac{(\mu + \delta + \lambda_i, \alpha)}{(\mu + \delta, \alpha)},$$

where $c(i)$ is the multiplicity of the weight λ_i in $V(\lambda)$.

Equation (43) is clearly a generalization of Eq. (16) and may be used in a similar way (at least for unit multiplicities). One may choose a basis for $V(\lambda)$ to be a weight basis (i.e., a basis of simultaneous eigenvectors of the Cartan subalgebra). Suppose that the basis is arranged so that the ith basis vector has weight λ_i. Then as before one may show that the operator $T[i]_j$ takes a maximal weight vector of L of weight μ into a maximal weight vector of weight $\mu + \lambda_i$. By this means we may construct generalization raising and lowering operators for more general groups which may be normalized using the equation

$$T[i]^j T^\dagger[i]_j = M(i)(P[i])^i_i .$$

In particular, such considerations are important in the labeling problems where a Lie algebra is embedded in a larger Lie algebra K which is separated from K by an irreducible tensor operator

$$K = L \oplus_\perp T$$

where

$$[L,T] \subseteq T, \quad [T,T] \subseteq L.$$

ACKNOWLEDGMENTS

The author would like to thank Professor H.S. Green who provided the foundations on which much of our work has been based. He would also like to thank Professor C.A. Hurst for his encouragement and for proof reading the manuscript. Finally, the author acknowledges the financial support of a Commonwealth Postgraduate Research Award.

[1] B. Kostant, J. Funct. Anal. **20**, 257 (1975).
[2] D.M. O'Brien, A. Cant, and A.L. Carey, Ann. Inst. Henri Poincaré, Section A, Physique Théorique, 405 (1977).
[3] M.D. Gould, "A General Character Formula for Semi-Simple Lie Algebras," Univ. Adelaide Preprint (1978).
[4] G.E. Baird and L.C. Biedenharn, J. Math. Phys. **12**, 1723 (1964).
[5] H.S. Green, J. Math. Phys. **12**, 2106 (1971); and A.J. Bracken and H.S. Green, J. Math. Phys. **12**, 2099 (1971).
[6] K.C. Hannabus, "Characteristic Equations for Semi-Simple Lie Groups," Preprint, Math. Inst. Oxford (1972).
[7] C.O. Nwachuku and M.A. Rashid, J. Phys. A **11**, 457 (1978); J.D. Louck, Am. J. Phys. **31**, 378 (1963); N. Makunda, J. Math. Phys. **8**, 1069 (1967); J.D. Louck and H. Galbraith, Rev. Mod. Phys. **44**, 540 (1972); S. O'Kubo, J. Math. Phys. **16**, 528 (1975).
[8] P.A.M. Dirac, Proc. R. (London) Soc. Ser. A **155**, 447 (1936).
[9] J.D. Louck, Los Alamos Scientific Laboratory Report LA2451 (1960).
[10] A. Klein, J. Math. Phys. **4**, 1238 (1963); and V. Bargman and M. Moshinsky, Nucl. Phys. **18**, 697 (1960).
[11] S.A. Edwards, J. Math. Phys. **1**, 164 (1978); S. Okubo, J. Math. Phys. **12**, 2382 (1977); H.S. Green, J. Aust. Math. Soc. Series B, 129 (1972).
[12] V.S. Popov and A.M. Perelomov, Sov. J. Nucl. Phys. **5**, 89 (1967); M.K.F. Wong and H.Y. Yeh, J. Math. Phys. **16**, 1239 (1975); C.O. Nwachuku and M.A. Rashid, J. Math. Phys. **17**, 1611 (1976).
[13] J.D. Louck and L.C. Biedenharn, J. Math. Phys. **11**, 2368 (1970).
[14] U. Fano, Eastern Theoretical Physics Conference, 1962.
[15] L.C. Biedenharn and J.D. Louck, Commun. Math. Phys. **8**, 89 (1968).
[16] G.E. Baird and L.C. Biedenharn, J. Math. Phys. **4**, 1449 (1963).
[17] M.D. Gould, "On Tensor Operators and Characteristic Identities for Semisimple Lie Algebras," J. Aust. Math. Soc. Ser. B (1978) (to be published).
[18] J.G. Nagel and M. Moshinsky, J. Math. Phys. **6**, 682 (1965).
[19] M.K.F. Wong, J. Math. Phys. **8**, 1899 (1967).
[20] S.C. Pang and K.T. Hecht, J. Math. Phys. **8**, 1233 (1967).
[21] J.D. Louck, Los Alamos Scientific Laboratory Report LA2451 (1960).
[22] M.D. Gould, On the Characteristic Identity and Reduced Matrix Elements of U(n) and O(n), J. Aust. Math. Soc. Ser. B. (to be published).
[23] L.C. Biedenharn, A. Giovannini and J.D. Louck, J. Math. Phys. **8**, 691 (1967); G.E. Baird and L.C. Biedenharn, J. Math. Phys. **5**, 1730 (1964); **6**, 1847 (1965).
[24] I.M. Gel'fand and M.L. Zetlin, Dokl. Akad. Nauk. SSSR **71**, 825 (1950).
[25] J.E. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer, New York-Heidelberg-Berlin, 1972).
[26] A.U. Klimyk, Am. Math.Soc. Transl. Ser. 2 **76**, xx (1968).

The KdV prolongation algebra[a)]

W. F. Shadwick

Department of Mathematics, King's College, Strand, London WC2R 2LS
(Received 20 October 1978)

It is shown that the Wahlquist-Estabrook prolongation of the KdV equation determines a Lie algebra $\mathscr{L}$ and that $\mathscr{L}$ is infinite-dimensional. Finite-dimensional prolongation algebras are shown to be homomorphic images of $\mathscr{L}$. Two such algebras—SL (2,$\mathbb{R}$) and a five-dimensional nilpotent algebra $\mathscr{H}$ are obtained as quotients of $\mathscr{L}$ by ideals which are respectively maximal and maximal for a given property $\mathscr{L}$ determines a connection on a trivial vector bundle and the connection is flat iff the KdV equation is satisfied. The symmetries of the KdV equation are used to construct automorphisms of $\mathscr{L}$ and it is shown that $\mathscr{H}$ is distinguished by these automorphisms. It is demonstrated that $\mathscr{H}$ is closely related to the three classical KdV conservation laws and that the sequential prolongation process for KdV yields only representations of $\mathscr{H}$.

1. INTRODUCTION

In their paper "Prolongation structures of nonlinear evolution equations,"[1] Wahlquist and Estabrook introduced a technique (referred to here as WE prolongation) for "prolonging" a partial differential equation and applied it to the Korteweg–de Vries (KdV) equation. They found that the prolongation determined a structure which "comes close to defining a Lie algebra" and that, by considering a special case, they could associate to the KdV equation an eight-dimensional Lie algebra denoted here by $\mathscr{A}$. This algebra [or rather a subalgebra isomorphic to SL(2,$\mathbb{R}$)] was then used to obtain the inverse scattering problem and Bäcklund transformation appropriate to the KdV equation.

The purpose of the present paper is to investigate the algebraic structure of the WE prolongation of KdV. The necessary material from Ref. 1 is summarized in Sec. 2 using the notation of Wahlquist and Estabrook.

It is demonstrated that the WE prolongation determines an infinite-dimensional Lie algebra $\mathscr{L}$ whose construction is given in Sec. 3. Finite-dimensional algebras associated with the KdV equation by prolonging are shown to be homomorphic images of $\mathscr{L}$. They may thus be identified with factor algebras $\mathscr{L}/K$ where the ideals K are the kernels of the respective homomorphisms. It is demonstrated that particular algebras of interest [such as SL(2,$\mathbb{R}$)] may be identified in terms of ideals of $\mathscr{L}$ which are maximal or maximal for some given property.

The symmetries of the KdV equation are used to construct two automorphisms of $\mathscr{L}$. These are used to show that $\mathscr{L}$ is infinite-dimensional and to show how the "eigenvalue parameter" λ arises in the algebra $\mathscr{A}$.

Section 4 contains a description of the WE prolongation in terms of a connection determined by $\mathscr{L}$, whose curvature vanishes if and only if the KdV equation is satisfied. The relation between this connection and those determined by finite-dimensional prolongation algebras is discussed. The problem of using such algebras to find an auto-Bäcklund transformation for KdV is briefly considered.

In Sec. 5 it is shown that a second subalgebra $\mathscr{H} \subset \mathscr{A}$ is closely related to the symmetries of the KdV equation and also to its "classical" conservation laws. $\mathscr{H}$ is shown to be the algebra determined by the sequential prolongation process of Ref. 1.

2. THE WE PROLONGATION

The first step in applying the Wahlquist–Estabrook prolongation procedure is to replace the differential equation by an exterior differential system. For the KdV equation

$$u_t + u_{xxx} + 12uu_x = 0 \tag{2.1}$$

Wahlquist and Estabrook proceed as follows:

Let $N \subset \mathbb{R}^2$ be coordinatized by x and t, and let $M \subset \mathbb{R}^5$ be a fibre space over N with projection π,

$$M \xrightarrow{\pi} N$$

and coordinates x, t, u, z, p, where, as is usual, x and t have been written for π^*x and π^*t respectively.

Define $I(\alpha^4)$ to be the differential ideal of 2-forms on M generated by

$$\alpha^1 := du \wedge dt - z\, dx \wedge dt,$$

$$\alpha^2 := dz \wedge dt - p\, dx \wedge dt,$$

and

$$\alpha^3 := -du \wedge dx + dp \wedge dt + 12uz\, dx \wedge dt.$$

If $s : N \to M$ is a cross section[2] of $M \xrightarrow{\pi} N$ [so that $s(x,t) = (x,t,u(x,t),z(x,t),p(x,t))$] with the property

$$s^*I(\alpha^4) = 0, \tag{2.2}$$

then it is easy to verify that

$$z(x,t) = \partial u(x,t)/\partial x,$$

$$p(x,t) = \partial z(x,t)/\partial x, \tag{2.3}$$

and

[a)]Supported by the National Research Council of Canada.

 0022-2488/80/030454-08$1.00

$$\partial u(x,t)/\partial t + \partial p(x,t)/\partial x + 12u(x,t)z(x,t) = 0$$

so that $u(x,t)$ is a solution of (2.1). Conversely, given any solution of (2.1), the map $s : N \to M$ by

$$(x,t) \mapsto (x,t,u(x,t),\partial u/\partial x,\partial^2 u/\partial x^2)$$

is a cross section of $M \xrightarrow{\pi} N$ which satisfies (2.2). Thus the differential equation (2.1) may be replaced by the differential ideal $I(\alpha^A)$ on M with solutions of (2.1) corresponding to cross sections of $M \xrightarrow{\pi} N$ which satisfy (2.2). The submanifold $S := s(N)$ determined by such a cross section will be referred to as a *solution submanifold of M*.

Wahlquist and Estabrook next introduce an undetermined number n of 1-form ω^k on the product space $M \times Y$, where $Y \subset \mathbb{R}^n$ has coordinates y^k, $k = 1,...,n$. The ideal $I(\alpha^A)$ is lifted from M to $M \times Y$ by $\mathrm{pr}_1{}^*$, where $\mathrm{pr}_1 : M \times Y \to M$ is projection on the first factor and as above α^A is written for $\mathrm{pr}_1{}^* \alpha^A$. The 1-forms ω^k have the form

$$\omega^k = dy^k + F^k\,dx + G^k\,dt,$$

where F^k and G^k are functions on $M \times Y$ *which are taken to be independent of x and t.*[3] The prolongation procedure consists in requiring that the forms $\{\omega^k, \alpha^A\}$ comprise a differential ideal $I(\alpha^A, \omega^k)$. In order that this be true, there must exist functions f^k_A and 1-forms η^k_j such that

$$d\omega^k = f^k_A \alpha^A + \eta^k_j \wedge \omega^j.$$

(The summation convention for repeated upper and lower indices is used here and below.) This yields a system of partial differential equations (which will be referred to as the WE prolongation equations) to be solved for F^k and G^k.

The solution found by Wahlquist and Estabrook[1] is

$$F^k = 2X^k_1 + 2uX^k_2 + 3u^2 X^k_3, \tag{2.4b}$$

$$G^A = -2(p + 6u^2)X^k_2 + 3(z^2 - 8u^3 - 2up)X^k_3 + 8X^k_4 + 8uX^k_5 + 4u^2X^k_6 + 4zX^k_7, \tag{2.4b}$$

where the X^k_a are functions of the y's alone and the vector fields $X_a := X^k_a \partial_k$ (∂_k denotes differentiation with respect to y^k) satisfy

$$X_7 = [X_2, X_1], \quad X_5 = [X_1, X_7], \quad X_6 = [X_2, X_7], \tag{2.5a}$$

$$[X_1, X_3] = [X_2, X_3] = [X_1, X_4] = [X_2, X_6] = 0,$$

$$[X_1, X_5] + [X_2, X_4] = 0, \tag{2.5b}$$

$$[X_1, X_6] + X_7 - [X_4, X_3] = 0$$

with $[X_a, X_b]^k := X^j_a \partial_j X^k_b - X^j_b \partial_j X^k_a$.

The solution to the prolongation equations given by (2.4) and (2.5) is *unique* in its dependence on u, z, and p and, in a way which will be made explicit below, unique in its dependence on the y's up to maps of the vector fields X_a, which preserve Eqs. (2.5).

It is important to note that (2.5a) may be regarded as *defining* X_5, X_6, and X_7 as vector fields generated by $\{X_1,...,X_4\}$, i.e., as iterated commutators of members of the set $\{X_1,...,X_4\}$. From this point of view (2.5b) is a set of relations on the vector fields generated by $\{X_1,...,X_4\}$. In order to simplify the notation in what follows X_8 and X_9 may be defined, following Ref. 1, by

$$X_8 := [X_4, X_3]$$

and

$$X_9 := [X_1, X_5].$$

Wahlquist and Estabrook note that the X_a do not seem to close off to a finite-dimensional Lie algebra but that if the further condition

$$X_9 = \lambda (X_7 - X_8) \quad (\lambda \text{ constant}) \tag{2.6}$$

is imposed, the vector fields $X_1,...,X_8$ comprise a basis for an eight-dimensional Lie algebra denoted here by $\mathscr{A}$ (see Appendix A for the multiplication table for $\mathscr{A}$ in this basis).

The problem of integrating (2.5) may thus be replaced by the problem of finding a finite-dimensional representation of $\mathscr{A}$. Denote by $\bar{X}_a$, $a = 1,...,8$, a basis of $\mathscr{A}$ as an abstract Lie algebra with multiplication table, Table A.1, in Appendix A. By construction the $\bar{X}_a$ obey (2.5) viewed as a set of commutation relations so a representation of $\mathscr{A}$ by vector fields $\bar{X}_a \mapsto \bar{X}^k_a(y)\,\partial_k$ necessarily provides functions $\bar{X}^k_a$, which satisfy (2.5) viewed as a set of differential equations. The number $n = \dim Y$, left free until this point, is taken to be the dimension of the representation space.

Having found a representation of $\mathscr{A}$, one obtains a solution of the prolongation equations by substituting the $\bar{X}^k_a$ for the X^k_a in (2.4) above. The solution will be denoted by $(\bar{F}^k, \bar{G}^k)$. This technique for solving the prolongation equations is formalized in Sec. 3 below.

A *solution* of the prolonged system $I(\alpha^A, \bar{\omega}^k)$ (where $\bar{\omega}^k = dy^k + \bar{F}^k\,dx + \bar{G}^k\,dt$) is a cross section of $M \times Y \xrightarrow{\pi \circ \mathrm{pr}_1} N$ with the property

$$s^* I(\alpha^A, \bar{\omega}^k) = 0.$$

Thus s is of the form $s : (x,t) \to (x,t,u(x,t),z(x,t), p(x,t)y^k(x,t))$ and in addition to Eqs. (2.3), $s^* I(\alpha^A, \bar{\omega}^k) = 0$ yields the set of first order equations

$$y^k{}_{,x} = -\bar{F}^k, \quad y^k{}_{,t} = -\bar{G}^k, \quad k = 1,...,n.$$

In Ref. 1 it is equations of this type which lead to the inverse scattering problem and Bäcklund transformation.

The algebra $\mathscr{A}$ may be decomposed as the direct sum of ideals $\mathrm{SL}(2,\mathbb{R}) \oplus \mathscr{H}$, where $\mathscr{H}$ is a five-dimensional nilpotent Lie algebra whose multiplication table is given in Appendix A.[4] By using this decomposition any vector $\bar{X}_a \in \mathscr{A}$, $a = 1,...,7$, may be expressed uniquely as $\bar{X}_a = S_a + H_a$, where $S_a \in \mathrm{SL}(2,\mathbb{R})$ and $H_a \in \mathscr{H}$. Since $\mathrm{SL}(2,\mathbb{R})$ and $\mathscr{H}$ are ideals in $\mathscr{A}$, $[\mathrm{SL}(2,\mathbb{R}), \mathscr{H}] = 0$ so that the vectors $\{S_a\}$ and $\{H_a\}$, $a = 1,...,7$ separately satisfy (2.5). It follows that the functions $\bar{F}^k$ and $\bar{G}^k$ may be chosen to have the form

$$\bar{F}^k = \bar{F}^k_S + \bar{F}^k_{\mathscr{H}},$$

$$\bar{G}^k = \bar{G}^k_S + \bar{G}^k_{\mathscr{H}},$$

where the pairs $(\bar{F}^k_S, \bar{G}^k_S)$ and $(\bar{F}^k_{\mathscr{H}}, \bar{G}^k_{\mathscr{H}})$ separately provide solutions of the prolongation equations.

These will be referred to as the $\mathrm{SL}(2,\mathbb{R})$ and $\mathscr{H}$ prolongations respectively since $\mathrm{SL}(2,\mathbb{R})$ and $\mathscr{H}$ are the algebras

generated by $\{S_a\}$ and $\{H_a\}$, $a = 1,...,7$ respectively. As has been well documented, it is the SL(2,$\mathbb{R}$) prolongation which provides the inverse scattering equations and Bäcklund transformation.[1,4,7,15] The role of the $\mathscr{H}$ prolongation will be examined in Sec. 5 below.

3. THE CONSTRUCTION OF THE LIE ALGEBRA $\mathscr{L}$ ALGEBRA $\mathscr{L}$

It is possible to construct a Lie algebra $\mathscr{L}$ which is determined by (2.5) without any additional relations such as (2.6) being added "by hand." Thus, just as $\mathscr{A}$ may be associated with the particular solution of (2.5) given by (2.6), $\mathscr{L}$ may be associated with the general solution of (2.5). A faithful representation of $\mathscr{L}$, if it could be found, would yield the general solution to the prolongation equations.

In this section the construction of $\mathscr{L}$ is given. It is demonstrated that an algebra is associated to a particular solution of (2.5) if and only if it is homomorphic to $\mathscr{L}$. The technique of solving the prolongation equations which was discussed in Sec. 2 is formalized in terms of $\mathscr{L}$ and its homomorphic images. By making use of the symmetries of the KdV equation two automorphisms of $\mathscr{L}$ which may be used to probe its structure are found. In particular they are used to show that $\mathscr{L}$ is infinite-dimensional.

The construction of $\mathscr{L}$ is as follows:

Let $\mathscr{L}_f$ denote the free Lie algebra over the span of $\{X_1,...,X_4\}$.[5] Define the set $\{X_{ab}\}$:

$$X_{13} := [X_1,X_3],$$
$$X_{14} := [X_1,X_4],$$
$$X_{15} := [X_1,X_5] + [X_2,X_4],$$
$$X_{16} := [X_1,X_6] + X_7 - X_8,$$
$$X_{23} := [X_2,X_3],$$

and

$$X_{26} := [X_2,X_6] \quad \text{[cf. Eq. (2.5b)]}$$

Let R denote the ideal in $\mathscr{L}_f$ generated by $\{X_{ab}\}$. Then $\mathscr{L} := \mathscr{L}_f/R$ is the Lie algebra with generators $X_a + R$, $a = 1,...,4$, and relations $X_{ab} = 0$.[5] [Less formally, $\mathscr{L}$ is the algebra generated by $X_1,...,X_4$ with the relations (2.5) and all of the relations obtained by taking brackets with (2.5).] In the following, $X_a + R$ will be written as X_a for notational convenience.

It is easy to see that a Lie algebra $\mathfrak{A}$ is associated with a particular solution of (2.5) iff there is a surjective homomorphism h: $\mathscr{L} \to \mathfrak{A}$.

Suppose first that h is such a homomorphism and let $K = \ker h$ be the span of some set $\{Y_\delta \in \mathscr{L} \mid \delta \in \Delta\}$ where Δ is an indexing set. The particular solution of (2.5) is given by the equations $Y_\delta = 0$, $\delta \in \Delta$. Conversely, if a particular solution of (2.5) is given by a set of relations $Z_\delta = 0$, $\delta \in \Delta$, let K be the ideal generated by $\{Z_\delta \mid \delta \in \Delta\}$ and let $\mathfrak{A}$ be $\mathscr{L}/K$. The required homomorphism is $X \to X + K$.

Thus the ideals of $\mathscr{L}$ yield particular solutions of (2.5), and, if representations of the factor algebras can be found, these yield solutions of the prolongation equations.

The ideals of $\mathscr{L}$ have a distinguished subset, viz., the maximal ideals. (Recall that an ideal $K \subset \mathscr{L}$ is maximal if the only ideal of $\mathscr{L}$ which properly contains K is $\mathscr{L}$ itself.)

Let K_0 be the ideal in $\mathscr{L}$ generated by $\{X_1 + X_5, X_2 - X_6, X_3, X_4\}$. Then $\mathscr{L}/K_0$ is isomorphic to SL(2,$\mathbb{R}$),[6] and it is easy to verify that K_0 is a maximal ideal of $\mathscr{L}$. [Recall that if J is an ideal of $\mathscr{L}$ and $K \subset J$, then $\mathscr{L}/J \approx (\mathscr{L}/K)/(J/K)$. Since SL(2,$\mathbb{R}$) has no proper ideals, there is no ideal other than $\mathscr{L}$ which properly contains K_0.] One may also consider a wider class of ideals by looking at those which are maximal for some given property. For example, let $\mathscr{K} = \{K \text{ an ideal in } \mathscr{L} \mid \mathscr{L}/K \text{ is non-Abelian}\}$. (Non-Abelian factor algebras are of most interest since the Abelian prolongations yield only conservation laws; cf. Sec. 5 and Ref. 7). An ideal K will be called maximal relative to $\mathscr{K}$ if no member of $\mathscr{K}$ properly contains it. If K_1 denotes the kernel of the homomorphism $h:\mathscr{L} \to \mathscr{H}$ given in Example 3.1 below, then $\mathscr{L}/K_1 \approx \mathscr{H}$. It is easy to verify that K_1 is not a maximal ideal of $\mathscr{L}$ since it is properly contained in the ideal J generated by $\{X_7\}$. J does not belong to $\mathscr{K}$, however, and K_1 *is* maximal relative to $\mathscr{K}$.

Suppose now that $h:\mathscr{L} \to \mathfrak{A}$ is a homomorphism and that $\mathfrak{A}$ has basis $\{A_\mu\}$. For any $X_a \in \mathscr{L}$, hX_a may be written as

$$hX_a = h_a^\mu A_\mu.$$

If $r:A_\mu \to A_\mu^k(y)\,\partial_k$ is a representation of $\mathfrak{A}$, the functions

$$\bar{X}_a^k := h_a^\mu A_\mu^k(y) \quad \text{satisfy} \quad (2.5),$$

and one obtains a solution $(\bar{F}^k, \bar{G}^k)$ of the prolongation equations by substituting these expressions for X_a^k, $a = 1,...,7$, in (2.4). As in Sec. 2 above, a solution of the prolonged system $I(\alpha^A, \bar{\omega}^k)$ yields the set of first order equations

$$y_{,x}^k = -\bar{F}^k, \quad y_{,t}^k = -\bar{G}^k. \tag{3.1}$$

The map $r \circ h$ is a representation of $\mathscr{L}$, and if $\mathfrak{A}$ is finite-dimensional, r may be assumed to give a faithful finite-dimensional representation of $\mathfrak{A}$.[8] In this case however, the representation $r \circ h$ cannot be faithful: as will be demonstrated below, $\mathscr{L}$ is infinite-dimensional so $\ker h \neq \{0\}$ when $\mathfrak{A}$ is finite-dimensional.

To illustrate the above, a solution of the prolongation equations associated with the Lie algebra $\mathscr{H}$ is constructed in Example 3.1.

Example 3.1: $\{H_\mu\}$ is the basis of $\mathscr{H}$ given in Appendix A. Define $h:\mathscr{L} \to \mathscr{H}$ by

$$hX_a = \delta_a^\mu H_\mu, \quad a = 1,2,3,4.$$

It follows from Table A.2 and the homomorphism property $h[X,Y] = [hX,hY]$ that $hX_7 = hX_8 = H_5$ and that $\ker h$ is generated by $X_7 - X_8$. Since H_5 commutes with the rest of $\mathscr{H}$, it follows that

$$hX_5 = hX_6 = 0.$$

Thus $\bar{F}^k$ and $\bar{G}^k$ are given by

$$\bar{F}^k = 2H_1^k + 2uH_2^k + 3u^2H_3^k,$$
$$\bar{G}^k = -2(p + 6u^2)H_2^k + 3(z^2 - 8u^3 - 2up)H_3^k + 8H_4^k + 4zH_5^k,$$

where $H_\mu \to H_\mu^k\,\partial_k$, $\mu = 1,...,5$, is any representation of $\mathscr{H}$—for example,

$$H_1 \mapsto \tfrac{1}{2}\partial_1,$$
$$H_1 \mapsto \tfrac{1}{2}\partial_1,$$
$$H_2 \mapsto \tfrac{1}{2}y^1\partial_3,$$
$$H_3 \mapsto \tfrac{1}{3}y^2\partial_3,$$
$$H_4 \mapsto -\tfrac{3}{4}\partial_2,$$
$$H_5 \mapsto -\tfrac{1}{4}\partial_3. \qquad \square$$

The symmetries of the KdV equation may be "lifted" in a natural way to provide automorphisms of $\mathscr{L}$ as follows:

A diffeomorphism ϕ of M with the property

$$\phi^* I(\alpha^A) = I(\alpha^A) \tag{3.2}$$

will be referred to as a *symmetry* of the KdV equation. In addition to the obvious symmetries of space and time translation, there are two one-parameter groups of symmetries of the KdV equation.[9] These are the Galilean and scale change transformations given by ϕ_G and ϕ_S with

$$\phi_G : (x, t,u,z, p) \to (x + 12at,\ t,u + a,z, p)$$

and

$$\phi_S : (x, t,u,z, p) \to (e^{-a}x, e^{-3a}t, e^{2a}u, e^{3a}z, e^{4a}p),$$

where $(x,t,u,z,p) \in M$.

Any symmetry ϕ may be lifted to $M \times Y$ as $\tilde{\phi} = \phi \times \mathrm{id}_Y$ so that the diagram

$$\begin{array}{ccc} M \times Y & \xrightarrow{\tilde{\phi}} & M \times Y \\ \downarrow{\scriptstyle \mathrm{pr}_1} & & \downarrow{\scriptstyle \mathrm{pr}_1} \\ M & \xrightarrow{\phi} & M \end{array}$$

commutes.

Let $(x,t,u,z,p,y^k) = \tilde{\phi}(x',t',u',z',p',y^k)$ and denote by ω'^k and α'^A the pullbacks under $\tilde{\phi}^*$ of the forms ω^k and α^A on $M \times Y$. Thus $\omega'^k = \tilde{\phi}^* \omega^k = dy^k + F'^k dx' + G'^k dt'$ where

$$F'^k := F^k \partial x/\partial x' + G^k \partial t/\partial x' \quad \text{(a)}$$

and (3.3)

$$G'^k := F^k \partial x/\partial t' + G^k \partial t/\partial t' \quad \text{(b)}$$

It follows from $d\omega'^k = \phi^* d\omega^k$ and $I(\alpha^A) = I(\alpha'^A)$ that the forms $\{\omega'^k, \alpha'^A\}$ constitute a differential ideal and that F'^k and G'^k satisfy the same equations with respect to the primed variables as do F^k and G^k with respect to the unprimed variables. If F'^k and G'^k are independent of x and t (as they are when ϕ is a KdV symmetry), it follows from (2.4) that

$$\begin{aligned} F'^k &= 2X'^k_1 + 2u' X'^k_2 + 3u'^2 X'^k_3, \\ G'^k &= -2(p' + 6u'^2)X'^k_2 + 3(z'^2 - 8u'^3 - 2u'p')X'^k_3 \\ &\quad + 8X'^k_4 + 8u' X'^k_5 + 4u'^2 X'^k_6 + 4z' X'^k_7, \end{aligned} \tag{3.4}$$

where the X'_a obey the relations (2.5).

Thus $\mathscr{L}'$ generated as above by $\{X'_1 \cdots X'_4\}$ is isomorphic to $\mathscr{L}$ by $\hat{\phi}: X_a \to X'_a$ and identifying $\mathscr{L}'$ with $\mathscr{L}$ one sees that $\hat{\phi}$ is an automorphism. Moreover, the automorphism may be found explicitly by comparing (3.3) and (3.4).

The results for $\hat{\phi}_G$ and $\hat{\phi}_S$ are given in (3.5) and (3.6) below.

The x and t translation symmetries both give rise to the identity automorphism of $\mathscr{L}$:

$$\begin{aligned} \hat{\phi}_G: X_1 &\to X_1 - aX_2 + \tfrac{3}{2}a^2 X_3, \\ X_2 &\to X_2 - 3a X_3, \\ X_3 &\to X_3, \\ X_4 &\to X_4 - 3a X_1 + \tfrac{3}{2}a^2 X_2 - \tfrac{3}{2}a^3 X_3 - aX_5 + \tfrac{1}{2}a^2 X_6, \\ X_5 &\to X_5 - aX_6, \\ X_6 &\to X_6, \\ X_7 &\to X_7, \\ X_8 &\to X_8, \\ X_9 &\to X_9 + 2a(X_7 - X_8); \end{aligned} \tag{3.5}$$

$$\begin{aligned} \hat{\phi}_S: X_1 &\to e^a X_1, \\ X_2 &\to e^{-a} X_2, \\ X_3 &\to e^{-3a} X_3, \\ X_4 &\to e^{3a} X_4, \\ X_5 &\to e^a X_5, \\ X_6 &\to e^{-a} X_6, \\ X_7 &\to X_7, \\ X_8 &\to X_8, \\ X_9 &\to e^{2a} X_9. \end{aligned} \tag{3.6}$$

($\hat{\phi}_S$ and $\hat{\phi}_G$ are of course determined by their effect on the generators $X_1, \ldots, X_4$. The effect on $X_5, \ldots, X_9$ is given here for later convenience).

It is possible to obtain some informaton about the structure of $\mathscr{L}$ by using $\hat{\phi}_G$ and $\hat{\phi}_S$. For example, one may show that the ideal in $\mathscr{L}$ generated by $[X_1, X_2]$ is infinite-dimensional (and hence that $\mathscr{L}$ is infinite-dimensional). The proof is given in Appendix B.

It is easily verified from Eqs. (3.5) and (3.6) that the kernel K_1 of the map $h: \mathscr{L} \to \mathscr{H}$ given in Example 3.1 is *invariant* under $\hat{\phi}_G$ and $\hat{\phi}_S$. Thus $\mathscr{H}$ is closely related to the symmetries of the KdV equation. It will be shown in Sec. 5 that it is likewise related to the "classical" KdV conservation laws.

The automorphisms $\hat{\phi}$ pass to the quotient $\mathscr{L}/K$, where K is any ideal in $\mathscr{L}$ and thus induce isomorphisms of homomorphic images of $\mathscr{L}$:

Let $h: \mathscr{L} \to \mathfrak{A}$ be a surjective homomorphism so that $\mathfrak{A} \approx \mathscr{L}/\ker h$. Then if $K = \ker h$,

$$\mathfrak{A} \approx \mathscr{L}/K \approx \hat{\phi}(\mathscr{L})/\hat{\phi}(K),$$

where the second isomorphism is given by

$$X + K \to \hat{\phi}X + \hat{\phi}(K).$$

It is shown in Example 3.2 that the parameter λ which appears in (2.6) and in the multiplication table for $\mathscr{A}$ may be introduced via the Galilean automorphism $\hat{\phi}_G$.[10]

Example 3.2: $\{\bar{X}_\mu\}$ is the basis of $\mathscr{A}$ given in Appendix A. Define $h: \mathscr{L} \to \mathscr{A}$ by

$$hX_a = \delta^\mu_a \bar{X}_\mu \quad a = 1,4.$$

It is easy to verify that $\ker h$ is generated by

$X_9 - \lambda (X_7 - X_8)$. Thus it follows from (3.5) that $\hat{\phi}_G(\ker h)$ is generated by

$$\hat{\phi}_G [X_9 - \lambda (X_7 - X_8)] = X_9 - (\lambda - 2a)(X_7 - X_8),$$

where a is the Galilean parameter.

Thus $\mathscr{A}$ is one member of a family of isomorphic algebras $\hat{\phi}_G(\mathscr{L})/\hat{\phi}_G(K)$ parametrized by the Galilean parameter a. □

4. THE CONNECTION DETERMINED BY $\mathscr{L}$

It has been pointed out by Estabrook, Wahlquist, and Hermann[7] that the WE prolongation may be interpreted in terms of a Cartan–Ehresmann connection[11] on the trivial vector bundle $S \times Y$, where $S = s(N)$ and s is a cross section of $M \xrightarrow{\pi} N$. (See also Refs. 12, 13, and 4). This connection is flat when S is a solution submanifold of M. Given a finite-dimensional Lie algebra $\mathfrak{A}$ associated with a particular solution of the WE prolongation equations, the curvature forms of the connection may be computed from the ω^k and the structure constants $\Gamma^{\mu}_{\rho\nu}$ of a basis for $\mathfrak{A}$.

It is shown in this section that the same procedure may be followed (at least formally) for the Lie algebra $\mathscr{L}$. The curvature forms Ω^c of the connection determined by $\mathscr{L}$ span $I(\alpha^A)$ and thus vanish if and only if the KdV equation is satisfied. The curvature forms of connections given by algebras homomorphic to $\mathscr{L}$ are linear combinations of the Ω^c.

It follows from (2.4) that the 1-forms ω^k may be rewritten in the form

$$\omega^k = dy^k + \beta^a X^k_a, \tag{4.1}$$

where β^a are the 1-forms on M defined by

$$\begin{aligned}
\beta^1 &= 2dx, \\
\beta^2 &= 2[u\,dx - (p + 6u^2)\,dt\,], \\
\beta^3 &= 3[u^2\,dx + (z^2 - 8u^3 - 2up)dt\,], \qquad \text{(a)(4.2)} \\
\beta^4 &= 8dt, \\
\beta^5 &= 8u\,dt, \\
\beta^6 &= 4u^2\,dt, \\
\beta^7 &= 4z\,dt.
\end{aligned}$$

The sum in 4.1 is over $a = 1,2,...,7$, but it may clearly be extended to all $a = 1,2,\cdots$ by taking

$$\beta^a = 0, \quad a > 7. \qquad \text{(b)(4.2)}$$

Let s be a cross-section of $M \xrightarrow{\pi} N$ and let $S = s(N)$. Then the 1-form given by (4.1) may be interpreted as a connection on $S \times Y$. (If there is a Lie group $\mathscr{G}$ whose Lie algebra is $\mathscr{L}$, this is a $\mathscr{G}$-connection, but since $\mathscr{L}$ is infinite-dimensional such a group may not exist).

The curvature forms[11] Ω^c of the connection are defined by

$$\Omega^c := d\beta^c + \tfrac{1}{2} C^c_{ab}\, \beta^a \wedge \beta^b, \quad c = 1,2,\cdots,$$

where C^c_{ab} are the structure constants of $\mathscr{L}$. At this point one of the differences from the finite-dimensional case appears. The vectors $\{X_1,...,X_4\}$ may be taken as linearly independent and hence as part of a basis for $\mathscr{L}$, but not all multiple commutators of these generators will be linearly independent. It is easy to check (by using $\hat{\phi}_G$ and $\hat{\phi}_S$ as in Appendix B) that the set $\{X_1,...,X_9\}$ is linearly independent and hence may be taken as part of a basis. Thus the brackets $[X_a, X_b]$ with $a,b \in \{1,2,...,9\}$ give some of the structure constants of $\mathscr{L}$. It follows from (4.2) that $\beta^a \wedge \beta^b$ vanishes unless $a \in \{1,2,3\}$ and $b \in \{2,3,...,7\}$ so that, to compute the Ω^c, only those structure constants C^c_{ab} with a and b in these ranges are needed.

Also, since $\{[X_a, X_b] \mid a = 1,2,3,\ b = 2,3,...,7\} = \{X_5, X_6, X_7, X_8, X_9\}$ $C^c_{ab} = 0$ when $a \in \{1,2,3\}$, $b \in \{2,...,7\}$, and $c > 9$. Thus $\Omega^c = 0$ if $c > 9$, and the remaining curvature forms are easily calculated. They are

$$\begin{aligned}
\Omega^1 &= 0, \\
\Omega^2 &= -2(\alpha^3 + 12u\,\alpha^1), \\
\Omega^3 &= -6u(\alpha^3 + 12u\,\alpha^1) + 6z\,\alpha^2 - 6p\,\alpha^1, \\
\Omega^4 &= 0, \\
\Omega^5 &= 8\alpha^1, \\
\Omega^6 &= 8u\alpha^1, \\
\Omega^7 &= 4\alpha^2, \\
\Omega^8 &= 0.
\end{aligned} \tag{4.3}$$

The Ω^c generate $I(\alpha^A)$, and, if $S = s(N)$ is a solution submanifold of M, then $s^*\Omega^c = 0$ so the connection is *flat*.

Suppose now that h is a homomorphism, $h: \mathscr{L} \to \mathfrak{A}$, where $\mathfrak{A}$ has basis $\{A_\mu\}$ and structure constants $\Gamma^{\rho}_{\mu\nu}$ in this basis. The 1-forms $h\omega^k := dy^k + \beta^a (hX_a)^k$ may be written in terms of the basis $\{A_\mu\}$ as

$$h\omega^k = dy^k + \tau^\mu A^k_\mu \tag{4.4}$$

where

$$hX_a = h^\mu_a X_a \quad \text{and} \quad \tau^\mu := h^\mu_a \beta^a.$$

The curvature forms for the connection given by (4.4) are

$$\Lambda^\mu = d\tau^\mu + \tfrac{1}{2} \Gamma^\mu_{\nu\rho}\, \tau^\nu \wedge \tau^\rho$$

so that from the definition of τ^μ above and the homomorphism property $h^\nu_a h^\rho_b \Gamma^\mu_{\nu\rho} = h^\mu_c C^c_{ab}$ it follows that

$$\Lambda^\mu = h^\mu_a \Omega^a$$

(where the sums on a, b, c range over 1,...,9). Thus the curvature forms of the connection determined by $\mathfrak{A}$ are linear combinations of the Ω^c and so lie in $I(\alpha^A)$. It follows that the connection is flat on $S \times Y$ when S is a solution submanifold. It will not always be the case, however, that the 2-forms Λ^μ generate $I(\alpha^A)$. When they do, the connection is flat if and only if S is a solution submanifold of M and in this case the KdV equation is equivalent to the Maurer–Cartan equations for $\mathfrak{A}$. It is well known that this is the case for $\mathfrak{A} = \mathrm{SL}(2,\mathbb{R})$,[14,15] It is also true for $\mathfrak{A} = \mathscr{H}$,[16] as is demonstrated in Example 4.1 below.

Example 4.1: the notation is as in Example 3.1. The homomorphism $h: \mathscr{L} \to \mathscr{H}$ is given by

$$h^\mu_a = \delta^\mu_a, \quad a = 1,...,4,\ \mu = 1,...,5,$$

$$\begin{cases} h^{\mu}_{5} = h^{\mu}_{6} = h^{\mu}_{9} = 0 \\ h^{\mu}_{7} = h^{\mu}_{8} = \delta^{\mu}_{5} \end{cases}, \quad \mu = 1,...,5.$$

Thus the curvature forms for the $\mathscr{H}$ connection are

$$\Lambda^{1} = 0,$$
$$\Lambda^{2} = -2(\alpha^{3} + 12u\,\alpha^{1}),$$
$$\Lambda^{3} = -6u(\alpha^{3} + 12u\,\alpha^{1}) + 6z\,\alpha^{2} - 6p\,\alpha^{1},$$
$$\Lambda^{4} = 0,$$
$$\Lambda^{5} = 4\alpha^{2},$$

and these generate $I(\alpha^{A})$ unless $p \equiv 0$. □

If a flat connection is constructed as above and $\mathfrak{A}$ is the Lie algebra of some Lie group $\mathscr{G}$, then one may consider the vector bundle $S \times Y$ as associated with a principal $\mathscr{G}$ bundle over S with a "pure gauge" connection given by $g^{-1}\,dg$, $g \in \mathscr{G}$.[15]

Two such connections corresponding to solutions u_1 and u_2 of the KdV equation are then related by a "gauge transformation," and, *in the case where the gauge terms can be eliminated*, one may obtain a Bäcklund transformation connecting u_1 and u_2.[13,15,4] It is by no means clear, however, under what conditions this will occur. In the case of the $\mathscr{H}$ prolongation, for example, it does not seem to be possible to obtain a nontrivial KdV Bäcklund transformation in this way.

If the representation

$$H_1 \mapsto \tfrac{1}{2} y^1 \partial_3,$$
$$H_2 \mapsto \tfrac{1}{2} \partial_1,$$
$$H_3 \mapsto \tfrac{1}{3} \partial_2,$$
$$H_4 \mapsto -\tfrac{3}{4} y^2 \partial_3,$$
$$H_5 \mapsto \tfrac{1}{4} \partial_3$$

is used, Eqs.(3.1) for the $\mathscr{H}$ prolongation become

$$\begin{aligned} y^1_{,x} &= -u, \\ y^1_{,t} &= p + 6u^2, \\ y^2_{,x} &= = -u^2, \\ y^2_{,t} &= -u^2_x + 8u^2 + 2uu_{xx}, \\ y^3_{,x} &= -y^1, \\ y^3_{,t} &= 6y^2 - u_x. \end{aligned} \tag{4.5}$$

One might look for a (gauge) transformation

$$u' = f(u, y^1, y^2, y^3)$$

such that if u satisfies the KdV equation and the y's satisfy (4.5), u' is a second solution of KdV.[1,4] In terms of differential forms, the requirement is that the 2-forms

$$\alpha'^{1} := du' \wedge dt - z'\,dx \wedge dt,$$
$$\alpha'^{2} := dz' \wedge dt - p'\,dx \wedge dt,$$
$$\alpha'^{3} := -du' \wedge dx + dp' \wedge dt + 12u'z'\,dx \wedge dt$$

lie in $I(\alpha^{A}, \omega^{1}, \omega^{2}, \omega^{3})$. Unfortunately, the only solutions are

$$u' = u \quad \text{and} \quad u' = 0,$$

although the effect of the choice of representation of $\mathscr{H}$ on this result has not yet been determined.

5. THE ROLE OF THE $\mathscr{H}$ PROLONGATION

In Ref. 1 Wahlquist and Estabrook briefly consider a sequential prolongation procedure in which the new variables y^k are added one at a time. It will be shown in this section that the $\mathscr{H}$ prolongation is the result of this procedure.

The first step in the sequential process is to prolong the ideal by adding a conservation law,[7] i.e., a 1-form ω^1 of the form

$$\omega^1 = dy^1 + F^1(u,z,p)dx + G^1(u,z,p)dt,$$

with $d\omega^1 \in I(\alpha^A)$. Next, a conservation law is sought for the prolonged ideal $I(\alpha^A, \omega^1)$, i.e., one looks for a 1-form

$$\omega^2 = dy^2 + F^2(u,z,p,y^1)dx + G^2(u,z,p,y^1)dt,$$

with $d\omega^2 \in I(\alpha^A, \omega^1)$. [Of course, if F^2 and G^2 are independent of y^1, ω^2 is a conservation law for the smaller system $I(\alpha^A)$. This case corresponds to an Abelian prolongation as one may write the F^k and G^k in terms of commuting vector fields.]

At the jth stage the new variable y^j plays the role of a potential for the system $(\alpha^A, \omega^1, ..., \omega^{j-1})$ but, since F^j and G^j may depend on $y^1 \cdots y^{j-1}$, it is a "pseudopotential"[1] for $I(\alpha^A)$.

The resulting set of 1-forms ω^i, $i = 1,...,N$, for some fixed N clearly yields a particular solution of the WE prolongation equations, i.e., since $d\omega^i \in I(\alpha^A, \omega^1, ..., \omega^{i-1})$, it follows that $d\omega^i \in I(\alpha^A, \omega^1, ..., \omega^N)$. Thus the functions F^i and G^i are given by (2.4) and (2.5) with the additional condition that $\partial_j F^i = \partial_j G^i = 0$ if $j \geqslant i$. The vector fields X_a have the form

$$X_a = X^1_a \partial_1 + X^2_a(y^1)\partial_2 + \cdots + X^N_a(y^1 \cdots y^{N-1})\partial_N,$$

where X^1_a is constant. It follows that the commutator of X_a and X_b is given by

$$\begin{aligned} [X_a, X_b] &= (X^1_a \partial_1 X^2_b - X^1_b \partial_1 X^2_a)\partial_2 + [(X^1_a \partial_1 + X^2_a \partial_2)X^3_b \\ &\quad - (X^1_b \partial_1 + X^2_b \partial_2)X^3_a]\partial_3 + \cdots \\ &\quad + [(X^1_a \partial_1 + \cdots + X^{N-1}_a \partial_{N-1})X^N_b - (X^1_b \partial_1 + \cdots \\ &\quad + X^{N-1}_b \partial_{N-1})X^N_a]\partial_N. \end{aligned} \tag{5.1}$$

This equation together with (2.4) may be used to deduce that $X^i_7 = X^i_8$, $i = 1,2,...,N$ {since $X_7 = [X_2, X_1]$ and $X_8 = [X_4, X_3]$, it follows from (5.1) that $X^1_7 = X^1_8 = 0$ Repeated application of (5.1) together with the relations $[X_6, X_7 - X_8] = X_6$ and $[X_6, X_1] = X_7 - X_8$ then yield in turn $X^2_7 - X^2_8 = 0, ..., X^N_7 - X^N_8 = 0$}. The vector fields $X^k_a \partial_k$ obey (2.4) with the additional relation $X_7 = X_8$ and hence span an algebra which is homomorphic to $\mathscr{H}$ by $H_i \to X_i$, $i = 1,2,...,5$. Thus the sequential process produces a representation of the $\mathscr{H}$-prolongation. The lowest dimension N in which $\mathscr{H}$ has a faithful representation is 3. A convenient choice is the representation given in Example 3.1, in which case Eq. (3.1) become

$$y^1_{,x} = -1,$$
$$y^1_{,t} = 0, \tag{5.2a}$$
$$y^2_{,x} = 0,$$
$$y^2_{,t} = 6, \tag{5.2b}$$
$$y^3_{,x} = -y^1 \cdot u - y^2 \cdot u^2,$$
$$y^3_{,t} = y^1(u_{xx} + 6u^2) + y^2(u_x^2 - 8u^3 - 2uu_{xx}) + u_x. \tag{5.2c}$$

Here y^1 and y^2 are potentials for $I(\alpha^A)$ while y^3 is a potential for $I(\alpha^A, \omega^1, \omega^2)$. It follows from (5.2) that $y^1 = -(x + c_1)$ and $y^2 = 6(t + c_2)$, where c_1 and c_2 are constants. Thus y^3 could have been obtained as a potential rather than a pseudopotential for $I(\alpha^A)$ if explicit $x - t$ dependence were allowed in F^k and G^k. The integrability conditions for (5.2c) yield the conservation law

$$[(x + c_1)u - 6(t + c_2)u^2]_t + [(x + c_1)(u_{xx} + 6u^2) + 6(t + c_2)(u_x^2 - 8u^3 - 2uu_{xx}) - u_x]_x = 0.$$

Since c_1 and c_2 are arbitrary constants this contains the three "classical" KdV conservation laws

$$(u)_t + (6u^2 + u_{xx})_x = 0,$$
$$(u^2)_t + (-u_x^2 + 8u^3 + 2uu_{xx})_x = 0,$$

and

$$(xu - 6tu^2)_t + (x(u_{xx} + 6u^2) + 6t(u_x^2 - 8u^3 - 2uu_{xx}) - u_x)_x = 0.$$

The third of these is the conservation law which Noether's theorem associated with the Galilean invariance of the KdV Lagrangian.[10]

Note that this representation of $\mathscr{H}$ contains *all* of the conservation laws $T_t + X_x = 0$ for KdV in which T and X are functions of the coordinates of M, i.e., of x, t, u, u_x, and u_{xx},[17] and that only the first two could be obtained from an Abelian prolongation with no explicit $x - t$ dependence.

ACKNOWLEDGMENTS

Useful discussions with Felix Pirani and David Robinson are gratefully acknowledged. The author wishes to thank B. Kupershmidt for pointing out an error in the original version of the proof that $\mathscr{L}$ is infinite-dimensional.

APPENDIX A: MULTIPLICATION TABLES FOR $\mathscr{A}$ AND $\mathscr{H}$

TABLE A.1. The Lie algebra $\mathscr{A}$. λ is a constant. The entry in row $\bar{X}_a$, column $\bar{X}_b$ is $[\bar{X}_a, \bar{X}_b]$.

	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$	$\bar{X}_5$	$\bar{X}_6$	$\bar{X}_7$	$\bar{X}_8$
$\bar{X}_1$	0	$-\bar{X}_7$	0	0	$\lambda(\bar{X}_7 - \bar{X}_8)$	$\bar{X}_8 - \bar{X}_7$	$\bar{X}_5$	0
$\bar{X}_2$		0	0	$\lambda(\bar{X}_8 - \bar{X}_7)$	$\bar{X}_8 - \bar{X}_7$	0	$\bar{X}_6$	0
$\bar{X}_3$			0	$-\bar{X}_8$	0	0	0	0
$\bar{X}_4$				0	$\lambda^2(\bar{X}_8 - \bar{X}_7)$	$\lambda(\bar{X}_7 - \bar{X}_8)$	$-\lambda\bar{X}_5$	0
$\bar{X}_5$					0	$\bar{X}_7 - \bar{X}_8$	$-\bar{X}_5 - \lambda\bar{X}_6$	0
$\bar{X}_6$						0	$\bar{X}_6$	0
$\bar{X}_7$							0	0
$\bar{X}_8$								0

TABLE A.2. The Lie algebra $\mathscr{H}$. The entry in row H_a, column H_b is $[H_a, H_b]$.

	H_1	H_2	H_3	H_4	H_5
H_1	0	$-H_5$	0	0	0
H_2		0	0	0	0
H_3			0	$-H_5$	0
H_4				0	0
H_5					0

APPENDIX B: A PROOF THAT $\mathscr{L}$ IS INFINITE-DIMENSIONAL

The proof proceeds by assuming that the ideal generated by $Y_0 := [X_7 - X_8]$ is finite-dimensional. Then, by using the automorphisms $\hat{\phi}_S$ and $\hat{\phi}_G$, an additional relation, not implied by (2.5), is derived, thus contradicting the assumption.

Suppose the ideal generated by Y_0 is finite-dimensional. Then, by constructing a finite number of vectors Y_k, $k = 1,...,n$ with $Y_{k+1} := [X_1, Y_k]$, a linearly dependent set must be obtained. Thus, for some n.

$$Y_n = \sum_{i=1}^{n-1} c^i Y_i$$

Y_k, $k = 1,...,n-1$, linearly independent. It is clear from (3.6) that each of these scales by a different factor from Y_n under $\hat{\phi}_S$. Thus $c^i = 0$, $i = 1,...,n-1$, and $Y_n = 0$.

Let a be the Galilean parameter and define the operator $D: \mathscr{L} \to \mathscr{L}$ by

$$D := \frac{-d}{da}\bigg|_{a=0} \circ \hat{\phi}_G.$$

It follows from (3.5) that $DX_1 = X_2$, $DX_2 = 0$, and $DY_0 = 0$

so, from the definition of Y_n,

$$\hat{\phi}_G Y_n = Y_n - aDY_n \frac{+a^2}{2!} D^2 Y_n + \cdots + \frac{(-a)^n}{n!} D^n Y_n. \tag{B1}$$

It may be assumed that n is even, say $n = 2m$, as $Y_n = 0$ implies $Y_{n+1} = 0$. It may be shown by a lengthy but straightforward induction[18] that (2.5) implies that

$$D^m Y_{2m} = (-2)^m m! Y_0, \quad m = 0,1,\cdots. \tag{B2}$$

It follows from (B1) and (B2) that $\hat{\phi}_G Y_n = 0$ identically in the parameter a only if $Y_0 = 0$, that is, only if $X_7 - X_8 = 0$.

[1]H.D. Wahlquist and F.B. Estabrook, J. Math. Phys. **16**, 1 (1975).
[2]All considerations are local. All maps are C^∞.
[3]The effect of this restriction is discussed in Sec. 5.
[4]F.A.E. Pirani, D.C. Robinson, and W.F. Shadwick, "Local Jet Bundle Formulation of Bäcklund Transformations" preprint (1978).
[5]J.E. Humphreys, *Introduction to Lie Algebras and Representation Theory* (Springer-Verlag, New York, Heidelberg, Berlin, 1972).
[6]H. Levy (unpublished).
[7]F.B. Estabrook, H.D. Wahlquist, and R. Hermann, in *The* 1976 *Ames Research Center (NASA) Conference on the Geometric Theory of Non-Linear Waves.* Lie Groups: History, Frontiers and Applications, Vol. VI (Math. Sci. Press, Brookline, Mass., 1977).
[8]This follows from Ado's theorem—see, e.g., N. Jacobson, *Lie Algebras* (Interscience, New York and London, 1962), p. 202.
[9]B.K. Harrison and F.B. Estabrook, J. Math. Phys. **12**, 653 (1971).
[10]H. Steudel, Ann. Physik **32** (7), 445 (1975).
[11]R. Hermann, *Gauge fields and Cartan-Ehresmann Connections Part A*, Interdisciplinary Mathematics, Vol. X (Math. Sci. Press, Brookline, Mass., 1975).
[12]M. Crampin, F.A.E. Pirani, and D.C. Robinson, Lett. Math. Phys. **2**, 15 (1977).
[13]J. Corones, B.L. Markovski, and V.A. Rizov, J. Math. Phys. **18**, 2207 (1977).
[14]R. Hermann, Phys. Rev. Lett. **36**, 385 (1976).
[15]M. Crampin, "Solitons and SL(2,ℝ)," preprint (1978).
[16]This was pointed out by F.A.E. Pirani.
[17]R. Miura, C.S. Gardner, and M.D. Kruskal, J. Math. Phys. **9**, 1204 (1968).
[18]W.F. Shadwick, "The Bäcklund Problem, Symmetries and Conservation Laws for some Nonlinear Differential Equations," Ph.D. thesis, University of London, 1979.

Integration of near-resonant systems in slow-fluctuation approximation

Marijke F. Augusteijn[a] and Ernst Breitenberger
Department of Physics, Ohio University, Athens, Ohio 45701

(Received 5 September 1979; accepted for publication 15 October 1979)

Autonomous, conservative, nonlinear oscillatory systems of many degrees of freedom with an "internal resonance," i.e., with the linearized ("normal") frequencies tied in a near-resonant relation $\Sigma g_i\omega_i = \epsilon$, can be completely integrated in a transparent approximation which is a variant of the "method of slowly variable amplitude and phase." It consists in developing the nonlinear coupling into a trigonometric polynomial and dropping all terms except the most slowly varying ones on the grounds that their perturbative effects are oscillatory but essentially not cumulative. In this "slow-fluctuation" technique, the amplitudes separate off the phases; they exhibit joint extrema governed by simple conservation laws and all follow from one single quadrature. The phases also require no more than quadratures. There are various possibilities of motion with all amplitudes constant and harmonic time variation at (generally) incommensurate frequencies. The approximation does not apply if a degree of freedom which occurs to the first power in the nonlinear coupling happens to dip down towards zero amplitude; abrupt phase changes can then take place and call for a separate, accurate integration covering the critical time interval. With this sole restriction the method is entirely general. It is exact in the small-amplitude limit; at finite amplitudes it differs from the exact solution for practical purposes by no more than minor high-frequency ripples; and it yields more physical insight than any other known method for less effort.

I. INTRODUCTION

We are concerned with conservative, autonomous, oscillatory systems of $n \geqslant 2$ degrees of freedom (d.f.). For simplicity, we shall speak of coupled anharmonic oscillators, but we mean no narrow restriction. The exact scope of the term "oscillatory system" is defined below, after Eq. (2.4).

It will be most convenient to choose coordinates q_i such that in the linearized system they represent normal modes, i.e., at infinitesimal amplitudes we shall have

$$q_i = A_i \cos(\omega_i t + \beta_i), \tag{1.1}$$

with constant A_i, β_i (always $A_i \geqslant 0$). Thus, the Hamiltonian will be

$$H(p,q) = \tfrac{1}{2}\sum_1^n (p_i^2/m_i + m_i\omega_i^2 q_i^2) + H_1(q_1,\ldots,q_n), \tag{1.2}$$

with the nonlinear coupling member H_1 vanishing to an order higher than the second. We assume that H_1 is analytic[1] so that it can be developed into powers; no first- or second-order terms will be present, and we shall tacitly assume that a small number of higher ones suffice to characterize the system, but we leave it open how many there are, and whether or not one amongst them is "dominant" in any sense. Considerations of magnitude are, in fact, of limited import in what follows. We might also allow H_1 to depend on momenta as well as on coordinates, but this makes no profound difference to our results and we omit it for economy, excepting a couple of asides.

When one naively attempts to find $q_i(t)$ in the form of (truncated) Fourier series[2] the nonlinear H_1 gives rise in a familiar way to combination frequencies[3]

$$\Omega_k = k_1\omega_1 + \cdots + k_n\omega_n, \quad k_i \gtrless 0 \text{ and integer.} \tag{1.3}$$

If the normal frequencies ω_i of the linearized system happen to have such values that one of these Ω_k vanishes, the calculation of the corresponding Fourier amplitudes leads to a vanishing denominator ("resonant denominator," or "small divisor").[4]

The physical background of this well-known mathematical nuisance is easy to understand. The frequencies of nonlinear oscillators depend in general on their amplitudes.[5] If several of them are coupled together in a regime which permits notable exchange of energy and therefore may rightly be called resonant, their frequencies will change because the amplitudes change. Fourier series, with their rigid frequencies and harmonics, are ill adapted to this process. One had better use a "method of variable amplitude and phase", i.e., seek solutions of the form (1.1) with A_i, β_i time dependent, or some close equivalent.

This was first attempted in a systematic way by Beth,[6] but his developments are cumbersome and seem to have remained virtually unknown. Later and independently, Mettler[7] obtained similar results by a trenchant canonical transformation technique; in a certain approximation, a family of first integrals can be found explicitly and leads to solutions $q_i(t)$ by *quadratures alone*, an inestimable advantage.

Both authors strictly treat only an *exact* resonance $\Omega_k = 0$ and try from there to arrive at an understanding of a moderately detuned system with, say,

$$g_1\omega_1 + \cdots + g_n\omega_n = \epsilon,$$
$$g_i \gtrless 0 \text{ integer and } |\epsilon| \ll |\omega_i| \text{ for all } i. \tag{1.4}$$

[a]This paper is based on a dissertation which will be submitted to Ohio University by M.F.A. in partial fulfillment of the requirements for the Ph.D. degree.

 0022-2488/80/030462-10$1.00

Of course, this situation is much more likely to occur in practice. Unfortunately, the attempts of Beth and Mettler[8] to deal with a near resonance are not tenable. At best, they involve a step of grafting the ϵ in Eq. (1.4) arbitrarily onto one of the ω_i and therefore lead to formulas dependent on the investigator's whim.

A transparent approximation especially suited to near resonances has been introduced in a problem of two d.f. by Mueller.[9] It can readily be generalized to several d.f. as follows: When one Ω_k vanishes exactly, the corresponding $\cos\Omega_k t$ is constant. When it does not vanish [$\Omega_k = \epsilon$ as by Eq. (1.4)], then the time variation $\cos\epsilon t$ is still only a slow one, or rather all the *other* combination frequencies in H_1 contribute *faster* variations. In the equations of motion these would contribute fluctuating perturbations the up and the down effects of which are likely to cancel if only the half-periods are short enough. Thus, drop them altogether and retain only the "slow-fluctuation" term with the frequency ϵ because it is the only one that can have a prominent *cumulative* effect on the motions. The idea is fairly obvious[10] and may well have been used by others; however, it has never been systematically developed.

Roughly speaking, Mueller's slow-fluctuation approximation is akin to the frequency-averaging techniques of engineering mechanics. They are well known to be broadly successful with one-dimensional oscillators and can sometimes be generalized to more dimensions, when only exact resonances occur so that an unambiguous averaging interval can be defined. Another paper by Mettler[11] contains examples; one of these deals with the same system as Mueller[9] and leads (at exact resonance!) to the same equations of motion[12] as the slow-fluctuation approach. More important for our present purposes is that at exact resonance Mueller's approximation also coincides with that used by Beth[6] and Mettler.[7] This observation has prompted our work.

In the present paper, we use Mettler's canonical transformations to deal in slow-fluctuation approximation with systems having near-resonances of the type (1.4). We claim no originality except for the generalization of the procedures of Mettler and Mueller. Nonetheless, we have good reasons for putting our work forward. First, the slow-fluctuation technique has been thoroughly tested in one case of two d.f.[9,13] and was found to be capable of uncanny accuracy. Thus, it may be confidently proposed for use elsewhere. Second, our methods are altogether general and apply to all such near-resonant systems regardless of the deep differences among their various subtypes. For instance, it is known[14] that in two d.f. the resonant energy exchange is prominent only for frequency ratios near 1:1, 1:2, and 1:3 but not for others, with consequent differences in phase space topology. Such differences are no obstacles to our approach, which therefore may even serve to explore them in turn. Indeed we shall develop details mainly in order to exhibit the potential of further applications.

The scope being so wide, we found it impossible to single out one particular, realistic system as a paradigm. We have therefore deliberately constructed an artificial example to highlight the main features of the slow-fluctuation technique. This is summarized in Sec. VII, which may be read in parallel with the rest of the exposition.

Further developments directed at the (orbital and Liapunov) stability of near-resonant motions will be described in a separate paper.

II. THE SLOW-FLUCTUATION APPROXIMATION

We begin with the well-known canonical transformations

$$q_i = (2\bar{p}_i/m_i\omega_i)^{1/2}\cos\bar{q}_i\,, \qquad p_i = -(2m_i\omega_i\bar{p}_i)^{1/2}\sin\bar{q}_i\,, \tag{2.1}$$

for every $i = 1,...,n$. From the expression for q_i it is seen that in terms of the previous notation (1.1) the new variables are

$$\bar{q}_i = \omega_i t + \beta_i, \quad \bar{p}_i = \tfrac{1}{2}m_i\omega_i A_i^2. \tag{2.2}$$

In the uncoupled, purely harmonic case $H_1 \equiv 0$, the A_i and β_i are constant; Eq. (2.2) shows they are then even equivalent to canonical integration constants. The method of variable amplitude and phase is therefore tantamount to a "variation of canonical constants" in the $\bar{p}, \bar{q}$ system. If desired it could be formulated as a canonical perturbation theory with no more than changes in phrasing, but we feel no need for this further step of abstraction.

The transformation (2.1) turns a single harmonic oscillator term $\frac{1}{2}(p_i^2/m_i + m_i\omega_i^2 q_i^2)$ into $\omega_i\bar{p}_i$. The terms in H_1, of the general form

$$\gamma q_1^{h_1}\cdots q_n^{h_n}\,, \qquad h_1,...,h_n \text{ integer } \geqslant 0 \text{ and } \sum_i^n h_i \geqslant 3, \tag{2.3}$$

turn into products of cosines which we convert at once into sums of cosines by means of the familiar trigonometric identities. One of the latter, namely, $2\cos^2 x = 1 + \cos 2x$, also contains a 1 besides cosines. A term (2.3) with all the exponents h_i even will therefore convert into a sum with exactly one member free of cosines (in fact, a product of even powers of the $\bar{p}_i^{1/2}$). Thus, the transformed Hamiltonian takes the general form

$$\bar{H}(\bar{p},\bar{q}) = \sum_1^n \omega_i\bar{p}_i + \bar{B}(\bar{p}) + \sum_k \bar{F}_k(\bar{p})\cos(k_1\bar{q}_1 + \cdots + k_n\bar{q}_n), \qquad k_i \gtrless 0 \text{ and integer}, \tag{2.4}$$

with $\bar{B}$ a polynomial, the $\bar{F}_k$ polynomials in integral and half-odd (but not negative) powers, and the summation depending on the detailed structure of $H_1(q)$ (and with sines present in the sum beside the cosines, in case H_1 should depend on momenta as well as on coordinates).

In practice, oscillatory Hamiltonians often arise from reductions and manipulations. The "oscillators" may be formal variables, not real system components. Likewise, the apparent mass m_i may be negative; the harmonic-oscillator part of H is then not positive definite. This happens in widely divergent areas: celestial mechanics,[15] plasma waves,[16] and even quantum electrodynamics.[17] The transformation (2.1)

can still be applied in such cases, provided we compensate for the negative m_i under the square roots by ascribing a negative sign to ω_i too, and then we automatically obtain the term $\omega_i\bar{p}_i$ in $\bar{H}$ with the required negative sign. Accordingly, we now *define* an *oscillatory system* for our purposes as one having a Hamiltonian of type (2.4), but with m_i and ω_i allowed to be jointly negative. Any such system can be translated back into simple oscillator language by means of Eq. (2.1) without ambiguity; note especially from Eq. (2.2) that $\bar{p}_i$ always remains nonnegative.

The exact equations of motion for the variable amplitudes are

$$\dot{\bar{p}}_i = -\frac{\partial \bar{H}}{\partial \bar{q}_i}. \tag{2.5}$$

The differentiation converts the cosines of Eq. (2.4) to sines, but does not affect the trigonometric arguments

$$k_1\bar{q}_1 + \cdots + k_n\bar{q}_n = \Omega_k t + k_1\beta_1 + \cdots + k_n\beta_n\,, \tag{2.6}$$

where Eqs. (1.3) and (2.2) have been used. Assume now that one of the Ω_k, with the particular coefficients $k_i = g_i$, is small in the sense of Eq. (1.4). Provided $g_1\beta_1 + \cdots + g_n\beta_n$ also varies slowly in time, the particular term with $\cos(g_1\bar{q}_1 + \cdots + g_n\bar{q}_n)$ is the only one in Eq. (2.5) that in the spirit of Mueller's slow-fluctuation approach needs to be retained for its cumulative effect.

On the other hand, the exact phase variabilities follow from $\dot{\bar{q}}_i = \partial\bar{H}/\partial\bar{p}_i$, or upon omission of an ω_i right and left:

$$\dot{\beta}_i = \frac{\partial \bar{B}}{\partial \bar{p}_i} + \sum_k \frac{\partial \bar{F}_k}{\partial \bar{p}_i}\cos(\Omega_k t + k_1\beta_1 + \cdots + k_n\beta_n). \tag{2.7}$$

At infinitesimal amplitudes $\bar{p}_i$, the β_i are constant. Moreover, our problem is "well posed" in the sense of Hadamard: If we change initial conditions continuously so as to increase the amplitudes to finite values, the β_i will develop continuously into time-varying functions as the $\dot{\beta}_i$ develop continuously from zero, provided only the $\partial\bar{B}/\partial\bar{p}_i$ and $\partial\bar{F}_k/\partial\bar{p}_i$ are continuous. Hence, at amplitudes $\bar{p}_i$ not too large all the β_i will indeed fluctuate slowly as well as little, and now we can safely argue that of all the cosines in Eqs. (2.5) and (2.7) it is indeed only the one slowly fluctuating $\cos(g_1\bar{q}_1 + \cdots + g_n\bar{q}_n)$ which deserves to be kept, all the others fluctuating so much faster that their net effects may be deemed ignorable.

There is one singular case, though. The $\partial\bar{B}/\partial\bar{p}_i$ are always bounded, $\bar{B}$ being a polynomial, but not necessarily the $\partial\bar{F}_k/\partial\bar{p}_i$. If an $\bar{F}_k$ contains one momentum, say $\bar{p}_l$, to the power $\frac{1}{2}$, then $\partial\bar{F}_k/\partial\bar{p}_l$ contains $\bar{p}_l^{-1/2}$; if furthermore the initial conditions happen to be such that at some time t_1 the amplitude $\bar{p}_l$ tends to zero, no Lipschitz condition is assured around $t = t_1$ and hence the integration of the exact system (2.5) and (2.7) near $t = t_1$ must be discussed separately regardless of any approximations employed at other times. The complication is accentuated when we study special states of motion with one or more amplitudes constant in time, for then we have to admit the possibility that $\bar{p}_l \equiv 0$ identically at all times, not merely at t_1. In this event, go back to the exact equations of motion in the original variables p, q and ask is $q_l \equiv 0$ a possible solution? It would be necessary that $\dot{p}_l = -\partial H/\partial q_l = -\partial H_1/\partial q_l \equiv 0$; thus, we arrive at a new problem in $n-1$ d.f., with all terms in p_l and q_l omitted from the Hamiltonian (1.2), and with an added constraint $\partial H_1/\partial q_l \equiv 0$. If this problem makes sense, it can be treated in its own right and we need concern ourselves no more with it here. However, it may be impossible to satisfy the constraint in a nontrivial manner; then $q_l \equiv 0$ is not a solution, $\bar{p}_l$ could vanish at most at isolated times t_1, and we find ourselves back at the previous, lesser difficulty. We shall deal with it in Sec. VI; until then we assume the absence of any exceptional low-amplitude conditions.

With this sole proviso we replace the exact and cumbersome Hamiltonian (2.4) by its slow-fluctuation approximation

$$\bar{S}(\bar{p},\bar{q}) = \sum_1^n \omega_i\bar{p}_i + \bar{B}(\bar{p}) + \bar{F}(\bar{p})\cos(g_1\bar{q}_1 + \cdots + g_n\bar{q}_n) \tag{2.8}$$

for the near-resonant case (1.4). While this step is plausible, it may not at a first glance seem fully convincing. However, we defer further justification until Sec. VIII at the end so as to have a wider perspective available.

The $\bar{F}$ in Eq. (2.8) is the $\bar{F}_k$ of the near-resonant cosine. It may have arisen from a single term in H_1; then from Eq. (2.3) simply

$$\bar{F}(\bar{p}) = \tilde{\gamma}\bar{p}_1^{\,h_1/2}\cdots\bar{p}_n^{\,h_n/2},$$
$$h_1,\ldots,h_n \text{ integer} \geqslant 0 \text{ and } \sum_1^n h_i \geqslant 3, \tag{2.9}$$

but there may be several terms in H_1 contributing. For example, if $H_1(q)$ contains both $q_1^3q_2$ and $q_1q_2^3$, both will yield the combination frequency $\omega_1 - \omega_2$, and if this almost vanishes (a situation somewhat like a subharmonic resonance), then $\bar{F}(\bar{p})$ will consist of two terms of the form (2.9). From the trigonometric identities it is easily seen quite in general that, if $q_1^{h_1}\ldots q_n^{h_n}$ and $q_1^{l_1}\ldots q_n^{l_n}$ have a combination frequency in common, the exponents h_i and l_i differ by some multiple of 2 for any i. As a consequence we have a fact which will be helpful in conjunction with Eq. (3.9). Take the common factor of all terms contributing to $\bar{F}(\bar{p})$ outside parentheses: It may contain half-odd powers, but inside the parentheses we shall have only integral powers of the $\bar{p}_i$ left; hence, the square $[\bar{F}(\bar{p})]^2$ is always a polynomial in integral powers of the $\bar{p}_i$, and its degree is seen to be at least 3.

In Eq. (2.8) we assumed the presence of only a single near resonance. Unless stated otherwise we restrict ourselves to this case; as we shall briefly show in the next section the behavior of a system with two simultaneous resonances is different. Note that in any case the coefficients g_i in a resonance need not all be different from zero. The d.f. with $g_i \neq 0$ and those with $g_i = 0$ form two distinct groups whose resonant behavior will be found to be quite different.

III. INTEGRATION OF THE AMPLITUDES

The Hamiltonian (2.8) immediately invites the point transformation

$$\bar{\bar{q}}_1 = g_1\bar{q}_1 + \cdots + g_n\bar{q}_n, \quad \bar{\bar{q}}_i = \bar{q}_i, \quad \text{for } i = 2,\ldots,n. \tag{3.1}$$

In order to make this canonical, we construct new momenta

by the standard technique.[18] Following Mettler,[7] a suitable generating function is

$$G(\bar{q},\bar{p}) = \left(\sum_1^n g_i \bar{q}_i\right)\bar{p}_1 + \sum_2^n \bar{q}_i \bar{p}_i \,. \tag{3.2}$$

From the appropriate formula $\bar{p}_i = \partial G/\partial \bar{q}_i$ we have then

$$\bar{p}_1 = g_1 \bar{p}_1, \quad \bar{p}_i = g_i \bar{p}_1 + \bar{p}_i, \quad \text{for } i = 2,...,n. \tag{3.3}$$

The numbering of our variables is arbitrary, of course, except that Eq. (3.3) requires us to associate $\bar{p}_1$ with $g_1 \neq 0$ to make the transformation 1–1. If this should not already be the case, renumber the original coordinates q_i. We also find it convenient later on to have $g_1 > 0$; if this should not already be the case, multiply Eq. (1.4) through with a minus sign. Without loss of generality we therefore complete the canonical transformation (3.1) and (3.3) by the *convention* $g_1 > 0$.

In the new variables the Hamiltonian (2.8) will be cyclic in $\bar{q}_2,...,\bar{q}_n$. Hence, the conjugate momenta are conserved:

$$\bar{p}_i = \text{const} = \alpha_i, \quad \text{for } i = 2,...,n, \tag{3.4}$$

or alternatively,[19] using Eqs. (3.3) and (2.2),

$$m_i \omega_i A_i^2 - (g_i/g_1) m_1 \omega_1 A_1^2 = 2\alpha_i, \quad i = 2,...,n. \tag{3.5}$$

Note that a d.f. with $g_i = 0$ always moves at constant amplitude. It would nevertheless not be correct to say that it does not resonate in any sense, for its phase may still be time dependent, as we discuss explicitly in the next section, and therein lies the possibility of some energy exchange with the rest of the system. This is a specifically nonlinear process without analog in linear systems.

The d.f. with $g_i \neq 0$ fall into two families: According to whether g_i is positive or negative, A_i will vary in the same sense as A_1 or opposite to it. For example, with a resonance $\omega_1 - \omega_2 = 0$ we may have A_1 and A_2 seesawing much like the amplitudes of two coupled harmonic oscillators, but with $\omega_1 + \omega_2 - 2\omega_3 = 0$, say, A_1 and A_2 must grow and shrink together, while both are still moving in opposition to A_3. In any of these cases, all the amplitudes reach their extrema jointly, an outstanding feature.

Under special initial conditions it may happen that one of the d.f. with $g_i \neq 0$ also moves at constant amplitude. Then it follows from Eq. (3.5) that all amplitudes remain constant. These special motions will be discussed separately in Sec. V.

Using the canonical constants and the relation (1.4), we find the following for the Hamiltonian (2.8) after the transformation (3.1) and (3.3):

$$\bar{\bar{S}}(\bar{p}_1,\alpha,\bar{q}_1) = \sum_2^n \omega_i \alpha_i + \epsilon \bar{p}_1 + \bar{\bar{B}}(\bar{p}_1,\alpha) + \bar{\bar{F}}(\bar{p}_1,\alpha)\cos\bar{q}_1 \,. \tag{3.6}$$

This is conserved:

$$\bar{\bar{S}} = \text{const} = E. \tag{3.7}$$

The equation of motion of the remaining amplitude $\bar{p}_1 = \bar{p}_1/g_1$ is

$$\dot{\bar{p}}_1 = -\frac{\partial \bar{\bar{S}}}{\partial \bar{q}_1} = \bar{\bar{F}}(\bar{p}_1,\alpha)\sin\bar{q}_1 \,, \tag{3.8}$$

from which $\bar{q}_1$ can be eliminated by means of Eqs. (3.6) and (3.7) with the result

$$\dot{\bar{p}}_1^2 = \bar{\bar{F}}^2 - \left[E - \sum_2^n \omega_i \alpha_i - \epsilon \bar{p}_1 - \bar{\bar{B}}\right]^2 = f(\bar{p}_1). \tag{3.9}$$

$\bar{\bar{F}}(\bar{p})^2$, which descends from $\bar{F}(\bar{p})^2$ through the linear substitution (3.3), is still a polynomial of at least the third degree (see the end of the previous section). Likewise, $\bar{\bar{B}}$ is a polynomial; hence, $f(\bar{p}_1)$ is a polynomial. The degree of f will generally be at least three, but in exceptional cases it can be reduced by accidental cancellation of the highest powers in $\bar{\bar{F}}^2$ and $\bar{\bar{B}}^2$; for example, with[20] $H_1 = (q_1^2 + q_2^2)^2$, unexpectedly f becomes quadratic if $\omega_1 = \omega_2$ and also $m_1 = m_2$.

Equation (3.9) has the form of an energy law for a conservative d.f. with coordinate $\bar{p}_1$; its solution is the usual

$$t - t_0 = \int d\bar{p}_1 \Big/ \sqrt{f(\bar{p}_1)} \tag{3.10}$$

followed by inversion. If $f(\bar{p}_1)$ is of the third or fourth degree, $\bar{p}_1$ will be an elliptic function of time (including degenerate cases). At higher degrees, analytic reductions are often possible.[21] Sometimes a trick substitution can be found.[22]

Whether numerical inversion and/or integration must be undertaken or not, the single quadrature (3.10) always suffices to furnish all amplitudes of the system, by Eq. (3.5).

Even without integrating, much can be learned from the roots of $f(\bar{p}_1)$. Foremost,[23] if there are two first-order roots $0 \leqslant R_1 < R_2$, and if $f(\bar{p}_1) > 0$ for $R_1 < \bar{p}_1 < R_2$, then $\bar{p}_1(t)$ is periodic with successive turning points R_1, R_2, R_1, etc.; the full period T is

$$T = 2\int_{R_1}^{R_2} d\bar{p}_1 \Big/ \sqrt{f(\bar{p}_1)} \tag{3.11}$$

(and is the period of all other amplitudes as well). In the limit as R_1 and R_2 coalesce, $\bar{p}_1$ will become constant (for more, see Sec. V). Another limiting case is that one of the two turning points is a higher-order root; then the integral (3.11) diverges and the motion is somewhat reminiscent of a limiting cycle, or of the approach to a cycle. We may also have $f(R_1) = 0$ and $f(\bar{p}_1) > 0$ for all $\bar{p}_1 > R_1$; then monotonic growth results, an event conceivable with the Hamiltonian not positive definite, and so forth.

In passing, we note that rather the same results come forth if H_1 also depends on momenta $H_1 = H_1(p,q)$. The slow-fluctuation Hamiltonian (2.8), and with it the final Eq. (3.6), would contain a sine beside the cosine; the equation of motion (3.8) then has a cosine and a sine in the same places, but with one sign change, and so the square-and-add process succeeds again in eliminating $\bar{q}_1$ to yield an equation just like Eq. (3.9). Elimination is possible in still another case, this one exotic yet not inconceivable, namely, if H_1 contains terms with such a wide range of powers that in Eq. (2.4) harmonics of the resonant frequency occur. They cannot be dropped, of course, because they do not fluctuate fast enough; thus, we might end up with, say, both $\cos\bar{q}_1$ and $\cos 2\bar{q}_1$ in the Hamiltonian (3.6). Clearly we may convert the latter together with the equation of motion into a system of equations for the two unknowns $\cos\bar{q}_1$ and $\sin\bar{q}_1$, then solve, square, and add. The result may not be as simple as Eq. (3.9), but again we are left with a single quadrature.

The picture changes greatly when we admit two simultaneous resonances

$$\sum_1^n g_i\omega_i = \epsilon, \quad \sum_1^n g_i'\omega_i = \epsilon',$$
$$|\epsilon|, |\epsilon'| \ll |\omega_i| \text{ for all } i, \tag{3.12}$$

and accordingly introduce new coordinates

$$\bar{\bar{q}}_1 = \sum_1^n g_i\bar{q}_i, \quad \bar{\bar{q}}_2 = \sum_1^n g_i'\bar{q}_i,$$
$$\bar{\bar{q}}_i = \bar{q}_i, \quad \text{for } i = 3,...,n. \tag{3.13}$$

Proceed as above: A suitable generating function is

$$G = \left(\sum_1^n g_i\bar{q}_i\right)\bar{\bar{p}}_1 + \left(\sum_1^n g_i'\bar{q}_i\right)\bar{\bar{p}}_2 + \sum_3^n \bar{q}_i\bar{\bar{p}}_i$$

and yields the new momenta

$$\bar{p}_1 = g_1\bar{\bar{p}}_1 + g_1'\bar{\bar{p}}_2,$$
$$\bar{p}_2 = g_2\bar{\bar{p}}_1 + g_2'\bar{\bar{p}}_2, \tag{3.14}$$
$$\bar{p}_i = g_i\bar{\bar{p}}_1 + g_i'\bar{\bar{p}}_2 + \bar{\bar{p}}_i, \quad \text{for } i = 3,...,n.$$

The slow-fluctuation Hamiltonian $\bar{\bar{S}}$ then contains $\cos\bar{\bar{q}}_1$ and $\cos\bar{\bar{q}}_2$ but is still cyclic in the other coordinates; hence, there is one integral less:

$$\bar{\bar{p}}_i = \alpha_i, \quad i = 3,...,n. \tag{3.15}$$

After some algebra, Eq. (3.15) can be expressed in terms of the amplitudes as

$$m_i\omega_i A_i^2 - \frac{g_1g_i' - g_ig_1'}{g_1g_2' - g_2g_1'} m_2\omega_2 A_2^2 + \frac{g_2g_i' - g_ig_2'}{g_1g_2' - g_2g_1'} m_1\omega_1 A_1^2 = 2\alpha_i, \quad i = 3,...,n. \tag{3.16}$$

Here the coefficients contain 2×2 determinants formed from the $2\times n$ matrix of the coefficients g_i, g_i'. At least one of these determinants must be different from zero, otherwise the two vectors **g** and **g'** are linearly dependent and we would be dealing with only one resonance. Without loss of generality, and possibly after renumbering the original variables, we may assume that $g_1g_2' - g_2g_1' \neq 0$; this will also ensure that Eq. (3.13) and (3.14) are 1–1. However, linear independence still encompasses degenerate cases which cannot arise in a physical system. Thus, if we solve Eqs. (3.12) for ω_1 and ω_2, neither of the solutions should vanish identically in all ω_i, for with a zero frequency there is no motion. By the same token, if there is one frequency, say ω_l, which takes part in both resonances beside ω_1 and/or ω_2, then its (nonvanishing) coefficients g_l and g_l' must form nonvanishing 2×2 determinants with g_1, g_1', and g_2, g_2'. In this case we see from Eq. (3.16) that for $i = l$ the conservation law definitely involves three amplitudes, instead of only two as in Eq. (3.5). The simplest example is $\omega_1 - \omega_3 = 0$ and $\omega_2 - \omega_3 = 0$ in three d.f. However, there may be no such frequency occurring in both resonances; then we may have vanishing determinants. A very simple case is $\omega_1 - \omega_3 = 0$ and $\omega_2 - \omega_4 = 0$ in four d.f.; here we obtain from Eq. (3.16) for $i = 3$ and $i = 4$ two conservation laws linking the amplitudes in pairs, namely, A_3 with A_1 and A_4 with A_2.

The salient feature in either case is that there exists no general reason why all amplitudes should necessarily reach their extrema jointly. With the second resonance present, the interplay of the variables has become far more complex. The equations for $\dot{\bar{\bar{p}}}_1$ and $\dot{\bar{\bar{p}}}_2$ in analogy to Eq. (3.8) contain $\sin\bar{\bar{q}}_1$ and $\sin\bar{\bar{q}}_2$, respectively, in addition to both $\bar{\bar{p}}_1$ and $\bar{\bar{p}}_2$, but there is only one conservation law $\bar{\bar{S}}(\bar{\bar{p}}_1, \bar{\bar{p}}_2, \alpha, \bar{\bar{q}}_1, \bar{\bar{q}}_2) = E$ available for elimination; in general, the variables can therefore not be separated and no straightforward quadratures result.[24]

IV. INTEGRATION OF THE PHASES

Using Eq. (3.1), we have the following for the variability of the phases with numbers 2,...,n:

$$\dot{\bar{q}}_i = \dot{\bar{\bar{q}}}_i = \frac{\partial\bar{\bar{S}}}{\partial\alpha_i} = \omega_i + \frac{\partial\bar{\bar{B}}(\bar{\bar{p}}_1,\alpha)}{\partial\alpha_i} + \frac{\partial\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha)}{\partial\alpha_i}\cos\bar{\bar{q}}_1. \tag{4.1}$$

Elimination of $\bar{\bar{q}}_1$ by means of Eq. (3.6) and omission of ω_i in accordance with Eq. (2.2) leaves

$$\dot{\beta}_i = \frac{\partial\bar{\bar{B}}(\bar{\bar{p}}_1,\alpha)}{\partial\alpha_i} + \frac{E - \Sigma_2^n\omega_i\alpha_i - \epsilon\bar{\bar{p}}_1 - \bar{\bar{B}}(\bar{\bar{p}}_1,\alpha)}{\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha)} \cdot \frac{\partial\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha)}{\partial\alpha_i}, \quad i = 2,...,n. \tag{4.2}$$

It is instructive to consider at once a d.f. with $g_i = 0$, which moves at constant amplitude. Suppose $H_1 = q_1q_3^3$ and $\omega_1 - 3\omega_3 = 0$; trivially then $\bar{\bar{B}} \equiv 0$, $\bar{\bar{F}}$ does not contain $\bar{\bar{p}}_2$, and therefore $\bar{\bar{F}}$ does not contain α_2. Hence, $\dot{\beta}_2 \equiv 0$, and q_2 moves harmonically at its linearized frequency ω_2. On the other hand, let $H_1 = q_1q_2^2q_3^3$ with the same resonance $\omega_1 - 3\omega_3 = 0$ (which is now quasisubharmonic); now $\bar{\bar{F}}$ does depend on α_2, the equation for $\dot{\beta}_2$ contains $\bar{\bar{p}}_1(t)$, and so q_2 will exchange energy with the other d.f. via the phase $\beta_2(t)$ alone. It is clear, however, that this process depends delicately on given circumstances.

The equation for the remaining phase β_1 follows after precisely similar steps as

$$\sum_1^n g_i\dot{\beta}_i = \frac{\partial\bar{\bar{B}}(\bar{\bar{p}}_1,\alpha)}{\partial\bar{\bar{p}}_1} + \frac{E - \Sigma_2^n\omega_i\alpha_i - \epsilon\bar{\bar{p}}_1 - \bar{\bar{B}}(\bar{\bar{p}}_1,\alpha)}{\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha)} \cdot \frac{\partial\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha)}{\partial\bar{\bar{p}}_1}. \tag{4.3}$$

Thus, all phases can be obtained by quadratures alone, once $\bar{\bar{p}}_1(t)$ is known explicitly. General conclusions about the functions $\beta_i(t)$ can fortunately be drawn, because of the complexity of the equations (4.2) and (4.3).

If $\bar{\bar{p}}_1$ is constant, all β_i are constant. At constant amplitudes the system motion is therefore purely harmonic. The frequencies will be $\omega_i + \dot{\beta}_i$, and of course there is no general reason why they should be commensurate; for an exceptional case, see Sec. V.

If the amplitudes vary, they all have strictly the same period T given by Eq. (3.11). The phase variation $\dot{\beta}_i$ then contains a constant part and a part of period T; the former leads to a secular change in β_i but there is no reason why during T this change should amount to an exact multiple of 2π. Thus, the system as a whole is not strictly periodic, except under special circumstances.

The n quadratures (4.2) and (4.3) complete the previous

n integration constants $\alpha_2,...,\alpha_n$ and E with n which are not more simple additive phase constants. However, the new ones are not necessarily fully independent of one another because when the amplitudes are not constant a supernumerary integration constant arises on the way, namely, the t_0 in Eq. (3.10). From Eq. (3.8) it is seen that whenever $\bar{\bar{q}}_1$ equals a multiple of π, the varying amplitudes go through one of their joint extrema; thus,

$$\bar{\bar{q}}_1(t=t_e) = \epsilon t_e + \sum_1^n g_i\beta_i(t=t_e) = r\pi, \qquad (4.4)$$

r integer, t_e amplitude extremum time,

is the most general way to formulate the corresponding relation between the phase constants. Note that Eq. (4.4) involves only the resonant d.f. with $g_i \neq 0$; the phase constants of the nonresonant ones always remain arbitrary.

The arbitrariness of the additive phase constants, except for Eq. (4.4), brings an enormous simplification to the classification of the system motions. At given $\alpha_2,...,\alpha_n$ and E, knowledge of a single set of phase functions $\beta_i(t)$ suffices: All system motions for the given $\alpha_2,...,\alpha_n$, E can be derived from this one set merely by phase shifting the cosine waves under the amplitude modulation curves $A_i(t)$, subject to Eq. (4.4) only.

These statements again continue to hold for a momentum-dependent H_1 : A phase equation and the energy law are two linear equations for $\cos\bar{\bar{q}}_1$ and $\sin\bar{\bar{q}}_1$; hence, solve, square, and add to eliminate towards quadratures as before. With a double resonance, however, the behavior of the phases seems difficult to treat and opaque to understand.

V. SYSTEM MOTION AT CONSTANT AMPLITUDES

All amplitudes will be jointly constant if $\dot{\bar{\bar{p}}}_1 \equiv 0$, and only then. According to Eq. (3.8), this can happen in two ways:

(I) $\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha)\equiv 0$, (II) $\sin\bar{\bar{q}}_1 \equiv 0$.

We may also turn to Eq. (3.9) and conclude that the constant value $\bar{\bar{p}}_1 = R$ is a second- or higher-order root of $f(\bar{\bar{p}}_1)$, for at a first-order root $\bar{\bar{p}}_1$ cannot remain constant.[23] Thus, a study of the roots of f will readily expose all constant-amplitude motions; however, it confounds the two above cases which are of quite distinct nature.

In Case (I) it is sufficient to consider constant amplitudes only; admissible values are those which satisfy the condition $\bar{\bar{F}} \equiv 0$. The simplest instance is that $\bar{\bar{F}}(\bar{p})$ contains but one term of the form (2.9); then it is enough to make any one factor $\bar{p}_i \equiv 0$, or in terms of the canonical constants, by Eqs. (3.3) and (3.4), to select as the initial condition

$$\bar{p}_i = g_i\bar{\bar{p}}_1 + \alpha_i = 0. \qquad (5.1)$$

The Hamiltonian (3.6) reduces in Case (I) to

$$\bar{\bar{S}} = \sum_2^n \omega_i\alpha_i + \epsilon R + \bar{\bar{B}}(R,\alpha). \qquad (5.2)$$

The integration constants $\alpha_2,...,\alpha_n$ are now restricted by the condition $\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha) = 0$; the remaining nth constant is either $\bar{\bar{p}}_1 = R$ or $\bar{\bar{S}} = E$ as related by Eq. (5.2). All phases remain arbitrary; for the frequencies we have directly from Eq. (5.2)

$$\dot{\bar{q}}_i = \frac{\partial\bar{\bar{S}}}{\partial\alpha_i} = \omega_i + \frac{\partial\bar{\bar{B}}(R,\alpha)}{\partial\alpha_i}, \quad \text{for } i=2,...,n, \qquad (5.3)$$

as in Eq. (4.1), and similar for $\dot{\bar{\bar{q}}}_1$ as in Eq. (4.3). If $\bar{\bar{B}} \equiv 0$, then all frequencies are just the linearized ones, an obvious result, for if both $\bar{\bar{F}}$ and $\bar{\bar{B}}$ vanish the entire nonlinear coupling has been neutralized.

Case (II) is different. For $\sin\bar{\bar{q}}_1 \equiv 0$ we need $\bar{\bar{q}}_1 = r\pi$, r integer. The initial phases can not be arbitrary in this case but must fulfill, say,

$$\bar{\bar{q}}_1(t=0) = \sum_1^n g_i\beta_i(t=0) = r\pi, \quad r \text{ integer}. \qquad (5.4)$$

Since now $\cos\bar{\bar{q}}_1 \equiv \pm 1$, we have from the Hamiltonian (3.6)

$$\dot{\bar{\bar{q}}}_1 = \frac{\partial\bar{\bar{S}}}{\partial\bar{\bar{p}}_1} = \epsilon + \frac{\partial\bar{\bar{B}}(\bar{\bar{p}}_1,\alpha)}{\partial\bar{\bar{p}}_1} \pm \frac{\partial\bar{\bar{F}}(\bar{\bar{p}}_1,\alpha)}{\partial\bar{\bar{p}}_1} = 0, \quad r_{\text{odd}}^{\text{even}}. \qquad (5.5)$$

This is the necessary relation between the integration constants $\alpha_2,...,\alpha_n$ and $\bar{\bar{p}}_1 = R$. It may be quite an involved equation in these constants, in contrast to the condition for Case (I) which can often enough be satisfied by just the vanishing of one amplitude.

The Hamiltonian (3.6) can in Case (II) be written as

$$\bar{\bar{S}} = \sum_2^n \omega_i\alpha_i + \epsilon R + \bar{\bar{B}}(R,\alpha) \pm \bar{\bar{F}}(R,\alpha), \quad r_{\text{odd}}^{\text{even}}. \qquad (5.6)$$

It serves to calculate the energy $\bar{\bar{S}} = E$, or alternatively R, exactly as Eq. (5.2). The frequencies follow from Eq. (5.6) as

$$\dot{\bar{q}}_i = \omega_i + \dot{\beta}_i = \frac{\partial\bar{\bar{S}}}{\partial\alpha_i} = \omega_i + \frac{\partial\bar{\bar{B}}(R,\alpha)}{\partial\alpha_i} \pm \frac{\partial\bar{\bar{F}}(R,\alpha)}{\partial\alpha_i}, \quad \text{for } r_{\text{odd}}^{\text{even}} \text{ and } i=2,...,n. \qquad (5.7)$$

The remaining frequency is linked to the others by Eq. (5.5):

$$\dot{\bar{\bar{q}}}_1 = \sum_1^n g_i\dot{\bar{q}}_i = 0. \qquad (5.8)$$

In two d.f., Eq. (5.8) simply means $g_1\dot{\bar{q}}_1 = -g_2\dot{\bar{q}}_2$, i.e., the two frequencies $\bar{q}_1 = \omega_1 + \dot{\beta}_1$ and $\bar{q}_2 = \omega_2 + \dot{\beta}_2$ are commensurable. This is the only instance of *strictly* periodic system motion we have perceived. In three or more d.f. there is no reason to expect commensurability, except as an accident. In any event, note how in Eq. (5.7) the frequency shifts $\dot{\beta}_i$ are always functions of the amplitudes. In Case (II) the nonlinear coupling has not been nullified; it is only given a kind of special tuning so as not to affect the amplitudes, whereas the frequencies remain distinctly amplitude dependent.

VI. THE EXCEPTIONAL LOW-AMPLITUDE CONDITIONS

We now return to the singular case shunted aside in Sec. II. The equations of motion in the original variables

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = -m_i\omega_i^2 q_i - \frac{\partial H_1}{\partial q_i},$$
$$\dot{q}_i = \frac{\partial H}{\partial p_i} = p_i/m_i \qquad (6.1)$$

are in no way singular; safe Lipschitz conditions hold for every physical motion. A singularity in the new variables $\bar{p}, \bar{q}$ is therefore an artifact intrinsic to the transformation (2.1).

How it can come about is easy to see. If $\bar{H}_1(\bar{p},\bar{q})$ contains $\bar{p}_l^{1/2}$, as is necessary for the singularity, then the original coupling $H_1(q)$ contains q_l to the first power:

$$H_1(q) = H_1'(q) + q_l H_1''(q) + q_l^2 H_1'''(q) + \cdots,$$

where $H_1', H_1'', \ldots$ are all independent of q_l, and specifically $H_1'' \not\equiv 0$. Consequently, the equations of motion

$$\dot{p}_l = -m_l\omega_l^2 q_l - H_1''(q) - O(q_l), \quad \dot{q}_l = p_l/m_l \tag{6.2}$$

contain a driving term $H_1''(q)$ which does not depend on q_l, plus other couplings of order $O(q_l)$ which will have only a vanishing effect in case the amplitude of q_l vanishes. The singularity of Sec. II arises just when the initial conditions are such that $\bar{p}_l$, or A_l, tends to zero at some time $t = t_1$. We now see that in this situation q_l behaves around t_1 like an oscillator driven by an external force $-H_1''$ which first brings the oscillation to a stop, then starts it up again; thus, the oscillator's proper motion must first be out of phase with the driving force, then in phase, and hence must undergo an abrupt phase reversal at $t = t_1$. This phase jump is the essence of the singularity.

For a quantitative argument, let us choose $t_1 = 0$; then assume that $A_l(0) = 0$ but $H_1''(q) \neq 0$ at $t = 0$, and set in hybrid notation

$$q_l = A_l \cos\bar{q}_l, \quad p_l = -m_l\omega_l A_l \sin\bar{q}_l, \quad \bar{q}_l = \omega_l t + \beta_l . \tag{6.3}$$

These formulas are correct in the uncoupled case $H_1 \equiv 0$, when A_l and β_l are constants, and now serve to introduce the convenient new variables A_l, β_l in Eq. (6.2) by variation of constants.

The variable phase β_l is calculated from

$$\frac{q_l}{p_l} = -\frac{\cos\bar{q}_l}{m_l\omega_l \sin\bar{q}_l}. \tag{6.4}$$

According to Eqs. (6.3), both q_l and p_l must tend to zero together with A_l, regardless of the behavior of β_l; when $t = 0$ is approached in Eq. (6.4) either from the right or from the left we must therefore apply l'Hôpital's rule and find from Eq. (6.2)

$$\lim \frac{q_l}{p_l} = \lim \frac{\dot{q}_l}{\dot{p}_l} = 0 \tag{6.5}$$

because $H_1'' \neq 0$. It follows from Eqs. (6.5) and (6.4) that

$$\lim \cos\bar{q}_l = 0 \text{ as } t \to 0 \pm . \tag{6.6}$$

With the substitutions (6.3) the equations (6.2) become

$$\dot{A}_l \sin\bar{q}_l + A_l\dot{\beta}_l \cos\bar{q}_l = D, \tag{6.7}$$

$$\dot{A}_l \cos\bar{q}_l - A_l\dot{\beta}_l \sin\bar{q}_l = 0, \tag{6.8}$$

where we have abbreviated the driving terms as

$$D = [H_1''(q) + O(q_l)]/m_l\omega_l$$

and know that

$$\lim_{t \to 0} D = H_1''(t=0)/m_l\omega_l \neq 0. \tag{6.9}$$

We make no such substitutions in the other equations (6.1) and do not even consider them; it suffices to know that they furnish nonsingular functions for us to insert in the right-hand side of Eq. (6.7).

The system (6.7) and (6.8) can be solved for the unknowns $\dot{A}_l$ and $A_l\dot{\beta}_l$:

$$\dot{A}_l = D \sin\bar{q}_l , \tag{6.10}$$

$$A_l\dot{\beta}_l = D \cos\bar{q}_l . \tag{6.11}$$

If in Eq. (6.10) we approach $t = 0$ from the right or from the left, D goes to the finite limit (6.9) and $\sin\bar{q}_l$ must tend to $+1$ or -1 in accordance with Eq. (6.6). Thus, the amplitude A_l goes to zero with a finite slope. Of course, $\dot{A}_l$ must be negative to the left of $t = 0$ and positive to the right; hence,

$$\lim\dot{A}_l = \pm\, |H_1''(t=0)|/m_l\omega_l \text{ as } t \to 0 \pm . \tag{6.12}$$

Concurrently, we have from Eqs. (6.10) and (6.12)

$$\lim \sin\bar{q}_l = \lim \sin\beta_l = \pm\, \mathrm{sgn} H_1''(t=0) \text{ as } t \to 0 \pm . \tag{6.13}$$

This is the surmised phase jump through 180°. Note that because of it the cosine in Eq. (6.6) does not vanish with a continuous tangent but with slope $\mp\, \mathrm{sgn} H_1''(t=0)$ as $t \to 0 \pm$.

Despite these discontinuities the motion $q_l(t)$ itself exhibits no jittery features. From Eqs. (6.3) it is easy to verify that q_l has a double zero at $t = 0$, while $\dot{q}_l$ goes through zero with a smooth tangent of slope $\ddot{q}_l(0) = -H_1''(t=0)/m_l$. Incidentally, the initial condition $q_l(0) = \dot{q}_l(0) = 0$ is the easiest means to visualize the emergence of this state of motion.

An illuminating detail can still be learned from Eq. (6.11). With another application of l'Hôpital's rule[25] we find

$$\begin{aligned}\lim\dot{\beta}_l &= \lim \frac{D\cos\bar{q}_l}{A_l} \\ &= -\lim D \lim \frac{\dot{\bar{q}}_l \sin\bar{q}_l}{\dot{A}_l} \\ &= -\lim \dot{\bar{q}}_l\end{aligned}$$

whether we let $t \to 0$ from the right or from the left, because of Eq. (6.10). Since $\dot{\bar{q}}_l = \omega_l + \dot{\beta}_l$, it follows that

$$\lim\dot{\beta}_l = -\omega_l/2 \text{ as } t \to 0 \pm . \tag{6.14}$$

From the standpoint of the slow-fluctuation method, Eq. (6.14) must be called a *fast* phase variation. In addition to the phase jump at $t = 0$ we have here a sharp transient which accompanies the starting up and slowing down of the oscillation, and which will need a few periods to die out.

The special cases that H_1 also contains momenta, or that H_1' has a first- or higher-order zero at $t = 0$, can now be treated with obvious adaptations of the argument. We may omit these for brevity.

While the singularity is only a feature of the description in terms of amplitudes and phases, it clearly vitiates the slow-fluctuation approach. Whenever it arises, one must integrate the system (6.1) numerically or otherwise in terms of the nonsingular variables p and q, and convert the result to amplitudes and phases. After a few periods of the troublesome q_l, when the transient (6.14) has died away, the solution may be spliced to a slow-fluctuation solution which can

be legitimately developed from then onwards. Eventually this will again approach a minimum of A_l, and should again be spliced to a direct solution of Eq. (6.1), and so forth. There is no reason why at successive amplitude minima the very stringent relation $q_l = \dot{q}_l = 0$ should repeat itself exactly, of course; each case would need separate study. Where A_l no longer vanishes exactly but goes to a very low minimum, the phase will no longer be discontinuous and the transient (6.14) will be milder, but the splicing process would still be necessary.

VII. AN EXAMPLE

In view of the wide variety of phenomena possible in near-resonant systems, we shall attempt to exemplify main features only. We choose for this purpose a system of three d.f. with coupling

$$H_1(q) = \gamma_1 q_1^2 q_2^2 + \gamma_2 q_1^3 q_3 \ . \tag{7.1}$$

All manipulative details will be left out as they are entirely elementary.

Substitution of Eq. (2.1) followed by trigonometric conversion transforms Eq. (7.1) into

$$\begin{aligned}\bar{H}_1(\bar{p},\bar{q}) &= 4\bar{\gamma}_1\,\bar{p}_1\,\bar{p}_2\cos^2\bar{q}_1\cos^2\bar{q}_2 \\ &\quad +4\,\bar{\gamma}_2(\bar{p}_1^3\,\bar{p}_3)^{1/2}\cos^3\bar{q}_1\cos\bar{q}_3 \\ &= \tfrac{1}{2}\bar{\gamma}_1\,\bar{p}_1\,\bar{p}_2\,[2+\cos(2\bar{q}_1-2\bar{q}_2)+2\cos 2\bar{q}_1 \\ &\quad +2\cos 2\bar{q}_2+\cos(2\bar{q}_1+2\bar{q}_2)]+\tfrac{1}{2}\bar{\gamma}_2 \\ &\quad \times(\bar{p}_1^{\,3}\,\bar{p}_3)^{1/2}[\cos(3\bar{q}_1-\bar{q}_3)+3\cos(\bar{q}_1-\bar{q}_3) \\ &\quad +3\cos(\bar{q}_1+\bar{q}_3)+\cos(3\bar{q}_1+\bar{q}_3)],\end{aligned} \tag{7.2}$$

where

$$\bar{\gamma}_1 = \gamma_1/m_1\omega_1 m_2\omega_2, \quad \bar{\gamma}_2 = \gamma_2(m_1^3\omega_1^3 m_3\omega_3)^{-1/2}.$$

Compare this with Eq. (2.4).

Since q_3 occurs in Eq. (7.1) to the first power, $\bar{p}_3$ occurs in Eq. (7.2) to the power $\frac{1}{2}$. A state of motion with $\bar{p}_3$ (or A_3) dipping down toward zero at some time t_1 will therefore be singular, regardless of any resonance, and must be treated separately as in Sec. VI. We need not deal with this complication further.

Resonant behavior in our system will result from near vanishing of one or two of the combination frequencies

$$2\omega_1 \pm 2\omega_2, \quad \omega_1 \pm \omega_3, \quad 3\omega_1 \pm \omega_3 \ .$$

For the sake of discussion, we take ω_1 and ω_3 to be both positive and adopt the single resonance

$$3\omega_1 - \omega_3 = \epsilon \quad \text{with} \quad |\epsilon| \ll \omega_1, \ \omega_3; \tag{7.3}$$

thus, $g_1 = 3, g_2 = 0$ and $g_3 = -1$. The third frequency ω_2 may be negative jointly with m_2.

The slow-fluctuation Hamiltonian (2.8) then becomes, upon keeping only the leading terms in the brackets of Eq. (7.2),

$$\begin{aligned}\bar{S}(\bar{p},\bar{q}) &= \sum_1^3 \omega_i\bar{p}_i + \bar{\gamma}_1\,\bar{p}_1\,\bar{p}_2 \\ &\quad +\tfrac{1}{2}\bar{\gamma}_2\bar{p}_1^{3/2}\bar{p}_3^{1/2}\cos(3\bar{q}_1-\bar{q}_3),\end{aligned} \tag{7.4}$$

with the first term in Eq. (7.1) still represented by $\bar{\gamma}_1\bar{p}_1\bar{p}_2 = \bar{B}(\bar{p})$ although that term does not resonate according to Eq. (7.3). From Eq. (3.5) we have the two amplitude integrals

$$m_2\omega_2 A_2^{\,2} = 2\alpha_2, \tag{7.5}$$

$$m_3\omega_3 A_3^{\,2} + \tfrac{1}{3}m_1\omega_1 A_1^{\,2} = 2\alpha_3 \ . \tag{7.6}$$

By means of the second canonical transformation (3.1) and (3.3), and using the constants α_2 and α_3, the slow-fluctuation Hamiltonan (7.4) becomes

$$\begin{aligned}\bar{\bar{S}} &= \epsilon\bar{p}_1 + \omega_2\alpha_2 + \omega_3\alpha_3 + 3\bar{\gamma}_1\alpha_2\bar{p}_1 \\ &\quad +\tfrac{1}{2}\bar{\gamma}_2(3\bar{p}_1)^{3/2}(-\bar{p}_1+\alpha_3)^{1/2}\cos\bar{q}_1 \ .\end{aligned} \tag{7.7}$$

The equation of motion (3.8) for $\bar{\bar{p}}_1 = \bar{p}_1/3$ is

$$\dot{\bar{\bar{p}}}_1 = -\frac{\partial\bar{\bar{S}}}{\partial\bar{\bar{q}}_1} = \tfrac{1}{2}\bar{\gamma}_2(3\bar{p}_1)^{3/2}(\alpha_3-\bar{p}_1)^{1/2}\sin\bar{\bar{q}}_1 \tag{7.8}$$

and elimination of $\bar{\bar{q}}_1$ as in Eq. (3.9) leads to

$$\begin{aligned}\dot{\bar{\bar{p}}}_1^{\,2} &= f(\bar{\bar{p}}_1) \\ &= \tfrac{27}{4}\bar{\gamma}_2^2\bar{\bar{p}}_1^3(\alpha_3-\bar{\bar{p}}_1) \\ &\quad -[(E-\omega_2\alpha_2-\omega_3\alpha_3)-(\epsilon+3\bar{\gamma}_1\alpha_2)\bar{\bar{p}}_1]^2.\end{aligned} \tag{7.9}$$

The polynomial $f(\bar{\bar{p}}_1)$ is of the fourth degree; hence, $\bar{\bar{p}}_1$ will be an elliptic function of time.

For the reduction of the integral (3.10) to standard form, the roots of $f(\bar{\bar{p}}_1)$ must be found by some algebraic or numerical technique. Clearly, f is of the structure type $h(x) = -x^4 + \alpha x^3 - (a+bx)^2$, where $\alpha > 0$ [because $\alpha \leqslant 0$ means $\alpha_3 \leqslant 0$, which according to Eq. (7.6) is either impossible or trivial]. Let us for the sake of discussion choose first $\alpha = 2$, $a = 0$, and $b = 1$; then simply $h(x) = -x^2(x-1)^2$ and we have two double roots at $x = 0$ and $x = 1$ corresponding to constant-amplitude motions. Next go from $a = 0$ to some small negative value, say $a = -0.01$; then $h(x) = -x^2(x-1)^2 + 0.02x - 0.0001$ and by quick graphing it will be seen that there are now four distinct, positive roots, two near the origin and two straddling $x = 1$, with $h(x)$ positive inside each pair and negative between the pairs. Hence, in this instance, for certain values of the system parameters, motion in two *distinct* amplitude ranges is possible for the *same* set of values for α_2, α_3, and E.

After $\bar{\bar{p}}_1(t)$, or $A_1(t)$, has been determined explicitly, $A_3(t)$ follows from Eq. (7.6). $A_2(t)$ equals its initial value which is equivalent to α_2 by Eq. (7.5).

The phase changes are found from Eqs. (4.2) and (4.3). We quote only the one for $i = 2$:

$$\dot{\beta}_2 = 3\bar{\gamma}_1\,\bar{\bar{p}}_1, \tag{7.10}$$

which is obviously not constant unless $\bar{\bar{p}}_1$ is. Since q_2 does not take part in our assumed condition (7.3), this is one of the cases where a quasiresonant energy exchange can proceed by phase variation only; this still holds if m_2, ω_2 are negative. Note also that if we are in a regime of two possible amplitude ranges for the same values of α_2, α_3, and E, then two different functions $\bar{\bar{p}}_1(t)$ must be used to calculate the phase changes; the latter can therefore differ appreciably in the two ranges.

Motion at constant amplitudes can be read off Eq. (7.8). As in Sec. V, we distinguish Cases (I) and (II):

(Ia) Make $\bar{\bar{p}}_1(t) \equiv 0$. It suffices to stipulate the initial condition $\bar{\bar{p}}_1(0) = 0$; the other amplitudes and the energy are connected by Eq. (5.2), which reduces to

$$E = \omega_2\alpha_2 + \omega_3\alpha_3 = \tfrac{1}{2}m_2\omega_2^2A_2^2 + \tfrac{1}{2}m_3\omega_3^2A_3^2, \quad (7.11)$$

as expected if $q_1 \equiv 0$ in Eq. (7.1); likewise, from Eq. (5.3) the motion is seen to be harmonic at the normal frequencies ω_2 and ω_3. The phase constants are arbitrary.

(Ib) An almost trivial subcase of (Ia) is $A_1 \equiv 0$ and $A_3 \equiv 0$: harmonic motion of q_2 only.

(Ic) Try $\bar{p}_1(t) \equiv \alpha_3$. By Eq. (5.1), this means $\bar{p}_3 \equiv 0$. However, q_3 occurs in the coupling Hamiltonian (7.1) to the first power, so we must go back to the exact equation of motion

$$\dot{p}_3 = -\frac{\partial H}{\partial q_3} = -m_3\omega_3^2q_3 - \gamma_2q_1^3 \quad (7.12)$$

and now see that $p_3 \equiv 0$ is impossible unless $q_1 \equiv 0$. Thus, we are led back to (Ib).

(II) For $\sin\bar{q}_1 \equiv 0$ we must choose initial phases such that Eq. (5.4) is fulfilled. The amplitudes must satisfy Eq. (5.5), which becomes after some simplifications

$$m_1\omega_1\epsilon + \frac{3}{2}\gamma_1A_2^2 \pm \frac{9}{8}\gamma_2A_1A_3 \mp \frac{1}{8}\gamma_2\frac{m_1\omega_1A_1^3}{m_3\omega_3A_3} = 0, \quad (7.13)$$

where $A_3 \neq 0$ [see the argument following Eq. (7.12)]. The energy (5.6) becomes

$$E = \tfrac{1}{2}\sum_1^3 m_i\omega_i^2A_i^2 + \tfrac{1}{4}\gamma_1A_1^2A_2^2 \mp \tfrac{1}{8}\gamma_2A_1^3A_3\,. \quad (7.14)$$

The constant frequencies (5.7) and (5.8) are given in terms of the amplitudes by

$$\dot{\bar{q}}_1 = \dot{\bar{q}}_3/3, \quad (7.15)$$

$$\dot{\bar{q}}_2 = \omega_2 + \tfrac{1}{2}\gamma_1A_1^2/m_2\omega_2, \quad (7.16)$$

$$\dot{\bar{q}}_3 = \omega_3 \pm \tfrac{1}{8}\gamma_2A_1^3/m_3\omega_3A_3. \quad (7.17)$$

Thus, the frequencies of the first and third d.f. are commensurate, but not all three. Note also that A_2 does not cause any frequency shifts.

The example incidentally demonstrates an annoying need to go hence and forth between several sets of variables. There is no neat alternative. One wants the amplitudes $A_i(t)$, of course, but the all-important equation of motion (3.8) for $\bar{\bar{p}}_1 = \bar{p}_1/g_1$ holds for an amplitude *squared* [cf. Eq. (2.2)] and it is best obtained in coordinates which promote the critical elimination of $\bar{\bar{p}}_2$, etc. Our notation follows Mettler[7]; it has the merit that single and double bars clearly mark the levels of canonical operation.

VIII. QUALITATIVE ASSESSMENT OF THE METHOD

It remains to be discussed in what measure the exact solutions of our approximate equations are approximate solutions of the exact equations. A quantitative answer would presuppose, of course, that we make quantitative assumptions about the coupling Hamiltonian H_1, and thus that we limit ourselves to a particular class of systems. We do not wish to do that. Fortunately, a qualitative discussion is nonetheless possible owing to the transparent nature of the method.

The slow-fluctuation equations of motion result from the exact ones (2.5) and (2.7) by the omission of all terms which fluctuate "fast". Let us now restore these terms and regard them as an additive perturbation on the slow-fluctuation derivatives $\dot{\bar{p}}_i, \dot{\bar{\beta}}_i$. We may calculate the perturbation explicitly, in a proper first approximation, by inserting for all $\bar{p}_i, \bar{q}_i$ those functions of t obtained in the slow-fluctuation approximation. Now each of the fast terms consists of two factors: an $\overline{F}_k(\bar{p})$ or $\partial\overline{F}_k/\partial\bar{p}_i$, and a trigonometric factor, a sine or cosine of the argument (2.6). Let us develop each factor into a Fourier series. There is no difficulty at all with the first factor: Since all $\bar{p}_i$ have the same period T, given by Eq. (3.11), $\overline{F}_k(\bar{p})$ and $\partial\overline{F}_k/\partial\bar{p}_i$ have a straightforward Fourier development with period T; note that there will be a time-free (or zero-order) term in it unless accidentally the first factor has a zero time average. The second, trigonometric factor will not often be truly periodic, for as we have shown in Secs. IV and V it is a rare occurrence for all phases to be exactly commensurate (mod 2π). Still, we may select some stretch of duration τ and develop that; if we choose τ longer than T we shall obtain a series which safely allows us to study the perturbations over an entire period of the amplitudes. This series will contain no terms with low periods such as $1/\tau$, but many with periods around the relevant combination frequency Ω_k; there may also be a zero-order term, but it can hardly be large, for although the phase variability creates some anharmonicity, it is generally too slow (cf. Sec. II) to produce a large time average of the trigonometric factor (in close analogy to FM radio signals).

The effect of the perturbing terms can now be found by a simple time integration. It is clear that we obtain mostly harmonic high-frequency terms with periods around the various combination frequencies Ω_k; in other words, a complicated fast ripple superimposed on every slow-fluctuation amplitude and phase.

Prominent nonperiodic, or secular, effects can only arise exceptionally, in two ways. If a system is (in slow-fluctuation approximation) *strictly* periodic, i.e., if the phases remain mod 2π in step with the amplitudes, then the periods of the two Fourier series which are combined in a fast term are commensurate, and the time integral of the product can contain a large term linear in t. Such motions are certainly rare, as we emphasized in Sec. IV, but they need to be considered. Secondly, it is conceivable that two Fourier series to be multiplied together *both* have a significantly large zero-order term; then there is again a significant contribution linear in t in the time integral of their product. This case can evidently not arise unless there is a coincidence of phase variabilities with much anharmonicity of trigonometric factors resulting.

The last-mentioned process will, of course, give rise to secular terms in any event, but they must be small, indeed very small, as long as the phases vary slowly enough, or the amplitudes are not too large.

To sum up, the slow-fluctuation method is suspect in the vicinity of strict system periodicity, it may require significant secular corrections whenever there is an accidentally fast combined variability of the phases, and it is not valid in the neighborhood of exceptional low-amplitude conditions (Sec. VI). In all other circumstances, it deviates from

the exact solutions essentially by no more than high-frequency ripples missing from both amplitudes and phases. Furthermore, the underlying integration problem being well posed, the ripples all must tend to zero continuously in the limit of vanishing amplitudes.

Since the troublesome cases are few and moreover easy to spot, the slow-fluctuation technique recommends itself as a most convenient approach to all near-resonant systems. No other method known to us yields so many results so accurately (except for ripples) at the price of no more than a few quadratures.

[1]Nonanalytic systems occasionally appear in the literature. e.g., B. Tschechankow and W. Popow, Z. Angew. Math. Mech. **53**, 134 (1973).

[2]"Integration by Trigonometric Polynomials"; the classic paper is by J. Horn, Z. Math. Phys. (Leipzig) **48**, 400 (1903).

[3]First encountered in the physiology of hearing; see H.v. Helmholtz, *On the Sensations of Tone*, (translated from the 1877 German edition) (Dover, New York, 1954), Chap. 7 and Appendix 12.

[4]The mathematical consequences were first noticed in the context of celestial mechanics; see H. Poincaré, C. R. Acad. Sci. **95**, 766 (1882); on the current status, see B. Garfinkel, Astron. J. **71**, 657 (1966).

[5]Exceptional oscillators with constant frequencies have been constructed by H. Ehrmann, Z. Angew. Math. Mech. **32**, 307 (1952).

[6]H.J.E. Beth, Philos. Mag. **26**, 268 (1913). We owe this reference to E. Mettler.

[7]E. Mettler, Z. Angew. Math. Mech. **43**, T81 (1963).

[8]Beth, Ref. 6, Sec. 13; Mettler, Ref. 7, Sec. (c).

[9]R.D. Mueller, "Orbital Instability—A Case Study in Two Dimensions," Ph.D. Dissertation, Ohio University (1969, unpublished).

[10]Mueller himself quotes as precursors R.A. Struble and J.E. Fletcher, J. Math. Phys. (N.Y.) **2**, 880 (1961); and J. Soc. Ind. Appl. Math. **10**, 314 (1962). Similar ideas may be read into various other papers, some of which are mathematically quite elaborate and directed at other questions. We mention J.H. Heinbockel and R.A. Struble, Z. Angew. Math. Phys. **14**, 262 (1963); B.I. Cheshankov, J. Appl. Math. Mech. **33**, 1075 (1969); N. Rott, Z. Angew. Math. Phys. **21**, 570 (1970); K.B. Dysthe and O.T. Gudmestad, J. Math. Phys. (N.Y.) **16**, 56 (1975); A.H.P. van der Burgh, J. Sound Vib. **49**, 93 (1976). Very close to Mueller's approach is the "averaging" of D.L. Hitzl, Celestial Mech. **5**, 433 (1972).

[11]E. Mettler, Ing. Arch. **28**, 213 (1959).

[12]Mettler, Ref. 11, Eqs. (18).

[13]E. Breitenberger and R.D. Mueller, "The Elastic Pendulum: A Nonlinear Paradigm" (to be submitted to this journal).

[14]D.J. Korteweg, Verh. K. Akad. Wet. Amsterdam Afd. Natuurkd. **5**, 10 (1897); K.T. Alfriend, Celestial Mech. **3**, 247 (1971); and Int. J. Non-Linear Mech. **6**, 563 (1971); J.A. Sanders, Celestial Mech. **6**, 421 (1977).

[15]For example, K.T. Alfriend and D.L. Richardson, Celestial Mech. **7**, 408 (1973).

[16]For example, K.B. Dysthe, Int. J. Electron. **29**, 401 (1970).

[17]For example, W. Heitler, *The Quantum Theory of Radiation* (Oxford University, New York, 1954), 3rd edition, Eq. (42) on p. 47.

[18]For example, H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 1950), p. 244.

[19]More or less similar first integrals for oscillatory systems have resulted from a remarkable variety of approximations, e.g., Mettler, Ref. 11; Heinbockel and Struble, Ref. 10; Hitzl, Ref. 10; van der Burgh, Ref. 10; F.B. Bretherton, J. Fluid Mech. **20**, 457 (1964); T.R. Kane and M.E. Kahn, J. Appl. Mech. **35**, 547 (1968). They also arise from the technique of successive, approximate canonical transformations introduced by G.D. Birkhoff, *Dynamical Systems* (American Mathematical Society Coll. Publ. Vol. IX, Providence, 1927), p. 82. This was recently used by Alfriend and Richardson, Ref. 15; it led to formulas which in appearance are very similar to ours, but their variables do not have simple physical meanings. G. Hori, Publ. Astron. Soc. Jpn. **18**, 287 (1966); and **19**, 229 (1967) uses essentially the same technique and relates his integrals to the "third integral" of Contopoulos.

[20]This problem has been integrated in closed form by U. Kirchgraber, Celestial Mech. **4**, 340 (1971).

[21]For helpful formulas on hyperelliptic integrals, see P.F. Byrd and M.D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists* (Springer, Berlin, 1971), 2nd edition, pp. 252–271.

[22]For example, J. Biermanns, Arch. Elektrotech. **3**, 345 (1915) integrated the Duffing system with an additional quintic term in closed form.

[23]The best discussion of this theorem known to us is in P. Frank and R.v. Mises, *Die Differential- und Integralgleichungen der Mechanik und Physik* (Vieweg, Brunswick, 1930), 2nd edition, Vol. 1, pp. 167–169.

[24]Some numerical work on doubly resonant systems is available, e.g., G.H. Walker and J. Ford, Phys. Rev. **188**, 416 (1969); J.M.A. Danby, in *Periodic Orbits, Stability and Resonances*, edited by G.E.O. Giacaglia (Reidel, Dordrecht, 1970), pp. 272–285. Some analytical work on a system somewhat different from ours has been done by L. Lewin and J. Vagners, Celestial Mech. **14**, 429 (1976).

[25]For a careful demonstration of the validity of this procedure in the case of one-sided limits, see J. Pierpont, *Lectures on the Theory of Functions of Real Variables* (Ginn, Boston, 1905), Vol. 1, Sec. 450; or M.H. Protter and C.B. Morrey, *A First Course in Real Analysis* (Springer, New York, 1977), p. 89.

Path integral in the representation of SU(2) coherent state and classical dynamics in a generalized phase space

Hiroshi Kuratsuji and Tōru Suzuki[a),b)]
Department of Physics, Kyoto University, Kyoto, Japan

(Received 7 May 1979; accepted for publication 11 October 1979)

Path integral in the representation of coherent state for the simplest semisimple Lie group SU(2) and its classical consequences are investigated. Using the completeness relation of the coherent state, we derive a path integral expression for the transition amplitude which connects a pair of SU(2) coherent states. In the classical limit we arrive at a canonical equation of motion in a "curved phase space" (two-dimensional sphere) which reproduces the ordinary Euler's equation of a rigid body when applied to a rotator.

1. INTRODUCTION

One of the main motives for the use of Feynman's path integral in quantum mechanics lies in its intuitive way of describing the correspondence between classical and quantum concepts. Especially the integration over paths in phase space[1] gives Hamilton's equation of motion in the classical limit. According to this the system is firstly supposed to propagate through infinite sequence of coordinate eigenstate, and then via the transformation to momentum representation at each time interval the transition amplitude is brought into the form of an integration over the paths in phase space. There is, however, another way of deriving the phase-space path integral through the introduction of the coherent state.[2]

The coherent state (hereafter abbreviated as CS), which is usually defined as an eigenstate of boson annihilation operators, is represented by the wave packet with minimal uncertainty the center of which corresponds to the point of phase space, and hence in the CS representation the quantum equation of motion just reduces to a canonical equation of motion. In this sense the path integral formulation through CS may provide us with the natural way of describing the quantal system in its intimate connection with the corresponding classical system, since the latter propagates through a definite path in the phase space. Thus the phase space path integral in the coherent state representation seems to be worthwhile for further investigation since the concept of CS has recently proved applicable to a wide class of physical systems through its extension to generalized coherent state.

As the ordinary CS is closely related with the unitary representation of the Heisenberg–Weyl group, so the generalized coherent state has been introduced by Perelomov[3] in relation to the unitary representation of an arbitrary Lie group. The simplest example is the CS based on a unitary irreducible representation of SU(2) group [hereafter called SU(2) CS], which coincides with the so-called Bloch state and serves as a powerful device for the superradiance phenomena, statistical mechanics of spin systems and so on.[4]

The main purpose of this paper is to derive the path integral representation of the transition amplitude for the case of SU(2), with the aid of general properties of coherent state.[5] It will also be shown that in the classical limit one arrives at an equation of motion in a "generalized phase space" which describes a classical motion of spin (angular momentum vector). Extension to the more general case is straightforward and will be treated in a separate paper with an application to many fermion systems.[6]

Section 2 is devoted to the brief review of the properties of SU(2) CS. In Sec. 3 we derive the path integral expression for the transition amplitude in the space of SU(2) CS. In Sec. 4 we investigate the classical equation of motion obtained in the limit of $\hbar \to 0$ and its relation to the geometrical structure of the phase space. An application to rigid rotator is given in Sec. 5 where the "quantization" procedure of the classical system is also discussed.

2. PROPERTIES OF THE SU(2) COHERENT STATES

In this section we recapitulate the basic properties of the SU(2) CS with slight modifications.

A. Construction of the SU(2) CS

According to Ref. 3 the generalized CS is given by the set $\{U(g)|0\rangle, g\in G\}$, where $U(g)$ is the unitary representation of the Lie group G acting on a Hilbert space and $|0\rangle$ is a fixed vector in this space. In the case of $G = \mathrm{SU}(2)$, $U(g)$ can be parametrized as $\exp(\mu\hat{J}_+ - \mu^*\hat{J}_- + \lambda\hat{J}_z)$, where $\hat{J}_\pm \equiv \hat{J}_x \pm i\hat{J}_y$, $\hat{J}_z$ are the infinitesimal generators of SU(2) and the complex parameter μ takes the value in $|\mu| \leqslant \pi/2$. We can put the real parameter λ to be equal to zero since the above form of SU(2) CS is only defined on the coset space SU(2)/U(1). This is easily seen from the fact that the element of the type $\exp(\chi\hat{J}_z)$, χ being real, form the invariant subgroup U(1) of SU(2) if $|0\rangle$ is chosen as the eigenvector of $\hat{J}_z$.

Now taking the application of SU(2) CS to a rigid rotator into consideration, we shall modify the su(2) algebra as follows: We introduce the signature σ which takes the value

a)Research Institute for Fundamental Physics, Kyoto University, Kyoto, Japan.
b)Present Address: Institut für Kernphysik Kernforschungsanlage Jülich, 5170 Jülich, West Germany.

 0022-2488/80/030472-05$1.00

± 1 and redefine the operators $\hat{J}_\pm$, $\hat{J}_z$ as satisfying

$$[\hat{J}_+, \hat{J}_-] = 2\sigma\hat{J}_z, \quad [\hat{J}_z, \hat{J}_\pm] = \pm\sigma\hat{J}_\pm. \tag{2.1}$$

In the case of rigid rotator, the angular momentum operators $\hat{I}_x = [(\hat{J}_+ + \hat{J}_-)/2]\hbar$, etc., will be defined with respect to the "body-fixed" frame, so that the sign of their commutator shoud be changed, i.e., $\sigma = -1$. Corresponding to the above definition, we choose the fixed vector $|0\rangle$ as

$$|0\rangle = |J, -\sigma J\rangle, \tag{2.2}$$

where $|J,M\rangle$ is the simultaneous eigenvector of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ with eigenvalues $J(J+1)$ and M, respectively. Note that the relation

$$\hat{J}_-|0\rangle = 0 \tag{2.3}$$

holds irrespective of the value of σ. Hence, the SU(2) CS can be written as

$$|\zeta\rangle \equiv U(\mu)|0\rangle = (1+|\zeta|^2)^{-J}\exp(\zeta\hat{J}_+)|0\rangle,$$
$$U(\mu) = \exp(\mu\hat{J}_+ - \mu^*\hat{J}_-). \tag{2.4}$$

The new parameter ζ takes an arbitrary value in the complex plane and is related to μ through $\zeta = (\mu/|\mu|)\tan|\mu|$. Another useful parametrization is given by stereographic projection of the complex ζ plane which reflects that the coset space SU(2)/U(1) is isomorphic to the two-dimensional sphere S^2. Thus we have

$$\mu = \tfrac{1}{2}\theta e^{-i\sigma\psi}, \quad \zeta = \tan(\tfrac{1}{2}\theta)\, e^{-i\sigma\psi}, \tag{2.5}$$

where the values of parameters are restricted as $0 \leqslant \theta \leqslant \pi$, $0 \leqslant \psi < 2\pi$, and σ is the signature in Eq. (2.1). In this case $U(\mu)$ produces a rotation through an angle θ about an axis $(\sigma \sin\psi, -\cos\psi, 0)$.

B. Completeness relation

The most important property of the CS $|\zeta\rangle$ is the completeness relation

$$\int |\zeta\rangle\, d\mu(\zeta)\langle\zeta| = 1, \tag{2.6}$$

where the invariant measure $d\mu(\zeta)$ on the coset space $SU(2)/U(1) \simeq S^2$ is given by

$$d\mu(\zeta) = \frac{2J+1}{\pi}\frac{d\,\mathrm{Re}\zeta\, d\,\mathrm{Im}\zeta}{(1+|\zeta|^2)^2}. \tag{2.7}$$

The right-hand side of the relation (2.6) denotes the unity in the representation space $\{|J,M\rangle; M = -J, -J+1, \ldots, J\}$ with fixed J. Although the relation (2.7) is usually derived directly from the definition (2.4) and the completeness relation $\Sigma_{M=-J}^{J}|J,M\rangle\langle J,M| = 1$, it can be shown that Eq. (2.6) proved also through the general group theoretical argument, especially with the aid of the invariance of the measure $d\mu(\zeta)$. The latter derivation of the completeness is also applicable to the generalized CS based on an arbitrary Lie group.

From the completeness relation we obtain the "reproducing relation"

$$f(\zeta) = \int K(\zeta,\zeta')f(\zeta')\, d\mu(\zeta'), \tag{2.8}$$

with $f(\zeta) \equiv \langle\zeta|f\rangle$, $|f\rangle$ being an arbitrary ket vector. The reproducing kernel K is defined by

$$K(\zeta,\zeta') \equiv \langle\zeta|\zeta'\rangle = (1+\zeta^*\zeta')^{2J}/\{(1+|\zeta|^2)(1+|\zeta'|^2)\}^J, \tag{2.9}$$

and plays a role of the δ function in Eq. (2.8). This function has the following property:

$$K(\zeta,\zeta') \leqslant 1, \tag{2.10}$$

where the equality holds only for $\zeta = \zeta'$. From Eq. (2.1) together with (2.4) and (2.9) we can easily derive formulas for the matrix elements of the product of $\hat{J}_\pm$, $\hat{J}_z$ between a couple of SU(2) CS. This will be given in the Appendix.

3. PATH INTEGRAL REPRESENTATION FOR THE TRANSITION AMPLITUDE

Let us consider a Hamiltonian $\hat{H}$ acting in our Hilbert space. We shall assume that $\hat{H}$ can be expanded as the finite polynomial of the infinitesimal operators $\hat{J}_\pm$ and $\hat{J}_z$ of SU(2). Since $\hat{H}$ conserves the quantum number J, we hereafter consider only the state vectors with fixed J. The transition amplitude (propagator) from the state $|\zeta\rangle$ at time t to the state $|\zeta'\rangle$ at time t' is given by

$$T(\zeta' t';\zeta t) = \langle\zeta'|\exp[-(i/\hbar)\hat{H}(t'-t)]|\zeta\rangle. \tag{3.1}$$

In order to derive the path integral form for the amplitude T, we divide $(t'-t)$ into n equal time intervals $\epsilon = (t'-t)/n$ and take the limit $n \to \infty$:

$$T = \lim_{\substack{n\to\infty \\ (\epsilon\to 0)}} \langle\zeta'|(1-(i/\hbar)\hat{H}\epsilon)^n|\zeta\rangle. \tag{3.2}$$

Inserting the completeness relation (2.6) in the space with fixed J into each time intervals of (3.2), we can rewrite T as

$$\begin{aligned} T &= \lim_{n\to\infty}\int\cdots\int\prod_{k=1}^{n-1} d\mu(\zeta_k) \\ &\quad\cdot\prod_{k=1}^{n}\left\langle\zeta_k\left|\left(1-\frac{i}{\hbar}\hat{H}\epsilon\right)\right|\zeta_{k-1}\right\rangle \\ &= \lim_{n\to\infty}\int\cdots\int\prod_{k=1}^{n-1} d\mu(\zeta_k)\cdot\prod_{k=1}^{n}\langle\zeta_k|\zeta_{k-1}\rangle \\ &\quad\cdot\prod_{k=1}^{n}\left\{1-\frac{i\epsilon}{\hbar}\frac{\langle\zeta_k|\hat{H}|\zeta_{k-1}\rangle}{\langle\zeta_k|\zeta_{k-1}\rangle}\right\}, \end{aligned} \tag{3.3}$$

where $\zeta_0 = \zeta$ and $\zeta_n = \zeta'$. As is performed in the path integral for the usual coherent state,[2] the integral (3.3) can be formally written in the following way: First the term in the curly bracket in (3.3) can be replaced by $\exp[-(i\epsilon/\hbar)\times\langle\zeta_k|\hat{H}|\zeta_{k-1}\rangle]$ in the limit $\epsilon \to 0$. Next, by using the identity $\langle\zeta_k|\zeta_{k-1}\rangle = \exp[\log\langle\zeta_k|\zeta_{k-1}\rangle]$ and the explicit form of the kernel function [i.e., Eq. (2.9)] the factor $\Pi_{k=1}^{n}\langle\zeta_k|\zeta_{k-1}\rangle$ is expressed as

$$\begin{aligned} \prod_{k=1}^{n}\langle\zeta_k|\zeta_{k-1}\rangle &= \exp\sum_{k=1}^{n}\epsilon\cdot\frac{1}{\epsilon}\log\langle\zeta_k|\zeta_{k-1}\rangle \\ &= \exp\sum_{k=1}^{n}\epsilon\left[\frac{J}{1+|\zeta_k|^2}\left(\zeta_k\frac{\Delta\zeta_k^*}{\epsilon}-\zeta_k^*\frac{\Delta\zeta_k}{\epsilon}\right)\right. \\ &\quad\left.+\frac{1}{\epsilon}O((\Delta\zeta_k)^2)\right] \\ &\to \exp\int_t^{t'}\frac{J}{1+|\zeta(t)|^2}\{\zeta(t)\dot{\zeta}^*(t)-\zeta^*(t)\dot{\zeta}(t)\}\,dt, \end{aligned} \tag{3.4}$$

where the dot denotes the time derivative. The integral over

the time in (3.4) is the formal expression in the limit $n \to \infty$. Now we obtain the final expression for the transition amplitude T:

$$T(\zeta' t', \zeta t) = \lim_{n\to\infty} \int\cdots\int \prod_{k=1}^{n-1} d\mu(\zeta_k) \times \exp \frac{i}{\hbar} \sum_{k=1}^{n} \epsilon \left[\frac{iJ\hbar}{1+|\zeta_k|^2} \left(\zeta_k^* \frac{\Delta\zeta_k}{\epsilon} - \zeta_k \frac{\Delta\zeta_k^*}{\epsilon} \right) - \langle \zeta_k | \hat{H} | \zeta_k \rangle \right]. \tag{3.5}$$

We may rewrite the expression (3.5) as the "formal" functional integral

$$T = \int d\mu[\zeta(t)] \exp[(i/\hbar)S], \tag{3.6}$$

$$S = \int_t^{t'} \mathscr{L}(\zeta(t), \dot{\zeta}(t), \zeta^*(t), \dot{\zeta}^*(t))\, dt,$$

where the "Lagrangian" $\mathscr{L}$ is given by

$$\mathscr{L} = i[J\hbar/(1+|\zeta|^2)](\zeta^*\dot{\zeta} - \dot{\zeta}^*\zeta) - \mathscr{H},$$
$$\mathscr{H} \equiv \langle \zeta | \hat{H} | \zeta \rangle, \tag{3.7}$$

which can be rewritten as

$$\mathscr{L} = \left\langle \zeta(t) \middle| \left(i\hbar \frac{\partial}{\partial t} - \hat{H} \right) \middle| \zeta(t) \right\rangle \tag{3.8}$$

with the aid of

$$\left\langle \zeta \middle| \frac{\partial}{\partial t} \middle| \zeta \right\rangle = \left\langle \zeta \middle| \left(\dot{\zeta} \frac{\partial}{\partial \zeta} + \dot{\zeta}^* \frac{\partial}{\partial \zeta^*} \right) \middle| \zeta \right\rangle = \frac{J}{1+|\zeta|^2} (\zeta^*\dot{\zeta} - \dot{\zeta}^*\zeta). \tag{3.9}$$

The path integral expression for the transition amplitude between arbitrary initial and final states is obtained from (3.5) by multiplying the wave functions of these states in the SU(2) CS representation and by integrating over the endpoint variables ζ and ζ'.

4. CLASSICAL EQUATION OF MOTION

In the case where $\hbar$ is extremely small compared with the action S in Eq. (3.5), the main contribution to the transition amplitude T comes from the path which makes the action stationary with fixed endpoint conditions $\zeta = \zeta(t)$, $\zeta' = \zeta(t')$:

$$0 = \delta S = \int_t^{t'} \left(\frac{\partial \mathscr{L}}{\partial \zeta} \delta\zeta + \frac{\partial \mathscr{L}}{\partial \dot{\zeta}} \delta\dot{\zeta} + \text{c.c.} \right) dt = \int_t^{t'} \left\{ \left[\frac{\partial \mathscr{L}}{\partial \zeta} - \frac{d}{dt} \left(\frac{\partial \mathscr{L}}{\partial \dot{\zeta}} \right) \right] \delta\zeta + \text{c.c.} \right\} dt. \tag{4.1}$$

As the variations $\delta\zeta$ and $\delta\zeta^*$ are independent and arbitrary, we obtain

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{\zeta}}\right) - \frac{\partial \mathscr{L}}{\partial \zeta} = 0, \quad \frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial \dot{\zeta}^*}\right) - \frac{\partial \mathscr{L}}{\partial \zeta^*} = 0. \tag{4.2}$$

Using the expression (3.7) for $\mathscr{L}$, we cast Eq. (4.2) into the "canonical" form as

$$\dot{\zeta} = -i \frac{(1+|\zeta|^2)^2}{2J\hbar} \frac{\partial \mathscr{H}}{\partial \zeta^*},$$
$$\dot{\zeta}^* = i \frac{(1+|\zeta|^2)^2}{2J\hbar} \frac{\partial \mathscr{H}}{\partial \zeta}. \tag{4.3}$$

Now we define the Poisson bracket by

$$\{A,B\} = i \frac{(1+|\zeta|^2)^2}{2J\hbar} \left(\frac{\partial A}{\partial \zeta^*} \frac{\partial B}{\partial \zeta} - \frac{\partial A}{\partial \zeta} \frac{\partial B}{\partial \zeta^*} \right) \tag{4.4}$$

for arbitrary function A and B of ζ and ζ^*. As is easily verified, this bracket satisfies the antisymmetricity and Jacobi identity. Then Eq. (4.3) is brought into the form

$$\dot{\zeta} = \{\zeta, \mathscr{H}\}, \quad \dot{\zeta}^* = \{\zeta^*, \mathscr{H}\}. \tag{4.3'}$$

Alternatively, using the angle variables (θ,ψ) through stereographic projection, we can rewrite Eq. (4.3) as

$$\dot{\theta} = \{\theta, \mathscr{H}\}, \quad \dot{\psi} = \{\psi, \mathscr{H}\}, \tag{4.5}$$

where the Poisson bracket (4.4) in the (θ,ψ) representation is given by

$$\{A,B\} = \frac{\sigma}{J\hbar \sin\theta} \left(\frac{\partial A}{\partial \psi} \frac{\partial B}{\partial \theta} - \frac{\partial A}{\partial \theta} \frac{\partial B}{\partial \psi} \right), \tag{4.6}$$

which coincides with the one given by Berezin[7] in his discussion on the quantization of classical dynamics. The set of Eqs. (4.5) is just suitable for the description of the classical spin (angular momentum) of magnitude J (see Sec. 5).

The equations of motion (4.3) may be regarded as a natural extension of the ordinary canonical equation of motion written in terms of complex variable

$$\dot{\zeta} = -\frac{i}{\hbar} \frac{\partial \mathscr{H}}{\partial \zeta^*}, \quad \dot{\zeta}^* = \frac{i}{\hbar} \frac{\partial \mathscr{H}}{\partial \zeta}, \tag{4.7}$$

which can also be obtained from the contraction of SU(2) into the Heisenberg–Weyl group. The only difference lies in the factor $2J(1+|\zeta|^2)^2$ in (4.3) which is related to the geometrical structure of phase space. In order to observe this, let us consider the time derivative term in the Lagranian (3.7), the corresponding differential form to which is given by

$$i(J\hbar/2)(1+|\zeta|^2)^{-1}(\zeta^* d\zeta - \zeta d\zeta^*). \tag{4.8}$$

By taking the exterior derivative we obtain

$$\Omega = iJ\hbar(1+|\zeta|^2)^{-2} d\zeta \wedge d\zeta^*, \tag{4.9}$$

where $\wedge$ denotes outer product. The expression (4.9) is well known as a fundamental two-form associated with the Kähler metric[8] $ds^2 = (1+|\zeta|^2)^{-2}\, d\zeta \cdot d\zeta^*$ of two-dimensional sphere ($\simeq$one-dimensional complex projective space). Thus Eqs. (4.3) may be regarded as a set of canonical equations of motion in a "curved phase space" (two-dimensional sphere), whereas the ordinary canonical equations (4.7) are viewed as the one in the flat-phase space since the corresponding two-form become $\Omega = i\, d\zeta \wedge d\zeta^*$ and hence the associated Kähler metric is constant.

The above generalization of the concept of phase space is supported by the existence of the Liouville theorem: The volume element of this generalized phase space is given by

$$\delta V = \rho \delta\zeta \wedge \delta\zeta^*, \quad \rho = \text{const}\cdot(1+|\zeta|^2)^{-2}. \tag{4.10}$$

The time derivative of the variation $\delta\zeta$ ($\delta\zeta^*$) of ζ (ζ^*) is calculated from (4.3):

$$\delta\dot{\xi} = \rho^{-1}\left[-\frac{i}{2J\hbar}\left(\frac{\partial^2\mathscr{H}}{\partial\xi\partial\xi^*}\delta\xi + \frac{\partial^2\mathscr{H}}{\partial\xi^{*2}}\delta\xi^*\right) - \dot{\xi}\left(\frac{\partial\rho}{\partial\xi}\delta\xi + \frac{\partial\rho}{\partial\xi^*}\delta\xi^*\right)\right] \tag{4.11}$$

together with the one for $\delta\xi^*$. As the system develops according to the equations of motion (4.3), the time derivative of δV becomes

$$\begin{aligned}\delta\dot{V} &= \dot{\rho}\delta\xi\wedge\xi^* + \rho\delta\dot{\xi}\wedge\delta\xi^* + \rho\delta\xi\wedge\delta\dot{\xi}^* \\ &= \left(\frac{\partial\rho}{\partial\xi}\dot{\xi}\delta\xi + \rho\delta\dot{\xi}\right)\wedge\delta\xi^* + \text{c.c.}\end{aligned} \tag{4.12}$$

which vanishes by virtue of Eq. (4.11), i.e., δV is conserved under the equation of motion. In other words, the phase space volume element is an integral invariant of the infinitesimal canonical transformation given by (4.3). More generally we can expect that the phase space volume element is invariant under an arbitrary finite canonical transformation which assures the invariance of the Poisson bracket (4.4).

Finally we make a remark that the generalization of phase space may not be restricted to the present simple case. If we encounter the manifold possessing with the Kähler metric we may get the corresponding "generalized canonical equation," the dynamical consequences of which will be investigated in a separate paper.[6]

5. EXAMPLE: RIGID ROTATOR

In this section we apply the results obtained in the preceding section to rigid rotator and discuss the correspondence between classical and quantum mechanical description. The Hamiltonian is given by

$$\hat{H} = \sum_{k=x,y,z}\frac{\hat{I}_k^2}{2\mathscr{J}_k}, \quad \hat{I}_k = \hat{J}_k\hbar, \tag{5.1}$$

where $\mathscr{J}_k$ and $\hat{I}_k$ denote the moment of inertia and the projection of angular momentum with respect to the body-fixed frame, respectively [i.e., $\sigma = -1$ in Eq. (2.1)]. The expectation value of $\hat{H}$ in SU(2) CS splits into classical and quantum-fluctuation parts as

$$\begin{aligned}&\mathscr{H} = \langle\xi|\hat{H}|\xi\rangle = \mathscr{H}_{\text{cl}} + \Delta\mathscr{H}, \\ &\mathscr{H}_{\text{cl}} = \sum_k\frac{\langle\hat{I}_k\rangle^2}{2\mathscr{J}_k}, \\ &\Delta\mathscr{H} = \sum_k\frac{(\Delta I_k)^2}{2\mathscr{J}_k}, \quad \Delta I_k = [\langle\hat{I}_k^2\rangle - \langle\hat{I}_k\rangle^2]^{1/2}.\end{aligned} \tag{5.2}$$

The fluctuation part ΔI_k is of order J^{-1} compared with $\langle\hat{I}_k\rangle$, see the Appendix. Let us consider the "classical" situation which holds the approximation developed in Sec. 4. Since the magnitude of angular momentum $J\hbar$ should be large compared with $\hbar$ in this case, the fluctuation Hamiltonian $\Delta\mathscr{H}$, being the quantity of $O(1/J) = O(\hbar/J\hbar)$, can be neglected. The expectation value of $\hat{I}_k$ in the (θ,ψ) representation is given by

$$\begin{aligned}&\langle\hat{I}_x\rangle = J\hbar\sin\theta\cos\psi, \\ &\langle\hat{I}_y\rangle = \sigma J\hbar\sin\theta\sin\psi, \\ &\langle\hat{I}_z\rangle = -\sigma J\hbar\cos\theta \quad (\sigma = -1),\end{aligned} \tag{5.3}$$

with the aid of which we rewrite $\mathscr{H}_{\text{cl}}$ as

$$\mathscr{H}_{\text{cl}} = J^2\hbar^2\left(\frac{1}{2\mathscr{J}_x}\sin^2\theta\cos^2\psi + \frac{1}{2\mathscr{J}_y}\sin^2\theta\sin^2\psi + \frac{1}{2\mathscr{J}_z}\cos^2\theta\right). \tag{5.4}$$

Substituting this into the classical equation of motion (4.5) we obtain

$$\begin{aligned}\dot{\psi} &= -J\hbar\left(\frac{1}{\mathscr{J}_x}\cos^2\psi + \frac{1}{\mathscr{J}_y}\sin^2\psi - \frac{1}{\mathscr{J}_z}\right)\cos\theta, \\ \dot{\theta} &= -J\hbar\left(\frac{1}{\mathscr{J}_x} - \frac{1}{\mathscr{J}_y}\right)\sin\theta\sin\psi\cos\psi.\end{aligned} \tag{5.5}$$

This equation is cast into the more familiar form by introducing the angular velocity ω_k as follows:

$$\omega_k = \langle\hat{I}_k\rangle/\mathscr{J}_k \quad (k = x,y,z). \tag{5.6}$$

Thus the time derivative of ω_k is given by

$$\begin{aligned}\mathscr{J}_x\dot{\omega}_x &= (\mathscr{J}_y - \mathscr{J}_z)\omega_y\omega_z, \\ \mathscr{J}_y\dot{\omega}_y &= (\mathscr{J}_z - \mathscr{J}_x)\omega_z\omega_x, \\ \mathscr{J}_z\dot{\omega}_z &= (\mathscr{J}_x - \mathscr{J}_y)\omega_x\omega_y,\end{aligned} \tag{5.7}$$

which coincide with the Euler equations for a rigid body. Equations (5.7) as well as Eqs. (5.5) can be regarded as the set of canonical equations in the generalized phase space.

Next let us investigate the correspondence between the classical equation of motion obtained above and the quantum mechanical one. The quantum analogue of Eqs. (5.7) is directly derived by defining the quantum mechanical "angular velocity" as

$$\hat{\omega}_k = \hat{I}_k/\mathscr{J}_k \tag{5.8}$$

and then using the Heisenberg equations of motion:

$$\begin{aligned}\mathscr{J}_x\dot{\hat{\omega}}_x &= (1/i\hbar)[\mathscr{J}_x\hat{\omega}_x, \hat{H}] \\ &= (\mathscr{J}_y - \mathscr{J}_z)\cdot\tfrac{1}{2}(\hat{\omega}_y\hat{\omega}_z + \hat{\omega}_z\hat{\omega}_y),\end{aligned} \tag{5.9}$$

and cyclic permutations. On the other hand, if we wish to obtain these equations via the "quantization" of the classical relations, we should determine the commutation relations between angular velocities (momenta). This is carried out through the Poisson bracket defined in Sec. 4. In fact we can derive from Eqs. (4.6) and (5.3) with $\sigma = -1$

$$\{\langle\hat{I}_x\rangle,\langle\hat{I}_y\rangle\} = -\langle\hat{I}_z\rangle, \tag{5.10}$$

and so on. In order that Eq. (5.10) be quantized into the commutation relation of the angular momentum with $\sigma = -1$, it requires the following correspondence for the classical quantities A, B and the quantized ones $\hat{A}$, $\hat{B}$:

$$\{A,B\}\longleftrightarrow(1/i\hbar)[\hat{A},\hat{B}]. \tag{5.11}$$

Note that the Poisson bracket on the left-hand side does not depend on the representation (choice of variables).

Finally we discuss the quantum analogue of the angle (θ,ψ) as dynamical variables. Let us introduce the operators $\hat{\theta}$ and $\hat{\psi}$ with the aid of $\hat{J}_\pm$ and $\hat{J}_z$ as follows:

$$\cos\hat{\theta} = \hat{J}_z/J, \tag{5.12a}$$

$$\exp i\sigma\hat{\psi} = (J+\sigma\hat{J}_z)^{-1/2}\hat{J}_+(J-\sigma\hat{J}_z)^{-1/2}, \tag{5.12b}$$

where $\sigma = -1$ in our case and eigenvalues of $\hat{\theta}$ and $\hat{\psi}$ are restricted in the same region as θ and ψ, respectively. Except

for the cases where the square root in (5.12b) vanishes the operator $\exp i\hat{\psi}$ is unitary and satisfies

$$[\hat{J}_z, \exp i\hat{\psi}] = \exp i\hat{\psi}, \tag{5.13}$$

so that the operator $\hat{\psi}$ can be regarded as the conjugate angle to $\hat{J}_z$. Since the magnitude of angular momentum is commutable with $\hat{J}_\pm$ and $\hat{J}_z$, we can replace J in Eqs. (5.12) with the operator $\hat{J}$ which is defined through $\hat{J}(\hat{J}+1) = \hat{\mathbf{J}}^2$. Thus we obtain the following "polar" coordinate expression for the shift operators:

$$\begin{aligned} \hat{J}_z &= \hat{J}\cos\hat{\theta}, \\ \hat{J}_+ &= 2\hat{J}\cos(\hat{\theta}/2)\cdot\exp i\sigma\hat{\psi}\cdot\cos(\hat{\theta}/2), \\ \hat{J}_- &= 2\hat{J}\cos(\hat{\theta}/2)\cdot\exp(-i\sigma\hat{\psi})\sin(\hat{\theta}/2). \end{aligned} \tag{5.14}$$

It is easy to check the commutation relation (2.1) starting with the relations

$$[\cos\hat{\theta},\hat{\psi}] = -i/\hat{J}, \quad [\hat{J},\hat{\psi}] = [\hat{J},\hat{\theta}] = 0. \tag{5.15}$$

It can also be shown that the expectation values of $\hat{\theta}$ and $\hat{\psi}$ in SU(2) CS coincide with the parameters θ and ψ, respectively, except for the additional terms which vanish in the classical limit.

ACKNOWLEDGMENT

One of the authors (T.S) would like to thank Dr. T. Kashiwa for valuable discussions. He is also indebted to the Yukawa Foundation for the financial aid.

APPENDIX

Here we derive the formula for the matrix elements of the infinitesimal operators $\hat{J}_\pm$, $\hat{J}_z$ between SU(2) CS.

First differentiating the expression (2.4) for SU(2) CS we obtain

$$\begin{aligned} &\frac{d}{d\zeta_2}\{(1+|\zeta_2|^2)^J\langle\zeta_1|\zeta_2\rangle\} \\ &\quad = \left\langle\zeta_1\left|\frac{d}{d\zeta_2}e^{\zeta_2\hat{J}_+}\right|0\right\rangle \\ &\quad = (1+|\zeta_2|^2)^J\langle\zeta_1|\hat{J}_+|\zeta_2\rangle, \end{aligned}$$

then substituting the formula (2.9) we have

$$\langle\zeta_1|\hat{J}_+|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle = 2J\zeta_1^*/(1+\zeta_1^*\zeta_2), \tag{A1}$$

and also

$$\langle\zeta_1|\hat{J}_-|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle = 2J\zeta_2/(1+\zeta_1^*\zeta_2). \tag{A2}$$

In order to derive the matrix element of $\hat{J}_z$ we utilize the formula

$$e^{-\zeta\hat{J}_+}\hat{J}_z e^{\zeta\hat{J}_+} = \hat{J}_z + \sigma\zeta\hat{J}_+$$

from which follows

$$\begin{aligned} &\langle\zeta_1|\hat{J}_z|\zeta_2\rangle(1+|\zeta_2|^2)^{-J}\langle\zeta_1|e^{\zeta_2\hat{J}_+}(\hat{J}_z+\sigma\zeta_2\hat{J}_+)|0\rangle \\ &\qquad = -\sigma J\langle\zeta_1|\zeta_2\rangle + \sigma\zeta_2\langle\zeta_1|\hat{J}_+|\zeta_2\rangle. \end{aligned}$$

This gives

$$\langle\zeta_1|\hat{J}_z|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle = -\sigma J(1-\zeta_1^*\zeta_2)/(1+\zeta_1^*\zeta_2), \tag{A3}$$

where we used the expression (A1).

In a similar manner we can evaluate the matrix elements of operators which is higher power of $\hat{J}_\pm$, $\hat{J}_z$. Here we only list up the ones for the quadratic case which is used in Sec. 5:

$$\begin{aligned} &\langle\zeta_1|\hat{J}_+^2|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle = 2J(2J-1)[\zeta_1^*/(1+\zeta_1^*\zeta_2)]^2, \\ &\langle\zeta_1|\hat{J}_-^2|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle = 2J(2J-1)[\zeta_2/(1+\zeta_1^*\zeta_2)]^2, \\ &\langle\zeta_1|\hat{J}_+\hat{J}_-|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle \\ &\qquad = [4J^2\zeta_1^*\zeta_2 + 2J(\zeta_1^*\zeta_2)^2]/(1+\zeta_1^*\zeta_2)^2, \\ &\langle\zeta_1|\hat{J}_-\hat{J}_+|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle \\ &\qquad = [4J^2\zeta_1^*\zeta_2 + 2J]/(1+\zeta_1^*\zeta_2)^2, \\ &\langle\zeta_1|\hat{J}_z^2|\zeta_2\rangle/\langle\zeta_1|\zeta_2\rangle \\ &\qquad = [J^2(1-\zeta_1^*\zeta_2)^2 + 2J\zeta_1^*\zeta_2]/(1+\zeta_1^*\zeta_2)^2. \end{aligned}$$

[1] R.P. Feynman, Phys. Rev. **84**, 108 (1951); W. Tobocman, Nuovo Cimento **3**, 1213 (1956); C. Garrod, Rev. Mod. Phys. **38**, 483 (1966).
[2] J.R. Klauder, Ann. Phys. **11**, 123 (1960); S.S. Schweber, J. Math Phys. **3**, 831 (1962). Extension of "fermion coherent state" has been done with the aid of Grassmann variables by Y. Ohnuki and T. Kashiwa, Prog. Theor. Phys. **60**, 548 (1978); see also J.R. Klauder, *op. cit.*; F.A. Berezin and M.S. Marinov, Ann. Phys. **104**, 336 (1977).
[3] A.M. Perelomov, Commun. Math. Phys. **26**, 222 (1972). For the review articles, see A.M. Perelomov, Usp. Fiz. Nauk. **123**, 23 (1977) [Sov. Phys. Uspekhi **20**, 703 (1977); H. Bacry, A. Grossmann, and J. Zak, in *Lecture Notes in Physics* (Springer, Berlin 1975), Vol. 50, p. 249.
[4] J.M. Radcliffe; J. Phys. A **4**, 313 (1971); F.T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, Phys. Rev. A **6**, 2211 (1972).
[5] Very recently, Klauder has also investigated the path integral for the spin coherent state; J.R. Klauder, Phys. Rev. D **19**, 2349 (1979). He deals with it in the angle variable or spin variable form, while we treat it in the complex variable form and this seems to reveal directly the canonical structure of classical equation of motion. The authors thank Dr. J.R. Klauder for his sending of his work before publication.
[6] H. Kuratsuji and T. Suzuki (to be published).
[7] F.A. Berezin, Commun. Math. Phys. **40**, 153 (1975).
[8] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry* (Wiley, New York, 1969), Vol. 2.

Is the Feynman–Dyson series adequate for the aysmptotic expansion of the Green's function of the quantum mechanical anharmonic oscillator?

Wai Bong Yeung
Department of Physics, Queen Mary College, Mile End Rd., London E1 4NS, England
and Department of Physics, National Central University, Chung-Li, Taiwan 320, Republic of China

(Received 11 January 1979)

We discuss the asymptotic expansion of the Green's function of the quantum mechanical anharmonic oscillator by applying the method of steepest descents in the evaluation of the corresponding Euclidean path integrals. The contributions arising from the pseudoparticle solutions are then shown to be zero.

I. INTRODUCTION

It was three years ago that an exact classical solution to the SU(2) Yang–Mills field equations in four-dimensional Euclidean space with finite action, called the pseudoparticle solution, was found by Belavin, Polyakov, Schwartz, and Tyupkin.[1] Since then much attention has been paid in finding the proper role played by this classical Euclidean solution in quantum theory. Polyakov suggested that this solution be used to dominate the path integrals in the discussion fo quark confinement in Yang–Mills theory.[2] Then, 't Hooft found that it has something to do with the chiral symmetry breakdown observed in nature;[3] and later Jackiw and Rebbi,[4] and independently Callan, Dashen, and Gross,[5] have proposed that the pseudoparticle solution gives rise to tunneling between gauge inequivalent classical vacua of the theory.

In this paper we would like to investigate the influence of the classical Euclidean solution on the asymptotic expansion of the Green's functions which appear in quantum mechanics as the ground state expectation value of the time ordered products of the coordinate operators. We use the anharmonic oscillator as an example for demonstration because it possesses classical Euclidean solutions of finite actions.

We shall organize our paper in four sections. In Sec. II we use a simple example to illustrate the method of steepest descents[6] in obtaining the asymptotic expansion of an ordinary definite integrals. Section III discusses the case of the anharmonic oscillator in quantum mechanics: We first show the reasons why we need real Euclidean path integrals in the discussions of the Green's functions, and then we incorporate the classical Euclidean solutions into our asymptotic expansion by analytically continuing the real Euclidean path integrals into the complex functional space. The combination of the contributions from an infinite number of "loosely packed" classical solutions together with the contribution from the trivial solution will be calculated. Section IV gives some conclusions of what we have done. There is also an Appendix at the end of the paper.

II. A SIMPLE EXAMPLE

We start our discussion by looking at the evaluation of the ordinary definite integral

$$I = \int_{-\infty}^{+\infty} e^{-x^2 - gx^4}\, dx.$$

Unfortunately, this integral cannot be expressed in terms of known elementary functions. In the absence of an exact solution, we would like to make an expansion in a power series of g and hope that such an expansion is at least asymptotic.[7]

One of the most powerful methods in obtaining the asymptotic expansion of a definite integral is the method of steepest descents.[8] In applying the method of steepest descents, we analytically continue the integrating variable x into the complex plane z, and because the integrand has singularities only at some points in the infinity away from the real axis, we have

$$I = \int_C dz\, e^{-z^2 - gz^4}, \tag{2.1}$$

where C is an arbitrary contour beginning and ending at the end points of the real axis.

The saddle points of $f(z) = -z^2 - gz^4$ occur at

$$z^0 = 0, \quad z^+ = +i/\sqrt{2g}, \quad z^- = -i/\sqrt{2g}.$$

Notice that two of the saddle points are located on the imaginary axis, and they cannot be counted if we are looking for real solutions only. Near the saddle point z^0,

$$f(z) \approx -z^2; \tag{2.2}$$

near the saddle point z^+,

$$f(z) \approx 1/4g + 2(z - z^+)^2; \tag{2.3}$$

and near the saddle point z^-,

$$f(z) \approx 1/4g + 2(z - z^-)^2. \tag{2.4}$$

The idea of the method of steepest descents is to deform the contour C to pass through the saddle points in such directions that $f(z)$ is positive definite near the saddle points. The whole integral is approximated by the integrations near the saddle points because the contributions elsewhere will be cancelled away due to the rapid oscillations of the exponential. Because of the positivity of the coefficients of the quadratic terms in Eqs. (2.3) and (2.4), the obvious choices of contours around z^+ and z^- are

$$(z - z^+) = \pm i\rho, \quad (z - z^-) = \pm i\sigma, \tag{2.5}$$

respectively, where ρ and σ are real.

 0022-2488/80/030477-09$1.00

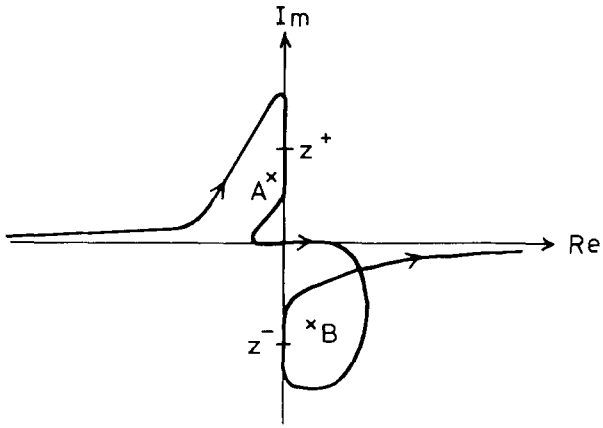

FIG. 1.

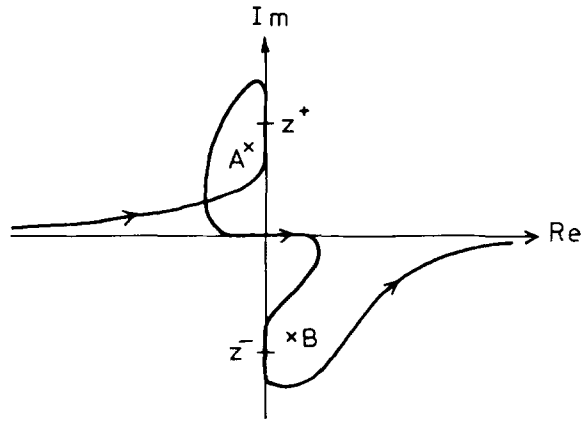

FIG. 2.

The choice of the signs in Eqs. (2.5) depend on the shape of the "mountain range" we are passing through. However, $(z - z^+)$ and $(z - z^-)$ must choose opposite signs. The reason is the following: If there is a "ridge" near A (see Fig. 1), there will also be a ridge near B because $f(z)$ is unchanged in $z \leftrightarrow -z$; then the right contour should be C_1. On the contrary, if there are "valleys" near A and B, the right contour should be C_2 (see Fig. 2). In both cases, $(z - z^+)$ and $(z - z^-)$ should choose opposite signs.

Then we have[9]

$$I = \int_{x\approx 0} dx\, e^{-x^2 - gx^4} \pm i \int_{\rho\approx 0} d\rho\, e^{1/4g - 2\rho^2} \mp i \int_{\sigma\approx 0} d\sigma\, e^{1/4g - 2\sigma^2}. \tag{2.6}$$

The last two terms cancel each other, and hence

$$\begin{aligned} I &= \int_{x\approx 0} dx\, e^{-x^2 - gx^4} \\ &= \int_{x\approx 0} dx \sum_n \frac{(-gx)^n}{n!} e^{-x^2} \\ &= \int_{-\infty}^{+\infty} dx \sum_n \frac{(-gx^4)^n}{n!} e^{-x^2}. \end{aligned} \tag{2.7}$$

The integral I is, in fact, equal to

$$\tfrac{1}{2}(1/g)^{1/2} \exp(1/8g) K_{1/4}(1/8g)$$

with $K_{1/4}(1/8g)$ expressible as an asymptotic series. If we expand $K_{1/4}(1/8g)$, we will see that it is indeed equal to Eq. (2.7) integrated term by term.

In the following, we will turn to quantum mechanics and see that the calculations in Eq. (2.6) will no longer hold.

III. THE ANHARMONIC OSCILLATOR

Let us now consider the quantum mechanical problem of the anharmonic oscillator whose Lagrangian is given by

$$\mathscr{L} = \tfrac{1}{2}\dot{\hat{x}}^2 - \tfrac{1}{2}\hat{x}^2 - (g/4)\hat{x}^4. \tag{3.1}$$

The coupling constant g is chosen to be positive.

The physical quantities that are always interesting to us are the Green's functions $G_n(t_1,\ldots,t_n)$ which are the ground state expectation values of the time-ordered products $T[\hat{x}(t_1)\cdots\hat{x}(t_n)]$. Unfortunately, these ground state expectation values cannot be represented as Feynman path integrals. What can be represented as Feynman path integrals are the expectation values of $T[\hat{x}(t_1)\cdots\hat{x}(t_n)]$ between two localized states.[10] For example, we have

$$\langle X'T' | T[\hat{x}(t_1)\hat{x}(t_2)] | XT\rangle = \int_{x(T)=X,\ x(T')=X'} \mathscr{D}[x]\, x(t_1)x(t_2) \exp\left(i\int_T^{T'} \mathscr{L}\, dt\right), \tag{3.2}$$

provided that $T \leqslant t_1, t_2 \leqslant T'$. The integration is supposed to be done over all real and continuous functions $x(t)$ satisfying $x(T) = X$, $x(T') = X'$. The above path integral is, in fact, a compact notation for the multiple integral

$$\lim_{n\to\infty} \int_{-\infty}^{+\infty} \Pi_j \frac{dx_j}{(2\pi i\epsilon)^{1/2}}\, x_i x_k \exp\left[i\Sigma\epsilon\mathscr{L}\left(\frac{x_i + x_{i-1}}{2}, \frac{x_i - x_{i-1}}{\epsilon}\right)\right],$$

where we have divided the interval between T and T' into $(n+1)$ subintervals of length $\epsilon = (T' - T)/(n+1)$ with t_1 and t_2 as two of the end points (points i and k) of the subintervals; x_i is the value of $x(t)$ at the point i.

To calculate the ground state expectation value of $T[\hat{x}(t_1)\hat{x}(t_2)]$, say, we insert complete sets of energy eigenstates $\{|n\rangle\}$ into the lhs of Eq. (3.2)[11]:

$$\begin{aligned} &\sum_{nm} \Phi_n(X')\Phi_m^*(X) e^{-iE_nT'} e^{iE_mT} \langle n | T[\hat{x}(t_1)\hat{x}(t_2)] | m\rangle \\ &= \lim_{n\to\infty} \int_{-\infty}^{+\infty} \Pi \frac{dx_j}{(2\pi i\epsilon)^{1/2}}\, x_i x_k \exp\left[i\Sigma\epsilon\mathscr{L}\left(\frac{x_i + x_{i-1}}{2}, \frac{x_i - x_{i-1}}{\epsilon}\right)\right], \end{aligned} \tag{3.3}$$

Φ_n here means the wave function of the nth excited state.

If we rotate the time variable into the imaginary axis by making $t_1 = -i\tau_1, t_2 = -i\tau_2, T = -i\mathscr{T}$, and $T' = -i\mathscr{T}'$, we will have

$$\sum_{nm} \Phi_n(X')\Phi_m^*(X)\, e^{-E_n\mathscr{T}' + E_m\mathscr{T}} \langle n|\, T[\hat{x}(\tau_1)\hat{x}(\tau_2)]\,|m\rangle$$

$$= \lim_{n\to\infty} \int_{-\infty}^{+\infty} \Pi \frac{dx_j}{(2\pi\epsilon')^{1/2}} x_i x_k \exp\left[\Sigma\epsilon' \mathscr{L}\left(\frac{x_i + x_{i-1}}{2}, \frac{x_i - x_{i-1}}{-i\epsilon'}\right)\right], \tag{3.4}$$

with $\epsilon' = (\mathscr{T}' - \mathscr{T})/(n+1)$. The last multiple integral is just the Euclidean path integral

$$\int_{x(\mathscr{T}')=X',\ x(\mathscr{T})=X} \mathscr{D}[x] x(\tau_1)x(\tau_2) \exp\left(+\int_{\mathscr{T}}^{\mathscr{T}'} d\tau \mathscr{L}^E\right),$$

integrating over all real functions $x(\tau)$ satisfying $x(\mathscr{T}') = X', x(\mathscr{T}) = X$. The weighting factor $\mathscr{L}^E$ is

$$-\left[\left(\frac{dx}{d\tau}\right)^2 + \frac{x^2}{2} + \frac{g}{4}x^4\right].$$

By letting $\mathscr{T} \to -\infty, \mathscr{T}' \to +\infty$, we pick up the ground state contribution only, and hence

$$\lim_{\substack{\mathscr{T}'\to\infty \\ \mathscr{T}\to-\infty}} \Phi_0(X')\Phi_0^*(X)\, e^{-E_0(\mathscr{T}'-\mathscr{T})} \langle 0|\, T[\hat{x}(\tau_1)\hat{x}(\tau_2)]\,|0\rangle = \int_{\substack{x(\infty)=X' \\ x(-\infty)=X}} \mathscr{D}[x] x(\tau_1)x(\tau_2) \exp\left(+\int_{-\infty}^{+\infty} \mathscr{L}^E d\tau\right). \tag{3.5}$$

Using similar arguments we can show that

$$\lim_{\substack{\mathscr{T}'\to\infty \\ \mathscr{T}\to-\infty}} \Phi_0(X')\Phi^*(X)\, e^{-E_0(\mathscr{T}'-\mathscr{T})} = \int \mathscr{D}[x] \exp\left(+\int_{-\infty}^{+\infty} \mathscr{L}^E d\tau\right). \tag{3.6}$$

Combining the above two equations, we obtain the Euclidean path integral representation of $\langle 0|\, T[\hat{x}(\tau_1)\hat{x}(\tau_2)]\,|0\rangle$ as

$$\int_{x(-\infty)=X,\ x(+\infty)=X'} \mathscr{D}[x] x(\tau_1)x(\tau_2) \exp\left(+\int_{-\infty}^{+\infty} d\tau\, \mathscr{L}^E\right) \Big/ \int_{x(-\infty)=X,\ x(+\infty)=X'} \mathscr{D}[x] \exp\left(+\int_{-\infty}^{+\infty} d\tau\, \mathscr{L}^E\right). \tag{3.7}$$

Notice that all the functions integrated must be real and satisfy the arbitrary but *fixed* boundary conditions.

In the above definition of the Euclidean path integral as a multiple integral, the integrating variables x_i are regarded as real and they range from $-\infty$ to $+\infty$. However, we can analytically continue each x_i into the complex plane without changing the value of the multiple integral provided that the new integrating variable z_i follows a contour which begins and ends at the negative and positive end points of the real axis, respectively.

There is an equivalent way of defining the path integral, namely, by choosing a complete set of normalized functions $\{\varphi_n(\tau)\}$ satisfying the required boundary conditions, and by expanding the path $x(\tau)$ as

$$x(\tau) = \sum_n \xi_n \varphi_n(\tau). \tag{3.8}$$

The measure $\mathscr{D}[x]$ of the Euclidean path integral can then be cast into the form $N[\pi_n\, d\xi_n]^{12}$ with a normalization constant N. In the calculation of the Green's functions, the constant N will be cancelled between the Euclidean path integrals in the numerator and denominator of Eq. (3.7), and hence we will drop the normalization constant in our subsequent discussion. The merit of the analytic continuation of $x(\tau)$ into the complex plane is that we will have more freedom in choosing the ξ's when $x(\tau)$ is complex than when it is real. That kind of freedom is crucial for the application of the method of steepest descents. The different choices of the ξ's will correspond to the different contours taken by the z's while giving the same integral.

Armed with all these, we can start calculating $\langle 0|\, T[\hat{x}(\tau_1)\hat{x}(\tau_2)]\,|0\rangle$ by evaluating the path integrals using the method of steepest descents. The first step is to solve the classical equation of motion

$$\ddot{x}_c = x_c + g x_c^3 \tag{3.9}$$

with boundary conditions $x_c(-\infty) = X, x_c(+\infty) = X'$. We choose these boundary conditions for x_c because we want the small oscillation $(x - x_0)$ to vanish at the infinities.

As we have said before, the result for $\langle 0|\, T[\hat{x}(\tau_1)\hat{x}(\tau_2)]\,|0\rangle$ is independent of our choice of X and X' as far as they are fixed at the same values for both the denominator and the numerator of Eq. (3.7). However, we must choose X and X' so that the classical equation of motion has solutions; and furthermore, these solutions must have finite Euclidean action $\int_{-\infty}^{+\infty} \mathscr{L}^E[x_c]\, d\tau$. A convenient choice is $X = X' = 0$; and we will stick with this choice in the subsequent discussion. With these boundary conditions, there are two kinds of solutions for the classical equation of motion with finite Euclidean action; they are $x_c^0 = 0$ and $x_c^{\pm}(\tau - a) = \pm i(2/g)^{1/2}[(1)/\cosh(\tau - a)]$, with a as an arbitrary constant.

If we expand $\int_{-\infty}^{+\infty} \mathscr{L}^E d\tau$ about $x_c(\tau)$, the result will look like the following:

$$x_c^0: \quad \int_{-\infty}^{+\infty} \mathscr{L}^E d\tau = +0 - \frac{1}{2}\int_{-\infty}^{+\infty} y_0(\tau)\left(\frac{d^2}{d\tau^2}+1\right)y_0(\tau)\,d\tau - \frac{g}{4}\int_{-\infty}^{+\infty} y_0^4(\tau)\,d\tau; \tag{3.10}$$

$$x_c^{\pm}: \quad \int_{-\infty}^{+\infty} \mathscr{L}^E d\tau = \frac{4}{3}\frac{1}{g} - \frac{1}{2}\int_{-\infty}^{+\infty} y_{\pm}(\tau)\left(-\frac{d^2}{d\tau^2}+1-\frac{6}{\cosh^2(-a)}\right)y_{\pm}(\tau)\,d\tau$$
$$-g\int_{-\infty}^{+\infty} \pm i\left(\frac{2}{g}\right)^{1/2}\frac{1}{\cosh(\tau-a)}y_{\pm}^3(\tau)\,d\tau - \frac{g}{4}\int_{-\infty}^{+\infty} y_{\pm}^4(\tau)\,d\tau, \tag{3.11}$$

where we have used the notation that $y^0(\tau) = x(\tau) - x_c^0(\tau)$, $y^{\pm}(\tau) = x(\tau) - x_c^{\pm}(\tau)$.

As a rule, we always represent the small oscillation $y(\tau)$ as the linear combinations of some complete set of functions satisfying the required conditions. For $y^0(\tau)$, we choose the set $\{\theta_n(\tau)\}$ satisfying

$$\left\{-\frac{d^2}{d\tau^2}+1\right\}\theta_n(\tau) = \omega_n^2\theta_n(\tau), \quad \theta_n(-\infty) = \theta_n(+\infty) = 0, \tag{3.12}$$

and for $y^{\pm}(\tau)$, we choose the set $\{\varphi_n(\tau)\}$ satisfying

$$\left\{-\frac{d^2}{d\tau^2}+1-\frac{6}{\cosh^2(\tau-a)}\right\}\varphi_n(\tau-a) = \Omega_n^2\varphi_n(\tau-a), \quad \varphi_n(-\infty) = \varphi_n(+\infty) = 0. \tag{3.13}$$

The reasons for these choices will become obvious later. The solutions for the first eigenvalue problem are trivial:

$$\theta_k(\tau) = e^{-i(k-i\epsilon)\tau} \quad \text{and} \quad \omega_k^2 = 1 + k^2 - i\epsilon,$$

with ϵ infinitesimal. The solutions for the second eigenvalue problem are also known[13]: *There is one bound state solution*

$$\varphi_{-1} = \frac{1}{(e^{(\tau-a)}+e^{-(\tau-a)})^2}F\left(0,5\,|\,3\,|\,\frac{1}{e^{2(\tau-a)}-1}\right),$$

with $\Omega_{-1}^2 = -3$; there are two zero frequency solutions $\varphi_0^{\pm} = (d/da)x^{\pm}(\tau-a)$ with $\Omega_0^2 = 0$; and there is a continuum of solutions

$$\varphi_k = e^{-i(k-i\epsilon)(\tau-a)}F(-2,3\,|\,1-ik\,|\,[(1)/1-e^{-2(\tau-a)}], \quad \text{with} \quad \Omega_k^2 = 1+k^2-i\epsilon.$$

The functions F are the hypergeometric functions multiplied by a constant so as to make the eigensolutions properly normalized. Notice that ω_k^2 is the same as Ω_k^2. In terms of these eigenfunctions, $y^0(\tau)$, $y^+(\tau)$, and $y^-(\tau)$ can be written as

$$y_0(\tau) = \sum \xi_n\theta_n(\tau), \quad y^+(\tau) = \sum \eta_n\varphi_n(\tau), \quad y^-(\tau) = \sum \rho_n\varphi_n(\tau). \tag{3.14}$$

Plugging Eq. (3.14) into Eqs. (3.10) and (3.11) and then into Eq. (3.7), we can write the measure for $\langle 0\,|\,T[\hat{x}(\tau_1)\hat{x}(\tau_2)]\,|\,0\rangle$ in terms of ξ, η, and ρ, if we forget about the translational invariance of the theory.

It is by now well known that the integration of paths around $x_c^{\pm}(\tau-a)$ will spoil the translational invariance of the theory, and this can be repaired by integrating over the collective coordinate a. Following the usual tricks,[14] we insert

$$\Delta^{\pm}[x]\int_{-\infty}^{+\infty} da\,\delta\left[\int_{\tau\simeq a} d\tau\varphi^{\pm}(\tau-a)x(\tau)\right] = 1$$

into the respective functional integral which integrate about $x_c^{\pm}$. After making a time translation, the numerator of Eq. (3.7) can be written as (leaving out the cubic and quartic terms temporarily)

$$\int\left(\prod_i d\xi_i\right)\left[\sum_n \xi_n\theta_n(\tau_1)\right]\left[\sum_m \xi_m\theta_m(\tau_2)\right]\exp\left(-\tfrac{1}{2}\sum_k \omega_k^2\xi_k\xi_{-k}\right)$$
$$+\int\left(\prod_{i\neq 0} d\eta_i\right)\left[\Delta^+\int_{-\infty}^{+\infty} da\,x_c^+(\tau_1-a)x_c^+(\tau_2-a)\right]\exp\left(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\eta_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\eta_k\eta_{-k}\right)$$
$$+\int\left(\prod_{i\neq 0} d\rho_i\right)\left[\Delta^-\int_{-\infty}^{+\infty} da\,x_c^-(\tau_1-a)x_c^-(\tau_2-a)\right]\exp\left(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\rho_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\rho_k\rho_{-k}\right), \tag{3.15}$$

where we have suppressed the infinitestimal ϵ so as to retain the orthonormality of the eigenfunctions. The denominator reads as

$$\int\left(\prod_i d\xi_i\right)\exp\left(-\tfrac{1}{2}\sum_k\omega_k^2\xi_k\xi_{-k}\right)+\int\left(\prod_{i\neq 0} d\eta_i\right)\left(\Delta^+\int_{-\infty}^{+\infty} da\right)\exp\left(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\eta_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\eta_k\eta_{-k}\right)$$
$$+\int\left(\prod_{i\neq 0} d\rho_i\right)\left(\Delta^-\int_{-\infty}^{+\infty} da\right)\exp\left(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\rho_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\rho_k\rho_{-k}\right). \tag{3.16}$$

The Jacobian $\Delta^{\pm}$ can be calculated to be

$$\Delta^{+} = \left| \int_{-\infty}^{+\infty} \left(\dot{x}_c^{+} + \eta_{-1}\dot{\varphi}_{-1} + \sum_k \eta_k \dot{\varphi}_k \right) \dot{x}_c^{+} \, d\tau \right|, \tag{3.17}$$

$$\Delta^{-} = \left| \int_{-\infty}^{+\infty} \left(\dot{x}_c^{-} + \rho_{-1}\dot{\varphi}_{-1} + \sum_k \rho_k \dot{\varphi}_k \right) \dot{x}_c^{-} \, d\tau \right|. \tag{3.18}$$

In employing the method of steepest descents to evaluate the Euclidean path integral appearing in Eqs. (3.15) and (3.16), we should choose a contour in the complex functional space such that the quadratic parts of the exponentials are negative definite. For the integration around $x_c = 0$, we have $\omega_k^2 > 0$, and hence we may choose the contour on which the functions satisfy $\xi_{-k} = \xi_k^*$ (* here means taking the complex conjugate); in other words, we are integrating over the real functions around $x_c = 0$. For the integration around x_c^{+}, we have $\Omega^2_{-1} < 0$ and $\Omega_k^2 > 0$ so we may choose η_{-1} as $i\sigma$ with σ real and $\eta_{-k} = \eta_k^*$. For the integration around x_c^{-}, we follow the suggestions given in Sec. II, i.e., we choose ρ_{-1} as $-i\sigma$ with σ real and $\rho_{-k} = (\rho_k)^*$.

With the above chosen contour, the numerator becomes

$$\begin{aligned}
&\int \left(\sum_k d\xi_k \, d\xi_k^* \right) \left[\sum_{k'} \xi_{k'} \theta_{k'}(\tau_1) \right] \left[\prod_{k'} \xi_{k'} \theta_{k'}(\tau_2) \right] \exp\left(-\tfrac{1}{2} \sum_k \omega_k^2 |\xi_k|^2 \right) + i \int d\sigma \left(\prod d\eta \, d\eta^* \right) \left| \int_{-\infty}^{+\infty} \left(\dot{x}_c^{+} + i\sigma\dot{\varphi}_{-1} \right.\right. \\
&\quad \left.\left. + \sum_k \dot{\varphi}_k \eta_k \right) \dot{x}_c^{+} \right| \int_{-\infty}^{+\infty} da \, x_c^{+}(\tau_1 - a) x_c^{+}(\tau_2 - a) \exp\left(\frac{4}{3g} + \frac{1}{2}\Omega^2_{-1}\sigma^2 - \frac{1}{2}\sum \Omega_k^2 |\eta_k|^2 \right) \\
&\quad - i \int d\sigma \left(\prod d\eta \, d\eta^* \right) \left| \int_{-\infty}^{+\infty} \left(\dot{x}_c^{-} - i\sigma\dot{\varphi}_{-1} + \sum_k \dot{\varphi}_k \rho_k \right) \dot{x}_c^{-} \right| \int_{-\infty}^{+\infty} da \, x_c^{-}(\tau_1 - a) x_c^{-}(\tau_2 - a) \\
&\quad \times \exp\left(\frac{4}{3g} + \frac{1}{2}\Omega^2_{-1}\sigma^2 - \frac{1}{2}\sum \Omega_k^2 |\rho_k|^2 \right),
\end{aligned} \tag{3.19}$$

and the denominator becomes

$$\begin{aligned}
&\int \left(\prod_k d\xi_k \, d\xi_k^* \right) \exp\left(-\tfrac{1}{2} \sum_k \omega_k^2 |\xi_k|^2 \right) + i \int d\sigma \left(\prod d\eta \, d\eta^* \right) \left| \int_{-\infty}^{+\infty} \left(\dot{x}_c^{+} + i\sigma\dot{\varphi}_{-1} + \sum_k \dot{\varphi}_k \eta_k \right) \dot{x}_c^{+} \right| \\
&\times \int_{-\infty}^{+\infty} da \exp\left(\frac{4}{3g} + \frac{1}{2}\Omega^2_{-1}\sigma^2 - \frac{1}{2}\sum \Omega_k^2 |\eta_k|^2 \right) \\
&- i \int d\sigma \left(\prod d\eta \, d\eta^* \right) \left| \int_{-\infty}^{+\infty} \left(\dot{x}_c^{-} - i\sigma\dot{\varphi}_{-1} + \sum_k \dot{\varphi}_k \eta_k \right) \dot{x}_c^{-} \right| \int_{-\infty}^{+\infty} da \exp\left(\frac{4}{3g} + \frac{1}{2}\Omega^2_{-1}\sigma^2 - \frac{1}{2}\sum \Omega_k^2 |\rho_k|^2 \right).
\end{aligned} \tag{3.20}$$

We write all these things in detail because we want to show that we have a situation completely different from that given in Sec. II: The contributions from x_c^{+} and x_c^{-} do *not* cancel each other as the contributions around $\pm i[1/(2g)^{1/2}]$ did in the simple example of Sec. II because $\Delta^{+} \neq \Delta^{-}$ after being rotated into the paths of steepest descents; this is a manifestation of the infinite degrees of freedom in quantum mechanics because we are doing integration in a functional space as compared with the single degree of freedom of our simple example. If we include the cubic and quartic terms in the exponential, it is even more transparent that the contributions from x_c^{+} and x_c^{-} do not cancel each other.

There is one more difficulty we have to overcome before we can do any computation: We have to dispose of the infinity coming from the integration of the collective coordinate. This difficulty is, however, fictitious because it is unrealistic to consider only the contributions coming from x_c^{+} and x_c^{-}; there is an infinite number of solutions of the classical equation of motion satisfying the required boundary conditions. They are the so called "multipseudoparticle" solutions. The contributions coming from the integration of the collective coordinates of the multipseudoparticle will exponentiate as we shall see from the following.

Let us now consider the solution $x_c^{N+M}(\tau; a_1, \ldots, a_{N+M})$ which is a patching of $N x_c^{+}$ solutions and $M x_c^{-}$ solutions; $a_1, \ldots, a_{N+M}$ are where these solutions are localized. There are $(N+M)!/N!M!$ of such solutions arising from the permutations of the configurations of x_c^{+} and x_c^{-}. A general path can be expanded around $x_c^{N+M}(\tau; a_1, \ldots, a_{N+M})$ as

$$x(\tau) = x_c^{N+M}(\tau; a_1, \ldots, a_{N+M}) + \sum_{n, \tau \approx a_1} \beta_n^1 \varphi_n(\tau - a_1) + \sum_{n, \tau \approx a_2} \beta_n^2 \varphi_n(\tau - a_2) + \cdots, \tag{3.21}$$

where $\Sigma_n \beta_n^j \varphi_n(\tau - a_j)$ is the deviation of $x(\tau)$ from the jth solution in the region somewhere around a_j. Then $\exp(+\int_{-\infty}^{+\infty} \mathscr{L}^E d\tau)$ will look like (ignoring the higher nonlinear terms)

$$\prod_{j=1}^{N} \exp\left(\frac{4}{3g} - \frac{1}{2}(\Omega_{-1})^2 (\eta^j{}_{-1})^2 - \frac{1}{2}\sum_k \Omega_k^2 \eta_k^j \eta^j{}_{-k} \right) \prod_{l=1}^{M} \exp\left(\frac{4}{3g} - \frac{1}{2}(\Omega_{-1})^2 (\rho^l{}_{-1})^2 - \frac{1}{2}\sum_k \Omega_k^2 \rho_k^l \rho^l{}_{-k} \right), \tag{3.22}$$

where we replace β_j by η^j when the jth solution is x_c^+ and β_l by ρ^l when the l th solution is x_c^-. To eliminate the zero frequency eigenfunction, we insert into the functional integral

$$\Delta^{N+M}[x]\int_{-\infty}^{+\infty} da_1\delta\Big[\int_{\tau\approx a_1} d\tau\varphi_0(\tau-a_1)x(\tau)\Big]\int_{a_1}^{+\infty} da_2\delta\Big[\int_{\tau\approx a_2} d\tau\varphi_0(\tau-a_2)x(\tau)\Big]\cdots=1; \tag{3.23}$$

because the solutions are highly localized, we have

$$\Delta^{N+M}[x]=\prod_{j=1}^{N}\Delta_j^+\prod_{l=1}^{M}\Delta_l^-. \tag{3.24}$$

The measure can then be written as

$$\prod_{j=1}^{N}\prod_{l=1}^{M}\Delta_j^+\,d\eta^j\Delta_l^-\,d\rho^l\,da_1\cdots da_{N+M},\quad a_1<\cdots<a_{N+M}. \tag{3.25}$$

Furthermore, because $x_c^+x_c^+=x_c^-x_c^-$, we have

$$x_c^{N+M}(\tau_1;a_1,\ldots,a_{N+M})x_c^{N+M}(\tau_2;a_1,\ldots,a_{N+M})\simeq x_c^+(\tau_1-a_1)x_c^+(\tau_2-a_1)+\cdots+x_c^+(\tau_1-a_{N+M})x_c^+(\tau_2-a_{N+M}), \tag{3.26}$$

which is symmetric in $a_1,\ldots,a_{N+M}$.

Hence, if we approximate $x(\tau_1)x(\tau_2)$ by $x_c^{N+M}(\tau_1;a_1,\ldots,a_{N+M})x_c^{N+M}(\tau_2;a_1,\ldots,a_{N+M})$, we can write the x_c^{N+M} contribution to $\int\mathscr{D}[x]x(\tau_1)x(\tau_2)\exp(+\int_{-\infty}^{+\infty}\mathscr{L}^E d\tau)$ as

$$\begin{aligned}&\frac{(N+M)!}{N!M!}\int\prod_{j=1}^{N}\prod_{l=1}^{M}\Delta_j^+\Delta_l^-\,d\eta^j\,d\rho^l\frac{da_1\ldots da_{N+M}}{(N+M)!}\,x_c^{N+M}(\tau_1;a_1,\ldots,a_{N+M})x_c^{N+M}(\tau_2;a_1,\ldots,a_{N+M})\\&\quad\times\exp\Big(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\eta_{-1}^2-\frac{1}{2}\sum\Omega_k^2\eta_k^j\eta_{-k}^j\Big)\exp\Big(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\rho_{-1}^2-\frac{1}{2}\sum\Omega_k^2\rho_k^j\rho_{-k}^j\Big)\\&\quad=\int_{-\infty}^{+\infty}dbx_c^+(\tau_1-b)x_c^+(\tau_2-b)\frac{(N+M)}{N!M!}\\&\qquad\times\Big[\int\Big(\prod_k d\eta_k\Big)\exp\Big(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\eta_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\eta_k\eta_{-k}\Big)\Delta^+\Big]^N\\&\qquad\times\Big[\int\Big(\prod_k d\rho_k\Big)\exp\Big(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\rho_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\rho_k\rho_{-k}\Big)\Delta^-\Big]^M\Big(\int da\Big)^{N+M-1},\end{aligned} \tag{3.27}$$

while the x_c^{N+M} contribution to $\int\mathscr{D}[x]\exp(\int_{-\infty}^{+\infty}\mathscr{L}^E d\tau)$ is

$$\begin{aligned}&\frac{1}{N!}\Big[\int\Delta^+ da\Big(\prod_k d\eta_k\Big)\exp\Big(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\eta_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\eta_k\eta_{-k}\Big)\Big]^N\\&\quad\times\frac{1}{M!}\Big[\int\Delta^- da\Big(\prod_k d\rho_k\Big)\exp\Big(\frac{4}{3g}-\frac{1}{2}\Omega_{-1}^2\rho_{-1}^2-\frac{1}{2}\sum_k\Omega_k^2\rho_k\rho_{-k}\Big)\Big]^M.\end{aligned} \tag{3.28}$$

In applying the method of steepest descents, we rotate each η_{-1} to $i\sigma$ and each ρ_{-1} to $-i\sigma$, and at the same time identifying η_{-k} with η_k^* and ρ_{-k} with ρ_k^*; we can write Eq. (3.27) as

$$\int_{-\infty}^{+\infty}db\,x_c^+(\tau_1-b)x_c^+(\tau_2-b)\frac{(N+M)}{N!M!}\Big(i\int\cdots\Big)^N\Big(-i\int\cdots\Big)^M\Big(\int da\Big)^{N+M-1}, \tag{3.29}$$

and Eq. (3.25) as

$$\frac{1}{N!}\frac{1}{M!}\Big(i\int\cdots\Big)^N\Big(-i\int\cdots\Big)^M. \tag{3.30}$$

Summing over the contributions from all multipseudoparticles (i.e., summing N and M from 0 to ∞ except when both N and M are zero),

$$\begin{aligned}&\Big[\int\mathscr{D}[x]x(\tau_1)x(\tau_2)\exp\Big(\int_{-\infty}^{+\infty}\mathscr{L}^E d\tau\Big)\Big]_{\text{from all pseudoparticles}}\\&\quad=\int_{-\infty}^{+\infty}x_c^+(\tau_1-b)x_c^+(\tau_2-b)\,dbi\int d\sigma\Big(\prod d\eta\Big)(\Delta^+-\Delta^-)\exp\Big(\frac{4}{3g}+\frac{1}{2}\Omega_{-1}^2\sigma^2-\frac{1}{2}\sum\Omega_k^2|\eta_k|^2\Big)\\&\qquad\times\exp\Big\{i\int da\Big[\int d\sigma\prod d\eta\Delta^+\exp\Big(\frac{4}{3g}+\frac{1}{2}\Omega_{-1}^2\sigma^2-\frac{1}{2}\sum\Omega_k^2|\eta_k|^2\Big)\\&\qquad-\int d\sigma\prod d\rho\Delta^-\exp\Big(\frac{4}{3g}+\frac{1}{2}\Omega_{-1}^2\sigma^2-\frac{1}{2}\sum\Omega_k^2|\rho_k|^2\Big)\Big]\Big\}\end{aligned} \tag{3.31}$$

and

$$\left[\int \mathscr{D}[x]\exp\left(\int \mathscr{L}^E d\tau\right)\right]_{\text{from all pseudoparticles}}$$
$$= \exp\left\{i\int da\left[\int d\sigma\prod d\eta\Delta^{+}\exp\left(\frac{4}{3g}+\frac{1}{2}\Omega^2_{-1}\sigma^2-\frac{1}{2}\sum\Omega_k^2|\eta_k|^2\right)\right.\right.$$
$$\left.\left.-\int d\sigma\prod d\rho\Delta^{-}\exp\left(\frac{4}{3g}+\frac{1}{2}\Omega^2_{-1}\sigma^2-\frac{1}{2}\sum\Omega_k^2|\rho_k|^2\right)\right]\right\}-1. \tag{3.32}$$

Now, $\int d\sigma\Pi\, d\eta\Delta^{+}\exp(4/3g+1/2\Omega^2_{-1}\sigma^2-1/2\Sigma\Omega_k^2|\eta_k|^2)$ is a high order zero because it is an infinite product of factors of the form $1/(1+k^2)$, and hence it is still zero even multiplied by $\int_{-\infty}^{+\infty} da$. The same is true for $\int d\sigma\Pi\, d\rho\Delta^{-}\exp(4/3g+1/2\Omega^2_{-1}\sigma^2-1/2\Sigma\Omega_k^2|\rho_k|^2)$. Thus, the whole thing of

$$\exp\left\{i\int da\left[\int d\sigma\prod d\eta\Delta^{+}\exp\left(\frac{4}{3g}+\frac{1}{2}\Omega^2_{-1}\sigma^2-\frac{1}{2}\sum\Omega_k^2|\eta_k|^2\right)\right.\right.$$
$$\left.\left.-\int d\sigma\prod d\rho\Delta^{-}\exp\left(\frac{4}{3g}+\frac{1}{2}\Omega^2_{-1}\sigma^2-\frac{1}{2}\sum\Omega_k^2|\eta_k|^2\right)\right]\right\}$$

is nothing but equal to 1.

With this important simplification, we can write our results as

$$\left[\int \mathscr{D}[x]\,x(\tau_1)x(\tau_2)\exp\left(\int \mathscr{L}^E d\tau\right)\right]_{\text{from all pseudoparticles}}$$
$$= \int_{-\infty}^{+\infty} x_c^{+}(\tau_1-b)x_c^{+}(\tau_2-b)\,db\int da\prod d\eta i(\Delta^{+}-\Delta^{-})\exp\left(\frac{4}{3g}+\frac{1}{2}\Omega^2_{-1}\sigma^2-\frac{1}{2}\sum\Omega_k^2|\eta_k|^2\right) \tag{3.33}$$

and

$$\left[\int \mathscr{D}[x]\exp\left(\int \mathscr{L}^E d\tau\right)\right]_{\text{from all pseudoparticles}} = 0.$$

If we add in the contribution from the trivial solution x_c^0, we will have

$$\int \mathscr{D}[x]\,x(\tau_1)x(\tau_2)\exp\left(\int \mathscr{L}^E d\tau\right)$$
$$= \int\prod d\xi\, d\xi^{*}\sum\xi_n\theta_n(\tau_1)\sum\xi_m\theta_m(\tau_2)\exp\left(-\frac{1}{2}\sum\omega_k^2|\xi_k|^2+I_4\right)+\int_{-\infty}^{+\infty} dbx_c^{+}(\tau_1-b)x_c^{+}(\tau_2-b)$$
$$\times\int d\sigma\prod d\eta\, d\eta^{*}i(\Delta^{+}-\Delta^{-})\exp\left(\frac{4}{3g}+\frac{1}{2}\Omega^2_{-1}\sigma^2-\frac{1}{2}\sum\Omega_k^2|\eta_k|^2\right) \tag{3.34}$$

and

$$\int \mathscr{D}[x]\exp\left(\int \mathscr{L}^E d\tau\right)=\int\prod d\xi\, d\xi^{*}\exp\left(-\frac{1}{2}\sum\omega_k^2|\xi_k|^2+I_4\right). \tag{3.35}$$

The term I_4 is the quartic term appearing in Eq. (3.10).

The ratio of Eqs. (3.34) and (3.35) corresponds to the Green's function $\langle 0|T[\hat{x}(\tau_1)\hat{x}\tau_2)]|0\rangle$. Before we calculate the ratio, however, we must make it clear that the result expressed in Eq. (3.34) is not the complete result for the path integral $\int\mathscr{D}[x]x(\tau_1)x(\tau_2)e^{\int_{-\infty}^{+\infty}\mathscr{L}^E d\tau}$. In Eq. (3.34) we have neglected the cubic and quartic terms as given by Eqs. (3.10) and (3.11). We have also neglected the small oscillational terms of Eq. (3.21). We deliberately neglected these terms in the above discussions because we don't want to make our equations look too clumsy. However, now we have to fill them in. If we include all these terms, and with exactly the same argument as given in the above, it can be easily shown that the complete expression for $\int\mathscr{D}[x]x(\tau_1)x\tau_2)\exp(\int_{-\infty}^{+\infty}\mathscr{L}^E d\tau)$ will then become

$$\int\prod d\xi\, d\xi^{*}\sum\xi_n\theta_n(\tau_1)\sum\xi_m\theta_m(\tau_2)\exp\left(-\frac{1}{2}\sum\omega_k^2|\xi_k|^2+I_4\right)+\left\{i\int_{-\infty}^{+\infty} dbx_c^{+}(\tau_1-b)x_c^{+}(\tau_2-b)\right.$$
$$\times\left.\int d\sigma\prod d\eta\, d\eta^{*}\delta^{+}-i\int_{-\infty}^{+\infty} dbx_c^{-}(\tau_1-b)x_c^{-}(\tau_2-b)\int d\sigma\prod d\eta d\eta^{*}\delta^{-}\right\}\exp(Q)+\left\{i\int_{-\infty}^{+\infty} db\,d\sigma\prod d\eta\, d\eta^{*}\right.$$
$$\times x_c^{+}(\tau_1-b)\left[+i\sigma\varphi_{-1}(\tau_2-b)+\sum\eta_k\varphi_k(\tau_2-b)\right]\delta^{+}-i\int_{-\infty}^{+\infty} db\,d\sigma\prod d\eta\, d\eta^{*}x_c^{+}(\tau_1-b)$$
$$\times\left[-i\sigma\varphi_{-1}(\tau_2-b)+\sum\eta_k\varphi_k(\tau_2-b)\right]\delta^{-}+(\tau_1\leftrightarrow\tau_2)\left.\right\}\exp(Q)+\left\{i\int_{-\infty}^{+\infty} db\,d\sigma\prod d\eta\, d\eta^{*}\right.$$
$$\times\left[i\sigma\varphi_{-1}(\tau_1-b)+\sum\eta_k\varphi_k(\tau_1-b)\right]\left[i\sigma\varphi_{-1}(\tau_2-b)+\sum\eta_k\varphi_k(\tau_2-b)\delta^{+}\right]-i\int_{-\infty}^{+\infty} db\,d\sigma\prod d\eta\, d\eta^{*}$$

$$\times\Big[-i\sigma\varphi_{-1}(\tau_1-b)+\sum\eta_k\varphi_k(\tau_1-b)\Big]\Big[-i\sigma\varphi_{-1}(\tau_2-b)+\sum\eta_k\varphi_k(\tau_2-b)\Big]\delta^-\Big\}\exp(Q), \tag{3.36}$$

where

$$\delta^+=\Delta^+\exp-i\sqrt{2g}\int_{-\infty}^{+\infty}\frac{1}{\cosh\tau}\Big(i\sigma\varphi_{-1}+\sum\eta_k\varphi_k\Big)^3d\tau-\frac{g}{4}\int_{-\infty}^{+\infty}\Big(i\sigma\varphi_{-1}+\sum\eta_k\varphi_k\Big)^4d\tau, \tag{3.37}$$

$$\delta^-=\Delta^-\exp i\sqrt{2g}\int_{-\infty}^{+\infty}\frac{1}{\cosh\tau}\Big(-i\sigma\varphi_{-1}+\sum\eta_k\varphi_k\Big)^3d\tau-\frac{g}{4}\int_{-\infty}^{+\infty}\Big(-i\sigma\varphi_{-1}+\sum\eta_k\varphi_k\Big)^4d\tau$$

$$=\Delta^-\exp-i\sqrt{2g}\int_{-\infty}^{+\infty}\frac{1}{\cosh\tau}\Big(i\sigma\varphi_{-1}-\sum\eta_k\varphi_k\Big)^3d\tau-\frac{g}{4}\int_{-\infty}^{+\infty}\Big(i\sigma\varphi_{-1}-\sum\eta_k\varphi_k\Big)^4d\tau$$

$$Q=\frac{4}{3g}+\frac{1}{2}\Omega^2_{-1}\sigma^2-\frac{1}{2}\sum_k\Omega_k^2|\eta_k|^2. \tag{3.38}$$

Regrouping, and taking into account of the fact that $x_c^+=-x_c^-$, we can write $\int\mathscr{D}[x]x(\tau_1)x(\tau_2)e^{\int(\frac{1}{2}\dot{x}^2+V)d\tau}$ as

$$\int\prod d\xi\,d\xi^*\sum\xi_n\theta_n(\tau_1)\sum\xi_m\theta_m(\tau_2)\exp\Big(-\frac{1}{2}\sum\omega_k^2|\xi_k|^2+I_4\Big)+i\int db\,d\sigma\prod d\eta\,d\eta^*$$

$$\times(\delta^+-\delta^-)\exp(Q)\Big[x_c^+(\tau_1-b)x_c^+(\tau_2-b)+x_c^+(\tau_1-b)i\sigma\varphi_{-1}(\tau_2-b)+x_c^+(\tau_2-b)i\sigma\varphi_{-1}(\tau_1-b)$$

$$+i^2\sigma^2\varphi_{-1}(\tau_1-b)\varphi_{-1}(\tau_2-b)+\sum\eta_k\varphi_k(\tau_1-b)\sum\eta_k\varphi_k(\tau_2-b)\Big]+i\int db\,d\sigma\prod d\eta\,d\eta^*(\delta^+-\delta^-)$$

$$\times\exp(Q)\Big[x_c^+(\tau_1-b)\sum\eta_k\varphi_k(\tau_2-b)+x_c^+(\tau_2-b)\sum\eta_k\varphi_k(\tau_1-b)+i\sigma\varphi_{-1}(\tau_1-b)\sum\eta_k\varphi_k(\tau_2-b)$$

$$+i\sigma\varphi_{-1}(\tau_2-b)\sum\eta_k\varphi_k(\tau_1-b)\Big]. \tag{3.39}$$

The first term of Eq. (3.39) together with Eq. (3.40) will give the conventional Feynman–Dyson series.[15] The second term which looks like $(\delta^+-\delta^-)\{\cdots\}$ with $\{\cdots\}$ containing expressions which are even in η and the third term which looks like $(\delta^++\delta^-)\{\cdots\}$ with $\{\cdots\}$ now containing expressions which are odd in η seemed very complicated at first glance. Their contributions are, however, equal to zero. The reasons are the following: To evaluate the integrations of $\Pi\,d\eta\,d\eta^*$, we use the same methods as given in Ref. 15 by making $\eta=re^{i\theta}$ and then writing $\Pi\,d\eta\,d\eta^*$ as $\Pi r\,dr\,d\theta$. Then, the integrations will be nonzero only if the angular integrations are nonzero, and this, for the second term, means that $(\delta^+-\delta^-)\{\cdots\}$ must contain equal number of η and η^* terms. However, expanding Δ^+ for δ^+ and Δ^- for δ^-, it can be seen that (see Appendix) $(\delta^+-\delta^-)\{\cdots\}$ does not contain terms which have equal number of η and η^*. Similarly, the third term which is of the form $(\delta^++\delta^-)\{\cdots\}$ does not contain terms which have equal number of η and η^*. Thus, the second integral and the third integral of Eq. (3.39) are equal to zero, even though their corresponding integrands are not. We note that it is crucial for the second term to be a difference $(\delta^+-\delta^-)$ and the third term to be a sum $(\delta^++\delta^-)$ in order to secure the zero contributions from both integrals.

IV. CONCLUSIONS

We conclude this paper by saying that the Green's functions of the quantum mechanical anharmonic oscillator is adequately described by the Feynman–Dyson series. We may ask the question, however, that why bother to spend so much effort in proving something that is so familiar to us. We would like to answer this question by mentioning the following facts: Firstly, there is no *a priori* reason for us to say that the Feynman–Dyson series is adequate. Secondly, the path integrals are mysterious objects. They are so mysterious that path integrals integrated over real functions which are defined on real arguments have contributions from functions which are defined over imaginary arguments in the discussions of tunnelings in quantum mechanics. Hence, there is no reason to rule out the significance of imaginary functions in the calculations of Green's functions in quantum mechanics. Thirdly, we know how to handle these pseudoparticle solutions and we know that the way we handle the pseudoparticle solutions are the right ways. We particularly note that all the solutions enter into our discussions in the same footings and can be generalized easily if there are more than one kind of pseudoparticle solutions. Lastly, the vanishing of the pseudoparticle contributions come from reasons that we have not met before, namely, that there is a negative frequency solution Ω^2_{-1} which forces us to rotate our integration into the imaginary axis and that $(\delta^+-\delta^-)$ has no even term in η^* while $(\delta^++\delta^-)$ has no odd terms in η^*.

With all the reasons given in the above, we think that our efforts are well justified.

APPENDIX

In this Appendix, we shall show that $(\delta^+-\delta^-)X$

expressions which are even powers in η does not contain terms which have equal number of η and η^*. We start with Δ^+ which can be written symbolically as $|1+i\eta|$ $=(1+i\eta-i\eta^*+\eta\eta^*)^{1/2}$. Expanding,

$$\Delta^+ = 1+\tfrac{1}{2}(i\eta - i\eta^* + \eta\eta^*) - \tfrac{1}{8}(i\eta - i\eta^* + \eta\eta^*)^2 + \cdots. \tag{A1}$$

Similarly,

$$\begin{aligned}\Delta^- &= |1-i\eta| = \sqrt{1-i\eta+i\eta^*+\eta\eta^*} \\ &= 1+\tfrac{1}{2}(-i\eta+i\eta^*+\eta\eta^*) \\ &\quad -\tfrac{1}{8}(-i\eta+i\eta^*+\eta\eta^*)^2+\cdots.\end{aligned} \tag{A2}$$

The δ^+ and δ^- can be written symbolically as

$$\begin{aligned}\delta^+ &= \Delta^+ \exp[-(i\sigma+\eta)^3-(i\sigma+\eta)^4], \\ \delta^- &= \Delta^- \exp[-(i\sigma-\eta)^3-(i\sigma-\eta)^4],\end{aligned} \tag{A3}$$

respectively. If we expand the exponentials that appear in Eq. (A3), then $(\delta^+ - \delta^-)$ will look like

$$\sum_N \eta^N(\Delta^+ \pm \Delta^-). \tag{A4}$$

The plus sign corresponds to the case when N is odd, and the minus sign corresponds to the case when N is even.

However, from Eqs. (A1) and (A2), we know that $(\Delta^+ + \Delta^-)$ contains only terms of the form $(-i\eta+i\eta^*)^{\text{even power}}$ which has the property that (the power of η^* − the power of η) = even number, and hence for N odd, $\eta^N(\Delta^+ + \Delta^-)$ has the property that (the power of η^*–the power of η) = odd number. We also see that $(\Delta^+ - \Delta^-)$ contains only terms of the form $(-i\eta+i\eta^*)^{\text{odd power}}$, hence for N even, $\eta^N(\Delta^+ - \Delta^-)$ has the same above property. Thus, we see that $(\delta^+ - \delta^-)X$ expressions which are even power in η cannot contain terms which have an equal number of η and η^*.

[1] A.A. Belavin, A.M. Polyakov, A.S. Schwartz, and Yu.S. Tyupkin, Phys. Lett. B **59**, 85 (1975).
[2] A.M. Polyakov, Phys. Lett. B **59**, 82 (1975); Nucl. Phys. B **120**, 429 (1977).
[3] G. 't Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. D **14**, 3432 (1976).
[4] R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 172 (1976).
[5] C. Callan, R. Dashen, and D. Gross, Phys. Lett. B **63**, 334 (1976).
[6] C.S. Lam, Nuovo Cimento A **47**, 451 (1966); **50**, 504 (1976).
[7] For the properties of asymptotic expansion, see the book of E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis* (Cambridge U.P., London, 1965).
[8] See the book of N.G. de Bruijn, *Asymptotic Methods in Analysis* (North-Holland, Amsterdam, 1961).
[9] We should use " " instead of " = " because we are talking about asymptotic expansions; however, we will stick with this convention in the subsequent discussion.
[10] R.F. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
[11] The rotation into the imaginary axis in order to project out the ground state expectation values was discussed in the report E.S. Abers and B.W. Lee, Phys. Rep. C **9**, 1 (1973).
[12] See Chap. VI. of the book of Jan Rzewnski, *Field Theory* (Hafner, New York, 1969), Vol. 2.
[13] P. Morse and H. Feshback, *Method of Theoretical Physics* (McGraw-Hill, New York, 1953), p. 1650.
[14] J.L. Gervais and B. Sakita, Phys. Rev. D **11**, 2943 (1975); A. Jevicki, Nucl. Phys. B **117**, 365 (1976) and references therein.
[15] W.B. Yeung, J. Phys. A **11**, 991 (1978).

A set of commuting operators and R(3) scalars for the complete classication of quadrupole phonon-states

H.E. De Meyer[a] and G. Vanden Berghe

Seminarie voor Wiskundige Natuurkunde, Rijksuniversiteit Gent, Krijgslaan 271-S9, B-9000 Gent, Belgium

(Received 25 June 1979; accepted for publication 3 October 1979)

A method is developed for constructing operator forms which commute with the Casimir operators of the groups appearing in the chain U(5)⊃SU(5)⊃R(5)⊃R(3)⊃(2), according to which the quadrupole-phonon states are usually classified. An expression for such operator is given.

1. INTRODUCTION

The quadrupole vibrations of the liquid drop, as discussed by Bohr[1] over twenty years ago, have become the natural starting point for the collective description of the positive-parity states of even–even nuclei. By using group theoretical techniques, it is possible to exactly define the wavefunctions of these quadrupole-phonon states for arbitrary angular momentum J_N and definite number of quanta N.[2-8] Five quantum numbers are needed to uniquely classify the states built up by N quadrupole phonons. Three of them are trivial: the quadrupole boson number N of the U(5) group, the angular momentum J_N of the R(3) group, and its third component M_N of the R(2) group. The fourth is the boson seniority v of the R(5) group. The fifth label however is not related with the eigenvalue of the Casimir operator of any group. One usually introduces a quantum number, which counts boson triplets coupled to zero angular momentum.[2,3]

Very recently Vanden Berghe and De Meyer[9] have introduced a new method for constructing N-quadrupole-phonon states. Starting from this new construction scheme, a method is developed for searching in a systematic way operators which commute with the Casimir operators of the U(5), R(5), R(3), and R(2) groups and which are independent of them. The eigenvalue of such operator can then be used as a fifth label which can serve for a unique identification of the considered phonon states. The general form of such operators and the conditions which have to be fulfilled such that the derived expression is useful for the purpose we achieve are firstly discussed. Afterwards operator forms built up with respectively two-, three-, and four-phonon creation and annihilation operators are considered until an acceptable expression has been obtained.

2. CANONICAL OPERATORS

In order to find in a systematic way, a complete minimum label description of U($2\lambda+1$) phonon-states in terms of eigenvalues of mutually commuting operators, we consider as a starting point the set of scalar operators defined by

$$O^{I_2 I_3 \cdots I_N \; J_N}_{J_2 J_3 \cdots J_N \; J_N} = \sqrt{2J_N+1}\,[(\underbrace{(\cdots((bb)^{I_2} b)^{I_3}\cdots)^{I_{N-1}} b}_{\times}\,)^{J_N}]^0 \quad (N\geqslant 2), \qquad (2.1)$$

$$O = \sqrt{2\lambda+1}\,(b^+ b)^0 \quad (N=1),$$

also called canonical operators as to refer to the property that all the elementary phonon creation operators b^+_μ (angular momentum λ with projection μ) stand to the left of the phonon annihilation operators $(-1)^\mu b_{-\mu}$. It is readily verified that (2.1) can be rewritten in the alternative form

$$\begin{aligned} O^{I_2\cdots J_N}_{J_2\cdots J_N} = & \sum_{m_1,m_2,\ldots,m_N} \sum_{M_2,\ldots,M_N} \sum_{M_N} \langle \lambda m_1 \lambda m_2 | J_2 M_2 \rangle \\ & \times \langle J_2 M_2 \lambda m_3 | J_3 M_3 \rangle \cdots \langle J_{N-1} M_{N-1} \lambda m_N | J_N M_N \rangle \\ & \times \langle \lambda m'_1 \lambda m'_2 | I_2 M'_2 \rangle \langle I_2 M'_2 \lambda m'_3 | I_3 M'_3 \rangle \cdots \\ & \times \langle I_{N-1} M'_{N-1} \lambda m'_N | J_N M_N \rangle \\ & \times b^+_{m_1} b^+_{m_2} \cdots b^+_{m_N} b_{m'_1} b_{m'_2} \cdots b_{m'_N} \quad (N\geqslant 2), \\ O = & \sum_m b^+_m b_m \quad (N=1), \end{aligned} \qquad (2.2)$$

which is the most useful form for computational purposes.

The set $\{O^{I_2\cdots J_N}_{J_2\cdots J_N} | N\geqslant 1\}$ is overcomplete, since for fixed N- and J_N-values the operators having different intermediate angular momentum values are not necessarily independent. In fact, the relations existing between these operators follow immediately from analogous relations between the different N-phonon states with total angular momentum J_N, hereafter called (N, J_N)-states, which, being labelled equally by the intermediate angular momentum values have been introduced and analyzed in detail in a previous paper.[9] As a consequence it is convenient to indicate for any fixed N and J_N, in so far as physical (N, J_N)-states exist, a minimal set of independent operators $O^{I_2\cdots J_N}_{J_2\cdots J_N}$, hereafter called (N, J_N)-operator basis, by using exactly the same selection criterion as the one that leads us to the particular choice of sets of independent (N, J_N)-states $|0\lambda J_2 \cdots J_N\rangle$, also called (N, J_N)-state basis. Remark that in labelling the operators, we have twice omitted the irrelevant momenta 0 and λ. The number of canonical operators in a (N, J_N)-operator basis is evidently the square of the number of states in the corresponding (N, J_N)-state basis, and this last number is usually derived from pure group-theoretical quantities, such as the group characters.[10] Finally, we define a (N, J_N)-operator (-state) basis to be a singleton-basis if there is only one operator (state) present in the basis.

In the present paper we are essentially concerned with the quadrupole ($\lambda = 2$) phonon case. Nevertheless, for further use, the integer λ will be kept arbitrary in most of the

[a] Aangesteld Navorser bij het Nationaal Fonds voor Wetenschappelijk Onderzoek.

 0022-2488/80/030486-07$1.00

following formulas. Finally, it has to be remarked that although the selection of the different operator bases is certainly not unique, the existence of a complete solution to the state labelling problem has to be independent of any such particular choice.

It is easily verified that the physical operators N (the number operator), J^2 (the angular momentum operator) and V^* [the Casimir operator of the R(2λ +1) group[7]] can be expressed in canonical form as follows

$$N = O, \tag{2.3}$$

$$J_2 = -\frac{\lambda(\lambda+1)(2\lambda+1)}{\sqrt{3}}[(b^+b)^1(b^+b)^1]^0$$
$$= \lambda(\lambda+1)O + \lambda(\lambda+1)(2\lambda+1)$$
$$\times \sum_J \begin{Bmatrix} \lambda & \lambda & J \\ \lambda & \lambda & 1 \end{Bmatrix} O^J, \tag{2.4}$$

$$V^* = \tfrac{1}{2}(2\lambda+1)(b^+b^+)^0(bb)^0 + \tfrac{1}{2}(1-2\lambda)N - \tfrac{1}{2}N^2$$
$$= -\lambda O + \tfrac{1}{2}(2\lambda+1)O_0^0 - \tfrac{1}{2}\sum_J O^J, \tag{2.5}$$

where in the last step we have used the relation

$$N^2 = O + \sum_J O^J, \tag{2.6}$$

and whereby the J-summations extend over the even integer values from $J = 0$ up to $J = 2\lambda$. The angular momentum projection operator J_0 is scalar but cannot be brought into canonical form. Fortunately this fact has no serious implications for our forthcoming analysis, since all operators of the form (2.1) trivially commute with N, J^2, and J_0.

Turning now in particular to the quadrupole case, we remark that the previously tabulated phonon-states[2,3,9] are all eigenstates of N, J^2, V^*, and J_0, but that for the sake of a complete and unambiguous classification of these states, a fifth label is required. As it is here our major concern to identify this extra label to the eigenvalue of a suitable operator, we will systematically investigate the operators of the form

$$X[\{f\}] = \sum_{N=2}^{N_{\max}} \sum_{J_N} \sum_{\substack{I_i J_i \\ (i=1,\ldots,N-1)}} f^{I_2 I_3 \ldots J_N}_{J_2 J_3 \ldots J_N} O^{I_2 I_3 \ldots J_N}_{J_2 J_3 \ldots J_N}, \tag{2.7}$$

by consecutively increasing $N_{\max}$ by one unit.

In (2.7) the X-operator is unambiguously determined by the real values of the variables $f^{I_2 \ldots J_N}_{J_2 \ldots J_N}$, occurring in the index-ordered set $\{f\}$, while on the right hand side the internal angular momentum summations are restricted to those values for which the associated canonical operators are elements of an (N, J_N)-operator basis.

In the first instance, the candidates for a suitable extra label generating operator, are those $X[\{f\}]$ for which

$$[V^*, X[\{f\}]] = 0, \tag{2.8}$$

since the commutation relations

$$[N, X[\{f\}]] \equiv 0,$$
$$[J^2, X[\{f\}]] \equiv 0, \tag{2.9}$$

and therefore also the relation

$$[N^2, X[\{f\}]] \equiv 0, \tag{2.10}$$

are identically satisfied. A second and necessary condition to be implied on those operators $X[\{f\}]$ which already fulfill (2.8), is that they should be independent of V^*, J^2, and N in the sense that they may not be a function (polynomial) F of V^*, J^2, and N alone; i.e.,

$$\forall F: X[\{f\}] \neq F(V^*, J^2, N). \tag{2.11}$$

It has to be noticed that if n different operators $X[\{f^{(1)}\}]$, $X[\{f^{(2)}\}]$, ..., $X[\{f^{(n)}\}]$, are found to satisfy (2.8) and (2.11)—this is *a priori* expected in higher multipole phonon-state classification problems whereby three or more labels are missing[9-11]—two supplementary conditions have to be fulfilled: No functional relationship may exist between these n operators, V^*, J^2, and N,

$$\forall F: F(X[\{f^{(1)}\}], \ldots, X[\{f^{(n)}\}], V^*, J^2, N) \neq 0, \tag{2.12}$$

and the operators have to commute mutually,

$$[X[\{f^{(i)}\}], X[\{f^{(j)}\}]] = 0 \quad (i,j = 1, 2, \ldots, n). \tag{2.13}$$

With the help of (2.5) the basic condition (2.8) can be rewritten as:

$$\sum_{N=2}^{N_{\max}} \sum_{J_N} \sum_{\substack{I_i J_i \\ (i=1,\ldots,N-1)}} f^{I_2 \ldots J_N}_{J_2 \ldots J_N} [O^0_0, O^{I_2 \ldots J_N}_{J_2 \ldots J_N}] = 0. \tag{2.14}$$

On the other hand, by using (2.4), (2.6) it follows from (2.9), (2.10) that

$$\sum_J [O^J, O^{I_2 \ldots J_N}_{J_2 \ldots J_N}] \equiv 0, \tag{2.15}$$

$$\sum_J \begin{Bmatrix} \lambda & \lambda & J \\ \lambda & \lambda & 1 \end{Bmatrix} [O^J, O^{I_2 \ldots J_N}_{J_2 \ldots J_N}] \equiv 0. \tag{2.16}$$

In the particular quadrupole problem the canonical operators of the form O^J_J are O^0_0, O^2_2, O^4_4, and therefore the basic condition (2.14) may be replaced by one of the three equations

$$\sum_{N=2}^{N_{\max}} \sum_{J_N} \sum_{\substack{I_i J_i \\ (i=1,\ldots,N-1)}} f^{I_2 \ldots J_N}_{J_2 \ldots J_N} [O^J_J, O^{I_2 \ldots J_N}_{J_2 \ldots J_N}] = 0 \quad (J = 0, 2, 4), \tag{2.17}$$

the remaining two equalities becoming then merely a consequence of the implied condition. It goes however without saying that in the context of Racah algebra the choice $J = 0$ in (2.17), and corresponding to (2.14), is the most convenient one.

In the following sections, the most important step is each time the transformation of an equation of the type (2.14) or (2.17) into a set of equations in the f-variables alone, by expressing the occurring operators in terms of canonical operators which belong to an operator basis. Since such operations necessarily require the knowledge of all existing relationships between the canonical operators, and therefore are quite involved, we have invoked computer assistance for numerical treatment with an accuracy up to four decimal positions.

3. THE CASE $N_{\max} = 2$

In this section we consider in particular the operators of the form (2.7) with $N = 2$, i.e.,

$$X[\{f\}] = \sum_{J_2} f^{J_2}_{J_2} O^{J_2}_{J_2}. \tag{3.1}$$

Since we require that X satisfies an equation of the type (2.14), or even more generally of the type (2.17) when $\lambda = 2$, it is convenient to deduce separately closed expressions for the occurring commutators. After straightforward calculation one obtains from (2.2) that

$$[O^J_J, O^{J_2}_{J_2}] = 4 \sum_{J_3} (-1)^{J+J_2} \sqrt{(2J+1)(2J_2+1)} \times \begin{Bmatrix} \lambda & J_3 & J_2 \\ \lambda & \lambda & J \end{Bmatrix} [O^{J_2 J_3}_{J J_3} - O^{J J_3}_{J_2 J_3}]. \quad (3.2)$$

Setting further $\lambda = 2$, the condition (2.17) becomes

$$\sum_{J_2} \sqrt{2J_2+1}\,(-1)^{J_2} f^{J_2}_{J_2} \sum_{J_3} \begin{Bmatrix} 2 & J_3 & J_2 \\ 2 & 2 & J \end{Bmatrix} \times [O^{J_2 J_3}_{J J_3} - O^{J J_3}_{J_2 J_3}] = 0. \quad (3.3)$$

It is clear that for any fixed J_3-value, the expression between square brackets in (3.3) vanishes identically if and only if the corresponding $(3, J_3)$-operator basis is a singleton-basis. Indeed, denoting the single basis element by $O^{K J_3}_{K J_3}$ it follows from the relations between $(3, J_3)$-quadrupole phonon states

$$|02 J J_3\rangle = \mu_{J J_3} |02 K J_3\rangle, \quad |02 J_2 J_3\rangle = \mu_{J_2 J_3} |02 K J_3\rangle,$$

whereby $\mu_{J J_3}$ and $\mu_{J_2 J_3}$ are real ratio factors, that

$$O^{J J_3}_{J_2 J_3} = O^{J_2 J_3}_{J J_3} = \mu_{J J_3} \mu_{J_2 J_3} O^{K J_3}_{K J_3}.$$

As it is well known that all quadrupole $(3, J_3)$-state bases are singleton-bases, there are nothing but singleton $(3, J_3)$-operator bases, and consequently (3.3) is satisfied for arbitrary f^J_J-values ($J = 0, 2, 4$), leaving us with three degrees of freedom:

$$X[f^0_0, f^2_2, f^4_4] = f^0_0 O^0_0 + f^2_2 O^2_2 + f^4_4 O^4_4 . \quad (3.4)$$

On the other hand, it is obvious from (2.3)–(2.6) that $X[f^0_0, f^2_2, f^4_4]$, whatever the values of f^0_0, f^2_2, and f^4_4 are, is a linear combination of N, N^2, J^2, and V^*, namely:

$$X[f^0_0, f^2_2, f^4_4] = \tfrac{1}{5} f^0_0 (N^2 + 2V^* + 3N) - \tfrac{1}{7} f^2_2 (J^2 + 4V^* + 4N - 2N^2) + \tfrac{1}{7} f^4_4 (J^2 + (18/5)N^2 + (6/5)V^* - (36/5)N). \quad (3.5)$$

It follows that $N_{\max} = 2$ does not provide us with a suitable operator for generating the fifth quadrupole classification label.

The situation is somewhat different in the octupole case. Indeed, as there exist two independent (3,3)-octupole phonon states,[9] built up in terms of the linearly independent basis states $|0303\rangle$ and $|0323\rangle$, the basic condition (2.14) becomes

$$\sum_{J_2} f^{J_2}_{J_2} \sqrt{2J_2+1} \begin{Bmatrix} 3 & 3 & 0 \\ 3 & 3 & J_2 \end{Bmatrix} [O^{03}_{J_2 3} - O^{J_2 3}_{03}] = 0, \quad (3.6)$$

a nonidentically vanishing equation which by expressing the different operators in terms of the (3,3)-operator basis elements, can be reduced into a homogeneous equation in the variables f^0_0, f^2_2, f^4_4, and f^6_6, restricting the four original degrees of freedom to three. It then turns out that the X-solutions of (2.14) can again be expressed as linear combinations of N, N^2, J^2, and V^*, making it as before necessary to consider higher $N_{\max}$-values. For details concerning the octupole state classification problem the reader is referred to a forthcoming paper.

4. THE CASE $N_{\max} = 3$

We continue our systematic search by studying in the present section the operators of the form (2.7) with $N_{\max} = 3$, i.e.,

$$X[\{f\}] = \sum_{J_2} f^{J_2}_{J_2} O^{J_2}_{J_2} + \sum_{\substack{I_2, J_2 \\ J_3}} f^{I_2 J_3}_{J_2 J_3} O^{I_2 J_3}_{J_2 J_3} . \quad (4.1)$$

In addition to the expression (3.2) for the commutator of two canonical operators belonging to $(2, J)$-operator bases, we need an analogous expression for the commutator of two canonical operators which are respectively elements of a $(2, J)$- and a $(3, J')$-operator basis. It is again straightforward to obtain that

$$\begin{aligned}
&[O^J_J, O^{I_2 J_3}_{J_2 J_3}] \\
&= +2(\delta_{J J_2} - \delta_{J I_2}) O^{J_2 J_3}_{I_2 J_3} + 4 \Bigg((-1)^J \hat{J} \hat{J}_2 \begin{Bmatrix} \lambda & J_3 & J \\ \lambda & \lambda & J_2 \end{Bmatrix} O^{I_2 J}_{J J_3} \\
&\quad - (-1)^{I_2} \hat{J} \hat{I}_2 \begin{Bmatrix} \lambda & J_3 & J \\ \lambda & \lambda & I_2 \end{Bmatrix} O^{J J_3}_{J_2 J_3} \Bigg) \\
&\quad + 4 \sum_{J' J_4} \Bigg((-1)^{J+J_2+J_3+J'} \hat{J} \hat{J}_2 \hat{J}_3 \hat{J}' \\
&\quad \times \begin{Bmatrix} \lambda & J' & J \\ \lambda & \lambda & J_2 \end{Bmatrix} \begin{Bmatrix} \lambda & J' & J_4 \\ \lambda & J_3 & J_2 \end{Bmatrix} O^{I_2 J_3 J_4}_{J J' J_4} \\
&\quad - (-1)^{J+I_2+J_3+J'} \hat{J} \hat{I}_2 \hat{J}_3 \hat{J}' \\
&\quad \times \begin{Bmatrix} \lambda & J' & J \\ \lambda & \lambda & I_2 \end{Bmatrix} \begin{Bmatrix} \lambda & J' & J_4 \\ \lambda & J_3 & I_2 \end{Bmatrix} O^{J J' J_4}_{J_2 J_3 J_4} \Bigg) \\
&\quad + 2 \sum_{J' J_4} \Bigg((-1)^{J+J_4} \hat{J} \hat{J}_2 \hat{J}_3 \hat{J}' \begin{Bmatrix} J_2 & J_4 & J \\ \lambda & \lambda & J_3 \end{Bmatrix} \\
&\quad \times \begin{Bmatrix} J & J_4 & J_2 \\ \lambda & \lambda & J' \end{Bmatrix} O^{I_2 J_3 J_4}_{J J' J_4} - (-1)^{I_2+J_4} \hat{J} \hat{I}_2 \hat{J}_3 \hat{J}' \\
&\quad \times \begin{Bmatrix} I_2 & J_4 & J \\ \lambda & \lambda & J_3 \end{Bmatrix} \begin{Bmatrix} J & J_4 & I_2 \\ \lambda & \lambda & J' \end{Bmatrix} O^{J J' J_4}_{J_2 J_3 J_4} \Bigg),
\end{aligned} \quad (4.2)$$

whereby we have introduced the shorthand notation

$$\hat{J} = \sqrt{2J+1} . \quad (4.3)$$

Since in (4.1) the multiple summations are restricted to the angular momentum values for which the associated $O^{I_2 J_3}_{J_2 J_3}$ operators are elements of a $(3, J_3)$-operator basis, one finds by careful inspection of the previously tabulated $(3, J_3)$-quadrupole state bases,[9] that $I_2 = J_2$, and with this restriction the first two expressions between round large braces in the r.h.s. of (4.2) vanish identically, as may be seen from a similar argument as the one following formula (3.3). Consequently, at this stage we can already conclude that there will not exist relations connecting $f^{J_2}_{J_2}$-coefficients with $f^{J_2 J_3}_{J_2 J_3}$-coefficients.

Furthermore, looking at the remaining expressions in (4.2) it is seen by taking into account that J is even, that the only nonzero contributions correspond to such J_4-values wherefore the $(4, J_4)$-operator basis is not a singleton basis.

TABLE I. The operators $O^J_J O^{J'}_{J'}$ expressed in canonical form.

	$O^0_0 O^0_0$	$O^0_0 O^2_2$	$O^0_0 O^4_4$	$O^2_2 O^2_2$	$O^2_2 O^4_4$	$O^4_4 O^4_4$
f^{20}_{20}	0	0	0	4.0	0	0
$f^{02}_{02} = x$	0.8	1.1429	2.0571	−0.3499	0.8396	0.0420
$f^{23}_{23} = y$	0	0	0	2.2857	1.7144	−0.1143
$f^{24}_{24} = z$	0	0	0	1.1429	2.8573	0.7799
f^{46}_{46}	0	0	0	0	0	4.0
$f^{020}_{020} = e$	1.0	0	0	0.4082	0	0.7347
$f^{022}_{022} = g$	0	1.0	0	0.4082	0.0204	1.1874
f^{202}_{202}	0	0	0	1.6326	0.5102	0.0742
$f^{024}_{024} = h$	0	0	1.0	0.4082	0.3711	0.4162
$f^{234}_{234} = c$	0	0	0	1.4286	0.2078	0.7992
f^{235}_{235}	0	0	0	0	1.3333	0
$f^{246}_{246} = d$	0	0	0	0	1.0	0.8
f^{468}_{468}	0	0	0	0	0	1.0
$f^{202}_{022} = a$	0	0	0	−0.8163	0.1020	−0.2968
$f^{022}_{202} = a$	0	0	0	−0.8163	0.1020	−0.2968
$f^{234}_{024} = b$	0	0	0	−0.7636	0.2777	−0.5767
$f^{024}_{234} = b$	0	0	0	−0.7636	0.2777	−0.5767

This is only the case for $J_4 = 2$ and $J_4 = 4$,[7,9] showing us also that J_3 may take any of its physical acceptable values between 0 and 6.

From the foregoing remarks, it follows that (4.1) can without loss of generality be restricted to:

$$X[f^{20}_{20}, f^{02}_{02}, f^{23}_{23}, f^{24}_{24}, f^{46}_{46}] = f^{20}_{20} O^{20}_{20} + f^{02}_{02} O^{02}_{02} + f^{23}_{23} O^{23}_{23} + f^{24}_{24} O^{24}_{24} + f^{46}_{46} O^{46}_{46}, \tag{4.4}$$

giving us *a priori* five degrees of freedom. It is now only a matter of straightforward calculation to find that the condition (2.14) is equivalent to the following system of equations:

$$14 f^{20}_{20} - 25 f^{23}_{23} + 11 f^{24}_{24} = 0,$$
$$55 f^{23}_{23} - 55 f^{24}_{24} + 28 f^{46}_{46} = 0. \tag{4.5}$$

The operators X of the form (4.4) which satisfy the condition (2.8), can therefore be written as

$$X[f^{20}_{20}, f^{02}_{02}, f^{23}_{23}, f^{24}_{24}, f^{46}_{46}] = X\left|\, (25y - 11z)/14, x, y, z, -\tfrac{55}{28}(y - z)\right| \quad (x,y,z \in R), \tag{4.6}$$

leaving us at this point with three degrees of freedom. As this is exactly the number required for explaining the existence in canonical form of the three operators N^3, J^2N, and V^*N, we can anticipate that the case $N_{\max} = 3$ does not leave room for a new label generating operator. For the sake of completeness, however, and to obtain at the same time a check proof of the validity of our method and of the result (4.6), we calculate consecutively from (2.3)–(2.6):

$$N^3 = N^2 + \sum_J O^J_J O = N^2 + 2\sum_J O^J_J + \sum_{JJ'} O^{JJ'}_{JJ'}$$
$$= 3N^2 - 2N + O^{20}_{20} + 2.1429\, O^{02}_{02} + 1.4\, O^{23}_{23} + 1.9092\, O^{24}_{24} + O^{46}_{46}, \tag{4.7}$$

$$J^2N = 6N^2 + 30 \sum_J \begin{Bmatrix} 2 & 2 & J \\ 2 & 2 & 1 \end{Bmatrix} O^J_J O$$
$$= 6N^2 - 12N + 2J^2 + 30 \sum_{JJ'} \begin{Bmatrix} 2 & 2 & J \\ 2 & 2 & 1 \end{Bmatrix} O^{JJ'}_{JJ'}$$
$$= 6N^2 - 12N + 2J^2 - 3 O^{20}_{20} - 4.2857\, O^{02}_{02} - 1.4\, O^{23}_{23} + 0.6367\, O^{24}_{24} + 4\, O^{46}_{46}, \tag{4.8}$$

$$V^*N = -2N^2 + \tfrac{5}{2} O^0_0 N - \tfrac{1}{2}(N^2 - N)N$$
$$= -0.5 N^3 - 1.5 N^2 + 5 O^0_0 + 2.5 O^{02}_{02}$$
$$= -0.5 N^3 - 0.5 N^2 + 3N + 2V^* + 2.5 O^{02}_{02}, \tag{4.9}$$

whereby the intermediate summations with respect to J and J' run over all possible values, and are thus not restricted to operator-basis values alone. The reader can now easily verify that the parts of the right-hand sides of (4.7), (4.8), and (4.9), which can be identified to operators of the general form (4.4), indeed satisfy (4.6), showing that the condition (2.11) is not fulfilled as we have stated before. Finally, it is worthwhile to note that in accordance to some general remarks given in Sec. 2, it turns out that the numerical evaluation of the extended condition (2.17) for $J = 2$ and $J = 4$ leads equally well in both cases to the solution (4.6).

The fact that at the present level no new label generating operator can be constructed is certainly not very surprising. From group-theoretical studies[3,7] one knows that the quadrupole-phonon states can be classified using the chain

$$U(5) \supset SU(5) \supset R(5) \supset R(3) \supset R(2).$$

The physical operators N, V^*, J^2, and J_0 can all be expressed in terms of the generators $(b^+ b)^{(2k+1)}$ $(k = 0,1)$ of the R(5) Lie algebra. The fifth invariant, we look for, should also be of

TABLE II. The operators $N^3, J^2N, V^*N, N^4, J^4, V^{*2}, J^2N^2, V^*N^2, J^2V^*$ expressed in canonical form.

	N^3	J^2N	V^*N	N^4	J^4	V^{*2}	J^2N^2	V^*N^2	J^2V^*
x	2.1429	−4.2857	1.4286	12.8571	−102.8557	−8.5715	− 8.5712	2.8571	25.7145
y	1.4	−1.4	−0.7	8.4	− 39.2	4.2	1.4	−6.3	1.4
z	1.9092	0.6367	−0.9546	11.4557	− 38.1695	5.7277	14.6402	−8.5916	− 8.2751
a	0	0	0	− 0.9091	− 14.5454	−0.2273	1.3637	0.4545	− 0.6818
b	0	0	0	− 0.7850	− 22.7640	−0.1962	0.2617	0.3925	− 0.1309
c	0	0	0	2.6434	20.6574	0.6608	− 0.8811	−1.3217	0.4406
d	0	0	0	2.8	− 11.2	0.7	4.2	−1.4	− 2.1
e	0	0	0	2.1429	51.4286	4.2857	− 4.2858	1.4286	−12.8571
g	0	0	0	3.6364	58.1816	−1.5909	− 5.4546	0.6818	− 4.2727
h	0	0	0	3.5665	− 46.5738	−1.6084	− 1.1888	0.7168	12.3776

that form. It is clear that an operator of the type (4.1) can be brought after some rearrangements to an expression built up by the just-mentioned generators. Nwachuku and Rashid[12] have noted that such invariants of odd order p in these generators are not independent and can be expressed in terms of those of even order $2q$ with $2q < p$. The relations (4.7)–(4.9) are examples of this statement.

5. The case $N_{max} = 4$

Continuing on the same lines of investigation as in the preceding sections, we start here from the operators of the form:

$$X[\{f\}] = \sum_{J_2} f^{J_2}_{J_2} O^{J_2}_{J_2} + \sum_{\substack{I_2 J_2 \\ J_3}} f^{I_2 J_3}_{J_2 J_3} O^{I_2 J_3}_{J_2 J_3} + \sum_{\substack{I_2 I_3 J_2 J_3 \\ J_4}} f^{I_2 I_3 J_4}_{J_2 J_3 J_4} O^{I_2 I_3 J_4}_{J_2 J_3 J_4} . \tag{5.1}$$

The supplementary general commutator needed for expressing the condition (2.14) in terms of a set of independent operators, is given by

$$\begin{aligned}
&[O^{J}_{J}, O^{I_2 I_3 J_4}_{J_2 J_3 J_4}] \\
&= \left(2\delta_{JJ_2} + 4(-1)^{J+J_2} \hat{J}\hat{J}_2 \begin{Bmatrix} \lambda & J_3 & J \\ \lambda & \lambda & J_2 \end{Bmatrix}\right) O^{I_2 I_3 J_4}_{J J_3 J_4} - \left(2\delta_{JI_2} + 4(-1)^{J+I_2} \hat{J}\hat{I}_2 \begin{Bmatrix} \lambda & I_3 & J \\ \lambda & \lambda & I_2 \end{Bmatrix}\right) O^{J I_3 J_4}_{J_2 J_3 J_4} \\
&\quad + \sum_{J'} \hat{J}\hat{J}'\hat{J}_2\hat{J}_3 \left(2(-1)^{J_4} \begin{Bmatrix} J & J_4 & J_2 \\ \lambda & \lambda & J' \end{Bmatrix} \begin{Bmatrix} J_2 & J_4 & J \\ \lambda & \lambda & J_3 \end{Bmatrix} + 4(-1)^{J+J_2+J_3+J'} \right. \\
&\quad \left. \times \begin{Bmatrix} \lambda & J' & J_2 \\ \lambda & \lambda & J \end{Bmatrix} \begin{Bmatrix} \lambda & J' & J_4 \\ \lambda & J_3 & J_2 \end{Bmatrix}\right) O^{I_2 I_3 J_4}_{J J' J_4} \\
&\quad - \sum_{J'} \hat{J}\hat{J}'\hat{I}_2\hat{I}_3 \left(2(-1)^{J_4} \begin{Bmatrix} J & J_4 & I_2 \\ \lambda & \lambda & J' \end{Bmatrix} \begin{Bmatrix} I_2 & J_4 & J \\ \lambda & \lambda & I_3 \end{Bmatrix} + 4(-1)^{J+I_2+I_3+J'} \begin{Bmatrix} \lambda & J' & I_2 \\ \lambda & \lambda & J \end{Bmatrix} \begin{Bmatrix} \lambda & J' & J_4 \\ \lambda & I_3 & I_2 \end{Bmatrix}\right) O^{J J' J_4}_{J_2 J_3 J_4} \\
&\quad + \sum_{J_5 J' J''} \hat{J}\hat{J}_2\hat{J}_3\hat{J}_4\hat{J}'\hat{J}'' \left(2(-1)^{J_4} \begin{Bmatrix} J & J'' & J_2 \\ \lambda & \lambda & J' \end{Bmatrix} \begin{Bmatrix} J_2 & J'' & J \\ \lambda & \lambda & J_3 \end{Bmatrix} \begin{Bmatrix} \lambda & J_3 & J'' \\ \lambda & J_5 & J_4 \end{Bmatrix} + 4(-1)^{J+J_2+J_3+J_4+J'+J''} \right. \\
&\quad \left. \times \begin{Bmatrix} \lambda & J' & J_2 \\ \lambda & \lambda & J \end{Bmatrix} \begin{Bmatrix} \lambda & J'' & J' \\ \lambda & J_2 & J_3 \end{Bmatrix} \begin{Bmatrix} \lambda & J'' & J_5 \\ \lambda & J_4 & J_3 \end{Bmatrix}\right) O^{I_2 I_3 J_4 J_5}_{J J' J'' J_5} \\
&\quad - \sum_{J_5 J' J''} \hat{J}\hat{I}_2\hat{I}_3\hat{J}_4\hat{J}'\hat{J}'' \left(2(-1)^{J_4} \begin{Bmatrix} J & J'' & I_2 \\ \lambda & \lambda & J' \end{Bmatrix} \begin{Bmatrix} I_2 & J'' & J \\ \lambda & \lambda & I_3 \end{Bmatrix} \begin{Bmatrix} \lambda & I_3 & J'' \\ \lambda & J_5 & J_4 \end{Bmatrix} + 4(-1)^{J+I_2+I_3+J_4+J'+J''} \right. \\
&\quad \left. \times \begin{Bmatrix} \lambda & J' & I_2 \\ \lambda & \lambda & J_4 \end{Bmatrix} \begin{Bmatrix} \lambda & J'' & J' \\ \lambda & I_2 & I_3 \end{Bmatrix} \begin{Bmatrix} \lambda & J'' & J_5 \\ \lambda & J_4 & I_3 \end{Bmatrix}\right) O^{J J' J'' J_5}_{J_2 J_3 J_4 J_5} \\
&\quad + 2 \sum_{J_5 J'} \hat{J}^2\hat{J}_4\hat{J}' \left(\begin{Bmatrix} J_3 & J_5 & J \\ \lambda & \lambda & J' \end{Bmatrix} \begin{Bmatrix} J_3 & J_5 & J \\ \lambda & \lambda & J_4 \end{Bmatrix} O^{I_2 I_3 J_4 J_5}_{J_2 J_3 J' J_5} - \begin{Bmatrix} I_3 & J_5 & J \\ \lambda & \lambda & J' \end{Bmatrix} \begin{Bmatrix} I_3 & J_5 & J \\ \lambda & \lambda & J_4 \end{Bmatrix} O^{I_2 I_3 J' J_5}_{J_2 J_3 J_4 J_5} \right).
\end{aligned} \tag{5.2}$$

By taking into account (3.2) and (4.2), the expression (5.2) clearly demonstrates the property that there cannot exist relations connecting $f^{I_2 I_3 J_4}_{J_2 J_3 J_4}$-coefficients with $f^{J_2}_{J_2}$-coefficients, so that without loss of generality the first operator sum in (5.1) may be omitted. On the other hand it is expected that from the condition (2.14) there will result equations containing both the $f^{I_2 J_3}_{J_2 J_3}$- and the $f^{I_2 I_3 J_4}_{J_2 J_3 J_4}$-variables. It follows then by inspection of the list of selected $(4,J_4)$ basis states,[9] that X in (5.1) can be restricted to the following particular form:

$$\begin{aligned}
X[\{f\}] = {} & f^{20}_{20} O^{20}_{20} + f^{02}_{02} O^{02}_{02} + f^{23}_{23} O^{23}_{23} + f^{24}_{24} O^{24}_{24} + f^{46}_{46} O^{46}_{46} + f^{020}_{020} O^{020}_{020} + f^{022}_{022} O^{022}_{022} + f^{202}_{202} O^{202}_{202} + f^{202}_{022} O^{202}_{022} \\
& + f^{022}_{202} O^{022}_{202} + f^{024}_{024} O^{024}_{024} + f^{234}_{234} O^{234}_{234} + f^{234}_{024} O^{234}_{024} + f^{024}_{234} O^{024}_{234} + f^{235}_{235} O^{235}_{235} + f^{246}_{246} O^{246}_{246} + f^{468}_{468} O^{468}_{468} .
\end{aligned} \tag{5.3}$$

After substituting (5.3) in (2.14) we also learn from (5.2) that the part of X containing the $(4,J_4)$ basis operators, gives rise to terms which, by reducing the occurring operators firstly to basis operators, do not vanish identically when the operators belong to singleton-bases, unless either

$$I_2 = J_2 \text{ and } I_3 = J_3, \tag{5.4}$$

or

$$f^{I_2 I_3 J_4}_{J_2 J_3 J_4} = f^{J_2 J_3 J_4}_{I_2 I_3 J_4}, \tag{5.5}$$

if (5.4) is not satified. In contrast to the operators treated in the previous sections, whereby a condition of the type (5.4) was generally fulfilled, we can now identify in X four constituant operators for which (5.4) does not hold. However, if one invokes the requirement that X should be a Hermitian operator, it follows as an initial condition that

$$f^{202}_{022} = f^{022}_{202}, \quad f^{234}_{024} = f^{024}_{234}. \tag{5.6}$$

Although (5.6) would have led us to considerable simplifications in the calculations, we have preferred not to impose in advance this condition. In doing so we have created the opportunity to observe that (5.6) follows naturally from (2.14), showing that at the present level the Hermiticity of X is merely a consequence of the commutator equality (2.8).

Indeed, after straightforward calculations, we have found that (2.14) is equivalent to the following system of independent equations:

$$\begin{aligned}
&-0.7667\, f^{202}_{202} + 0.5476\, f^{234}_{234} - 2.9999\, f^{022}_{202} + 1.8441\, f^{024}_{234} = 0,\\
&-1.2\, f^{202}_{202} + 0.7481\, f^{234}_{234} + 1.3364\, f^{235}_{235} - 1.1570 f^{246}_{246} - 0.8\, f^{022}_{202} - 0.3110\, f^{024}_{234} = 0,\\
&-0.4277\, f^{234}_{234} - 0.6414\, f^{235}_{235} + 1.1662\, f^{246}_{246} - 1.0526\, f^{468}_{468} - 0.8\, f^{024}_{234} = 0,\\
&f^{234}_{024} = f^{024}_{234},\\
&f^{202}_{022} = f^{022}_{202},\\
&-1.2\, f^{20}_{20} + 2.1429\, f^{23}_{23} - 0.9429\, f^{24}_{24} - 3.6\, f^{022}_{202} - 1.2\, f^{202}_{202} = 0,\\
&-1.0690\, f^{23}_{23} + 1.0690\, f^{24}_{24} - 0.5442\, f^{46}_{46} - 3.6\, f^{024}_{234} - 1.0690\, f^{234}_{234} = 0,
\end{aligned} \tag{5.7}$$

thus reducing the original degrees of freedom from 17 in number to ten. Consequently, the $X[\{f\}]$ solutions of (2.14) can be written as

$$\begin{aligned}
&X\left[f^{20}_{20}, f^{02}_{02}, f^{23}_{23}, f^{24}_{24}, f^{46}_{46}; f^{020}_{020}, f^{022}_{022}, f^{202}_{202}, f^{202}_{022}, f^{022}_{202}, f^{024}_{024}, f^{234}_{234}, f^{234}_{024}, f^{024}_{234}, f^{235}_{235}, f^{246}_{246}, f^{468}_{468}\right]\\
&\quad = X\left[\frac{25y-11z}{14} - 2.4053b - 0.7143c, x, y, z, -\frac{55}{28}(y-z) - 6.6147b - 1.9643c;\ \text{e.g.},\right.\\
&\qquad -3a + 2.4053b + 0.7143c, a, a, h, c, b, b, -2.0952a + 2.3926b + 0.0816c + 0.8658d,\\
&\qquad \left. d, 1.2768a - 2.2180b - 0.4559c + 0.5804d\right] \quad (x,y,z,a,b,c,d,e,g,h \in R).
\end{aligned} \tag{5.8}$$

Since of the remaining ten degrees of freedom only nine are needed to explain that the nine independent operators N^3, J^2, V^*N, $N^4, J^2N^2, V^*N^2, J^4, V^{*2}$, and J^2V^*— the order in which the operators N, V^* and J^2 occur in the products is irrelevant since they commute—are solutions of the form (5.8), we have at first sight obtained the possibility to construct a tenth independent operator satisfying both the conditions (2.8) and (2.11). In order to find such an operator explicitly in terms of $(3,J_3)$- and $(4,J_4)$-basis operators, we must first express the nine mentioned trivial operators in canonical form.

As it is obvious that by choosing

$$a = b = c = d = e = g = h = 0, \tag{5.9}$$

the solutions (5.8) become equivalent to the solutions (4.6), it follows from the results of the preceding section that the three operators N^3, J^2N, and V^*N have the form (5.8). On the other hand it is easily seen by taking into account the relation (2.3)–(2.6), that the six other operators $N^4, J^2N^2, V^*N^2, J^{44}, V^{*2}$, and J^2V^* contain a part which is quadratic in the O^J_J-operators.

Moreover, in the quadrupole case these parts are independent linear combinations of the six operator products

$$O^0_0O^0_0,\ O^0_0O^2_2,\ O^0_0O^4_4,\ O^2_2O^2_2,\ O^2_2O^4_4,\ O^4_4O^4_4, \tag{5.10}$$

and it is therefore convenient to bring first these products into canonical form.

In doing so, we find them to be of the form (5.3), whereby, the numerical values of the 17 coefficients are listed in Table I. The reader can easily convince himself that the equations (5.7) are satisfied when substituting the values of each of the six columns, showing that the six operator combinations (5.10) are indeed of the prescribed form (5.8).

It is now a matter of straightforward calculation to obtain also the values which take the ten independent parameters $x, y, z, a, b, c, d, e, g$, and h for $N^4, J^2N^2, V^*N^2, J^4, V^{*2}$, and V^*N^2. These values are given in Table II, whereby we have added for completeness the parameter values for N^3, J^2N, and V^*N too. It has to be noticed that by giving these coefficients the nine operators are only determined upon terms of the form N^2, J^2, V^* and N.

Any operator of the form (5.8) which is independent of the nine operators listed at the head of Table II, and satisfies the basic conditions (2.8) and (2.11), thus corresponds to a tenth column of values such that the determinant of the so obtained square (10,10)-matrix is different from zero. Obviously, there exists an infinity of such operators. However, if we invoke for the selection of one particular operator, the criterion that the number of canonical basis operators contained in the operator should be minimal, we must naturally arrive at the choice

$$d = 1, \quad a = b = c = e = g = h = x = y = z = 0, \tag{5.11}$$

giving as fifth label generating operator for the complete classification of the quadrupole phonon states

$$S = 0.8658O^{235}_{235} + O^{246}_{246} + 0.5804O^{468}_{468}. \tag{5.12}$$

6. CONCLUSION

Taking as a starting point a previously published[9] classification scheme for quadrupole phonon states, we have developed an efficient method for searching systematically operators which commute with N, J^2, J_0, and V^*, and which are independent of these four operators, which themselves provide physically interpretable labels. By carrying out this program, it turned out that there exists at least one operator S, given in (5.12), which contains four phonon creation and four annihilation operators and which fulfills all the requirements indicated. However, it is evident that S is not the unique operator having the mentioned properties since all polynomial combinations of S, V^*, N, and J^2 might do as well, but it is undoubtedly the simplest one. At present, we cannot assure that the quadrupole-phonon states which have been selected previously, are all eigenstates of the new operator S. Indeed, the basic conditions which led us to an operator such as S, only guarantee us that it is possible to construct a set of quadrupole-phonon states which are simultaneously eigenstates of the five operators N, J^2, J_0, V^*, and S, whereby the five corresponding eigenvalues provide us with five labels which can serve for uniquely identifying the eigenstates, which on their turn can then be used as good orthonormalized quadrupole-phonon states.

Since at first sight there is no indication that the fifth label associated to S, is in some way related to the number of phonon triplets coupled to zero angular momentum occurring in a state, previously used as fifth label[2,3] one should find a possibility to explore the physical content of the label generated by the S-operator. We are currently investigating this new problem and hope to report on it in the near future.

Another important feature of our technique described in the present paper, is that it can be used without relevant alterations to investigate the octupole-phonon state classification problem which in contrast to the quadrupole problem has not yet been solved in a completely satisfying way. Indeed, up to now the complete classification of octupole-phonon states has necessitated the introduction of a label associated to an arbitrary ordering of the states,[11] therefore deprived from any physical or group-theoretical meaning. We have good reasons to believe that our new mechanism of investigation offers the advantage of overcoming such serious basic difficulties.

[1] A. Bohr, Kgl. Dan. Vid. Selsk. Mat. Fys. Medd. **26**, 14 (1952).
[2] A. Arima and F. Iachello, Ann. Phys. (N.Y.) **99**, 253 (1976).
[3] E. Chacon, M. Moshinsky, and R.T. Sharp, J. Math. Phys. **17**, 668 (1976).
[4] N. Kemmer, D.L. Pursey, and S.A. Williams, J. Math. Phys. **9**, 1224 (1968).
[5] L. Von Bernus, A. Kappatsh, V. Rezwani, W. Scheid, V. Schneider, M. Sedlmayr, and R. Sedlmayr: in *Heavy-Ion, High Spin States and Nuclear Structure* (IAEA, Vienna, 1975), Vol. 2, p. 3.
[6] T.M. Corrigan, F.J. Margetan, and S.A. Williams, Phys. Rev. C **14**, 2279 (1976).
[7] H.J. Weber, M.G. Huber, and W. Greiner, Z. Phys. **190**, 25 (1966).
[8] S.A. Williams and D.L. Pursey, J. Math. Phys. **9**, 1230 (1968).
[9] G. Vanden Berghe and H. De Meyer, Nucl. Phys. A **323**, 302 (1979).
[10] H. Weyl, *The Classical Groups* (Princeton University, Princeton, New Jersey, 1946), Chapter 5.
[11] S.G. Rohozinski, J. Phys. G: Nucl. Phys. **4**, 1075 (1978).
[12] C.O. Nwachuku and M.A. Rashid, J. Math. Phys. **18**, 1387 (1977).

Inverse scattering. I. One dimension

Roger G. Newton
Physics Department, Indiana University, Bloomington, Indiana 47405

(Received 31 July 1979; accepted for publication 19 October 1979)

This paper presents two new methods of reconstructing an underlying potential in the one-dimensional Schrödinger equation from a given S matrix. One of these methods is basd on a Gel'fand-Levitan equation, the other on a Marchenko equation. A sequel of this paper will treat the three-dimensional case by similar methods.

1. INTRODUCTION

The inverse scattering problem for the Schrödinger equation has a rich history that need not be recounted here,[1] nor is there a need to describe its remarkable application over the last ten years to the solution of nonlinear evolution equations. If we confine our attention to potentials without symmetry properties, then the history becomes very much less rich, but most applications, both to solitons and to such fields as seismology, among others, are in that area. Solutions of the one-dimensional inverse problem go back to Kay and Moses.[2] A Marchenko-like solution was given by Faddeev,[3,4,5] and Kay,[6] and a quite different, nonlinear equation was recently introduced by Deift and Trubowitz.[7] The three-dimensional case (without spherical symmetry) resisted solution for many years. The first method was given by Kay and Moses,[8] but it contained certain flaws. The next step was taken by Faddeev,[5] who utilized a new Green's function for the Schrödinger equation, and independently on the basis of Faddeev's Green's function, by the present author.[9] The methods of Ref. 9 are relatively cumbersome and still contain important loose ends.

The present paper is the first in a series of two that treat both the one-dimensional and the three-dimensional inverse scattering problem by new methods that are analogous to the Gel'fand–Levitan and Marchenko procedures in the radial case. Part of the motivation for this first paper of the series is to set the stage in one dimension for similar methods in three dimensions. It does not purport to be simpler than the method of F. Whether its ideas are more fruitful than those of F or of D & T because they lead to three-dimensional extensions while the latter may not remains to be seen. In any event, they may be useful in their own right.

This paper presents two inversion methods for the one-dimensional Schrödinger equation. One is based on the introduction of a Jost matrix and its subsequent use in a procedure that is quite analogous to that of Gel'fand and Levitan. While Faddeev[5] calls his method a Gel'fand–Levitan procedure, this is a misnomer. His is really a Marchenko theory, as D & T recognize. The Gel'fand–Levitan theory given here is therefore new.

The second method given in this paper is a Marchenko theory that differs from Faddeev's in utilizing the entire S matrix rather than a reflection amplitude alone. While this may seem like a step backwards, it is the exact one-dimensional analog of a procedure that has been recently found to work in the three-dimensional case.[10]

It is a characteristic difference between the well worked-out radial scattering problem (i.e., on the half-line $0<r<\infty$) and the one-dimensional one (i.e., on the full line $-\infty<x<\infty$) that the S matrix in the former case is a complex number of unit modulus, and in the latter it is a unitary 2×2 matrix. The basic reason for this is the fact that in the half-line case the spectrum of the Schrödinger equation is simple, while in the full-line case it has the multiplicity 2. For the half-line equation, an acceptable solution has to vanish at the origin or to satisfy another boundary condition there[11]; for the full-line equation, both linearly independent solutions are acceptable. This makes the full one-dimensional case formally more complicated and leads towards the three-dimensional case, in which the spectrum has infinite multiplicity.

Section 2 deals with the direct problem only. Much of what it contains is not new but needed subsequently. The use of two "regular" solutions that are defined at a finite point is new. They are needed for the Gel'fand–Levitan method analogously to the radial case. It is their use that leads to the definition of a Jost matrix in terms of which the S matrix has a canonical product representation.[12]

Section 3 deals with bound states and their appearance in the Jost matrix. Because of the multiplicity of the spectrum there are *two* real numbers that characterize a bound-state eigenfunction: One is its normalization and the other may be taken to be its logarithmic derivative at the origin. However, once the residue of the transmission amplitude is known, these two numbers are not independent. Furthermore, in spite of the spectral multiplicity, there can be no bound-state degeneracy.

Section 4 defines the spectral function for the "regular solution" and Sec. 5 presents the Povsner–Levitan representation of that "regular solution." Section 6 discusses the "Hilbert problem" of constructing the Jost matrix from the S matrix. The method presented is quite similar to the Marchenko theory. In Sec. 7 we derive a Gel'fand–Levitan equation, and in Sec. 8, a Marchenko equation. In Sec. 9 we discuss the inversion procedures based on the equations of Secs. 6–8. There are two appendices. Appendix A shows that the determinant of the Jost matrix equals the Fredholm determinant of the integral equation for the "physical" solution. In Appendix B we discuss the effects of shifting the potential.

2. SCATTERING SOLUTIONS AND THE S-MATRIX

It will be useful to introduce several different solutions

 0022-2488/80/030493-13$1.00

of the Schrödinger equation

$$-y'' + V(x)\,y = k^2 y \tag{2.1}$$

that satisfy different boundary conditions. The first pair of solutions, the Jost solutions, satisfy the boundary conditions

$$\lim_{x\to\infty} f_1(x)\,e^{-ikx} = 1, \tag{2.2}$$

$$\lim_{x\to-\infty} f_2(x)\,e^{ikx} = 1. \tag{2.2'}$$

Hence, they solve the integral equations

$$f_1(x) = e^{ikx} - k^{-1}\int_x^\infty dy\,\sin k(x-y)\,V(y)f_1(y), \tag{2.3}$$

$$f_2(y) = e^{-ikx} + k^{-1}\int_{-\infty}^x dy\,\sin k(x-y)\,V(y)f_2(y), \tag{2.3'}$$

respectively. Assuming that the potential $V(x)$ obeys the restriction[13]

$$\int_{-\infty}^\infty dx(1+x^2)|V(x)| < \infty, \tag{2.4}$$

the functions f_1, f_1', f_2, and f_2' are easily proved[4] to exist and to be continuous for all k in the upper half-plane and on the real axis $\mathrm{Im}k \geqslant 0$, and to be regular analytic for $\mathrm{Im}k > 0$. The function f_1 approaches e^{ikx} there for large $|k|$:

$$f_1(x)\,e^{-ikx} = 1 + O(1/|k|), \tag{2.5}$$

while

$$f_2\,e^{ikx} = 1 + O(1/|k|). \tag{2.5'}$$

The solution $f_1(x)$ describes waves incident from the left and one easily reads off from Eq. (2.3) that

$$\begin{aligned} f_1(x) &= e^{ikx} + o(1), \quad \text{as } x\to\infty, \\ &= e^{ikx}(1/T^l) + e^{-ikx}(R^l/T^l) + o(1), \\ &\qquad\qquad \text{as } x\to-\infty, \end{aligned}$$

where the transmission and reflection amplitudes for left incidence are given by

$$1/T^l = 1 + (i/2k)\int_{-\infty}^\infty dx\,e^{-ikx}\,V(x)f_1(x),$$

$$R^l/T^l = -(i/2k)\int_{-\infty}^\infty dx\,e^{ikx}\,V(x)f_1(x), \tag{2.6}$$

respectively. One readily finds[4] from the first of these equations that T^l, as a function of k, is continuous on the real axis, and it is a meromorphic function in $\mathrm{Im}k > 0$, without zeros there.

Similarly, $f_2(x)$ describes waves that are incident from the right and

$$\begin{aligned} f_2(x) &= e^{ikx}(R^r/T^r) + e^{-ikx}(1/T^r) + o(1), \quad \text{as } x\to\infty, \\ &= e^{-ikx} + o(1), \quad \text{as } x\to-\infty, \end{aligned}$$

where the transmission and reflection amplitudes for right incidence are found from Eq. (2.3'):

$$1/T^r = 1 + (i/2k)\int_{-\infty}^\infty dx\,e^{ikx}\,V(x)f_2(x),$$

$$R^r/T^r = -(i/2k)\int_{-\infty}^\infty dx\,e^{-ikx}\,V(x)f_2(x), \tag{2.7}$$

respectively.

The Wronskian of f_1 and f_2 being independent of x, we can evaluate it at $x\to-\infty$ or at $x\to+\infty$, and we find

$$W(f_1,f_2) \equiv f_1 f_2' - f_1' f_2 = -2ik/T^l = -2ik/T^r, \tag{2.8}$$

from which we may conclude that

$$T^l = T^r \equiv T. \tag{2.9}$$

We shall combine the two functions f_1 and f_2 into a two-component row vector

$$\chi_k(x) \equiv (f_1,f_2). \tag{2.10}$$

In a similar manner we form another row vector of two solutions of Eq. (2.1):

$$\psi_k(x) = (u_1,u_2) \tag{2.11}$$

that satisfies the integral equation

$$\psi_k(x) = \psi_k^0(x) - (i/2k)\int_{-\infty}^\infty dy\,e^{ik|x-y|}V(y)\psi_k(y), \tag{2.12}$$

where

$$\psi_k^0(x) = (e^{ikx}, e^{-ikx}). \tag{2.13}$$

The asymptotic forms of u_1 and u_2 for real k are found from Eq. (2.12) to be

$$\begin{aligned} u_1(x) &= e^{ikx}T^l + o(1), \quad \text{as } x\to\infty, \\ &= e^{ikx} + e^{-ikx}R^l + o(1), \quad \text{as } x\to-\infty, \end{aligned} \tag{2.14}$$

$$\begin{aligned} u_2(x) &= e^{-ikx} + e^{ikx}R^r + o(1), \quad \text{as } x\to\infty, \\ &= e^{-ikx}T^r + o(1), \quad \text{as } x\to-\infty. \end{aligned} \tag{2.14'}$$

Comparison with f_1 and f_2 shows that

$$\psi_k(x) = T_k\chi_k(x). \tag{2.15}$$

The function $\psi_k(x)$ is the physical wave function that is usually denoted by ψ_k^+. We may also form

$$\psi_k^-(x) = (u_3,u_4),$$

where (for real k)[14]

$$u_3 = u_2^*, \quad u_4 = u_1^*.$$

This may be written

$$\psi_k^-(x) = \psi_{-k}^+(x)Q, \tag{2.16}$$

where

$$Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.17}$$

The two solution u_1 and u_2 of Eq. (2.1) must be expressible as linear combinations of u_3 and u_4, which defines the 2×2 S matrix:

$$\psi_k^+(x) = \psi_k^-(x)S_k, \tag{2.18}$$

or by Eq. (2.16),

$$\psi_k = \psi_{-k}QS_k. \tag{2.19}$$

The scattering amplitude, on the other hand, is defined as a 2×2 matrix A_k such that

$$\begin{aligned} \psi_k(x) - \psi_k^0(x) &= e^{ikx}(1,0)A_k + o(1), \quad \text{as } x\to\infty, \\ &= e^{-ikx}(0,1)A_k + o(1), \quad \text{as } x\to-\infty. \end{aligned} \tag{2.20}$$

One then obtains the expression[14]

$$A_k = -(i/2k)\int_{-\infty}^{\infty} dx\, \psi_k^{0\dagger} V \psi_k \tag{2.21}$$

from Eq. (2.12), in which $\psi_k^{0\dagger} = \tilde{\psi}_k^{0*}$. Insertion of Eq. (2.12) in (2.21) shows that A_k has the alternative representation

$$A_k = -(i/2k)\int_{-\infty}^{\infty} dx \psi_{k}^{-\dagger} V \psi_k^0\,, \tag{2.21'}$$

from which we may conclude by Eq. (2.16) that

$$A^{\dagger}_{-k} = QA_kQ\,. \tag{2.22}$$

If we insert the asymptotic form (2.20) in Eq. (2.19) and compare coefficients of e^{ikx} and e^{-ikx}, we obtain the relation

$$S_k = \mathbb{1} + A_k \tag{2.23}$$

between S and A, as well as the equation

$$QS_k^*QS_k = \mathbb{1}. \tag{2.24}$$

Since, furthermore, Eq. (2.12) implies that for real k

$$\psi_{-k} = \psi_k^*, \tag{2.25}$$

it follows from Eq. (2.19) that

$$S_{-k} = S_k^*. \tag{2.26}$$

However, Eqs. (2.22) and (2.23) imply

$$S^{\dagger}_{-k} = QS_kQ$$

and therefore

$$\tilde{S}_k = QS_kQ. \tag{2.27}$$

This is the *reciprocity* theorem of scattering theory. Together with Eq. (2.24) it leads to the unitarity of the S matrix

$$S_kS_k^{\dagger} = S_k^{\dagger}S_k = \mathbb{1}. \tag{2.28}$$

Comparison of Eq. (2.20) with Eqs. (2.14) and (2.14'), and the use of Eq. (2.23), show that

$$S = \begin{pmatrix} T^l & R^r \\ R^l & T^r \end{pmatrix}. \tag{2.29}$$

This shows that Eq. (2.9) is identical with Eq. (2.27). The unitarity equation (2.28), written out explicitly, then says

$$|R^l|^2 = |R^r|^2 = 1 - |T|^2, \tag{2.28'}$$

as well as

$$\frac{R^rR^l}{|R^rR^l|} = -\frac{T^2}{|T|^2}. \tag{2.28''}$$

It also follows that

$$\det S = T/T^*. \tag{2.30}$$

The function S_k is continuous[4] on the real axis and its asymptotic behavior has been established[4] to be [assuming Eq. (2.4)]

$$S_k = \mathbb{1} + O(1/k) \tag{2.31}$$

for real k; and for $\mathrm{Im}k \geqslant 0$ we have[4]

$$T_k = 1 + O(1/|k|). \tag{2.32}$$

The reflection amplitudes are known to have a further property if V is in the class (2.4). Defining the Fourier transform of each reflection amplitude by

$$F^j(x) = \int_{-\infty}^{\infty} dk\, R^j_k e^{ikx}, \quad j = l,r,$$

we have[14] for all $-\infty < a < \infty$

$$\int_a^{\infty} dx(1+x^2)|F^j(x)| \leqslant c(a) < \infty. \tag{2.33}$$

If $T_{k=0} = 0$, then it can be shown[16] that near $k = 0$

$$T_k = bk + o(k), \quad b \neq 0 \tag{2.34}$$

for $\mathrm{Im}k \geqslant 0$, and

$$R^j_k = -1 + a^jk + o(k), \quad j = r,l \tag{2.34'}$$

for $\mathrm{Im}k = 0$.

We shall refer to a 2×2 matrix-valued function (2.29) of k that has the properties (2.26)–(2.28) and (2.31)–(2.33), and if $T_0 = 0$, Eqs. (2.34) and (2.34'), that is continuous on the real axis, and whose diagonal elements are the boundary values of an analytic function regular in the upper half-plane, as in the *s class*. D & T have proved[17] that S_k is the S matrix of the equation (2.1) with a potential in class (2.4) *if and only if* it is in the *s* class.

It is useful to define a third set of solutions of Eq. (2.1) by the boundary conditions

$$g_1(0) = 1, \quad g_1'(0) = ik, \qquad g_2(0) = 1, \quad g_2'(0) = -ik, \tag{2.35}$$

and to combine them into a row vector

$$\varphi_k(x) = (g_1,g_2).$$

Then φ_k satisfies the integral equation

$$\varphi_k(x) = \psi_k^0 + (1/k)\int_0^x dy \sin k(x-y)V(y)\varphi_k(y), \tag{2.36}$$

where ψ_k^0 is defined by Eq. (2.13). It is easy to prove that if V obeys Eq. (2.4) then for each fixed x, $\varphi_k(x)$ and $\varphi_k'(x)$ are *entire* analytic functions of k.

We may express g_1 and g_2 as linear combinations of the two linearly independent solutions $f_1(x)$ and $f_2(x)$, or

$$\varphi_k(x) = \chi_k(x)F_k\,, \tag{2.37}$$

where F_k is a 2×2 matrix that is independent of x. We will get an expression for F_k by calculating the Wronskian matrix, using Eq. (2.8),

$$W(\tilde{\chi}_k,\chi_k) = -2ikP/T_k,$$

where

$$P = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.38}$$

Consequently, if we define the Jost matrix

$$J_k = F_k/T_k, \tag{2.39}$$

then

$$PW(\tilde{\chi}_k,\varphi_k) = 2ikJ_k. \tag{2.40}$$

We also find from Eq. (2.35) that

$$W(\tilde{\varphi}_k,\varphi_k) = -2ikP,$$

from which it follows that

$$\det F_k = T_k, \tag{2.41}$$

and by Eq. (2.39)

$$\det J_k = 1/T_k. \tag{2.41'}$$

Because χ_k and χ'_k are regular analytic in the upper half-plane and φ_k and φ'_k are entire analytic functions, Eq. (2.40) shows that the Jost matrix J_k is an analytic function of k, regular in the open upper half-plane and continuous on the real axis, except possibly at $k = 0$. Furthermore, it follows from Eqs. (2.5) and (2.5'), and the analogous expressions[17] for g_1 and g_2, that for $\mathrm{Im}k \geqslant 0$

$$J_k = \mathbb{1} + O(1/|k|). \tag{2.42}$$

Now the boundary conditions (2.35) imply that

$$\varphi_{-k}(x) = \varphi_k(x)Q. \tag{2.43}$$

Consequently, from Eqs. (2.15) and (2.37)

$$\psi_{-k} = \varphi_{-k}J_{-k}^{-1} = \varphi_k Q J_{-k}^{-1} = \psi_k J_k Q J_{-k}^{-1}.$$

Comparison with Eqs. (2.19) therefore leads to the decomposition of the S matrix into Jost matrices[12]

$$S_k = QJ_{-k}QJ_k^{-1}, \tag{2.44}$$

which is the exact analog of the canonical decomposition in the case of the radial Schrödinger equation.

It follows from Eq. (2.40) that for real k

$$J_{-k} = J_k^*, \tag{2.45}$$

and hence the unitarity (2.28) implies that the Jost matrix must satisfy the relation

$$Q\tilde{J}^*J = \tilde{J}J^*Q. \tag{2.46}$$

A further equation follows from Eq. (2.41'), which may be written

$$\mathrm{tr}S = 2/\det J. \tag{2.41''}$$

Let us define for any 2×2 matrix M

$$\mathring{M} = M^{-1}\det M = \begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix} \tag{2.47}$$

and insert Eq. (2.44) in (2.41''). We obtain

$$\mathrm{tr}QJ^*Q\mathring{J} = 2. \tag{2.48}$$

Finally, the Jost matrix is related to the Fredholm determinant Δ of the integral equation (2.12) by

$$\det J = \Delta, \tag{2.49}$$

which is proved in Appendix A.

3. BOUND STATES

Suppose that, for some value of k with $\mathrm{Im}k > 0$,

$$1/T_k = 0. \tag{3.1}$$

Then Eq. (2.8) shows that f_1 and f_2 are linearly dependent

$$f_2 = cf_1 \tag{3.2}$$

and Eqs. (2.2) and (2.2') show that they are equal to a function that decreases exponentially both as $x \to +\infty$ and as $x \to -\infty$. Thus, Eq. (3.1) defines an eigenvalue of the Schrödinger equation (2.1), or a bound state. It follows from the self-adjointness of Eq. (2.1) that Eq. (3.1) can hold only for $\mathrm{Im}k^2 = 0$. Furthermore, Eq. (3.1) cannot hold for real k because Eq. (2.28') implies that then $|T| \leqslant 1$.

The derivative of $1/T_k$ at $k = iK$, where Eq. (3.1) holds, is given by

$$\frac{\partial}{\partial k}(1/T_k)|_{k=iK} = -i\int_{-\infty}^{\infty} dx f_1 f_2.$$

This formula is easily derived by differentiating Eq. (2.1) with respect to k, using Eq. (2.8), and integrating. Thus, by Eq. (3.2)

$$\frac{\partial}{\partial k}(1/T_k)|_{k=iK} = -ic\int_{-\infty}^{\infty} dx f_1^2. \tag{3.3}$$

Since, at $k = iK$, f_1 is real and c cannot vanish, the zero of $(1/T_k)$ at $k = iK$ must be *simple*.

According to Eqs. (2.41') and (2.49) the bound state condition (3.1) is equivalent to

$$\Delta_k = \det J_k = 0. \tag{3.1'}$$

As the Fredholm determinant vanishes, the homogeneous form of Eq. (2.12), which we shall denote by (2.12)°, has a nontrivial solution. Conversely, if Eq. (2.1) has a normalizable solution for $k^2 = -K^2$, then Eq. (2.12)° must have a nontrivial solution and Eq. (3.1') must hold for $k = iK$. Thus, Eq. (3.1) is the necessary and sufficient condition for a bound state with the eigenvalue $k^2 < 0$.

Suppose now that Eq. (3.1') holds. Then there must exist a (real) nontrivial row vector r such that

$$J_{iK}\tilde{r} = 0. \tag{3.4}$$

Then, according to Eqs. (2.15) amd (2.37)

$$\varphi_k\tilde{r} = \psi_k J_k\tilde{r} = (\psi_k\Delta_k)(J_k\tilde{r}/\Delta_k).$$

As $k \to iK$ the left-hand side approaches a finite limit not identically zero as a function of x. We also know that $\psi_k\Delta_k$ approaches a finite limit. Since Δ_k has a simple zero at $k = iK$, $\lim J_k\tilde{r}/\Delta_k$ exists as $k \to iK$, because J_k is analytic. Hence, $\psi_k\Delta_k$ must approach a limit that is not identically zero. This limit is a solution of Eq. (2.12)°, i.e., a bound state eigenfunction:

$$\varphi_{iK}\tilde{r} = g_1 r_1 + g_2 r_2 = u_b(x).$$

It follows from the boundary conditions (2.35) that

$$\frac{u_b'(0)}{u_b(0)} = -K\frac{r_1 - r_2}{r_1 + r_2}. \tag{3.5}$$

Thus, the vector r defined by Eq. (3.4) determines the logarithmic derivative of the bound-state solution $u_b(x)$ at $x = 0$.

Suppose that there were two such linearly independent vectors r, so that there would be two linearly independent bound-state functions, i.e., the eigenvalue would be degenerate. Then it would follow from Eq. (3.4) that $J_{iK} = 0$, and hence from Eq. (2.40) that g_1 and g_2 would be linearly dependent. However, this would contradict the boundary conditions (2.35). Hence, such degeneracy cannot occur.

Let us now look at the pole of the solution ψ_k of Eq. (2.12) at $k = iK$:

$$\lim_{k \to iK}(k - iK)\psi_k = u_N d,$$

where d is a row vector $d = (d_1, d_2)$ and u_N is a normalized eigenfunction. Since the principal part of the resolvent kernel of Eq. (2.1) as a function of $E = k^2$ is

$$G(x,y) = \frac{u_N(x)u_N(y)}{E+K^2} + \cdots$$

and

$$\psi_k = \psi_k^0 + G_k V \psi_k^0 ,$$

it follows that

$$d_1 = -(i/2K)\int_{-\infty}^{\infty} dx u_N V e^{-Kx},$$

$$d_2 = -(i/2K)\int_{-\infty}^{\infty} dx u_N V e^{Kx}.$$

As a result we find from Eqs. (2.23), (2.29), and (2.21) that the residue of the transmission amplitude at the bound-state pole is given by

$$\mathrm{Res}T_{iK} = id_1 d_2. \tag{3.6}$$

Thus, the residue of T does not determine the "nature" of the eigenfunction in the sense of Eq. (3.5).[19]

We next consider the normalization of the eigenfunction. Equation (3.3) tells us that if we define

$$N^2 \equiv \int_{-\infty}^{\infty} dx f_1^2, \tag{3.7}$$

then

$$\mathrm{Res}T_{iK} = i/cN^2, \tag{3.8}$$

c being defined by Eq. (3.2). Let r be the row vector defined by Eq. (3.4), normalized so that

$$r\tilde{r} = 1.$$

Then, from Eqs. (2.37) and (2.39),

$$\varphi_k \tilde{r} = \chi_k J_k \tilde{r} T_k .$$

As $k \to iK$ we have

$$\lim_{k \to iK} J_k \tilde{r} T_k = \dot{J}_{iK} \tilde{r}\, \mathrm{Res}T_{iK} ,$$

where the dot indicates differentiation with respect to k, and

$$\chi_{iK}(x) = f_1(x)(1,c).$$

Therefore,

$$\varphi_{iK}(x)\tilde{r} = u_N(x)n, \tag{3.9}$$

with

$$n^2 = (1/c)[i(1,c)\dot{J}_{iK}\tilde{r}]^2(-i\,\mathrm{Res}T_{iK}), \tag{3.10}$$

because of Eq. (3.8). [Note that $\dot{J}_{iK}$ and $\mathrm{Res}T_{iK}$ are purely imaginary. It follows from Eq. (3.8) that the right-hand side of Eq. (3.10) is positive.] The residue of T_k may also be expressed in terms of $\dot{J}_{iK}$ by Eq. (2.41'):

$$\mathrm{Res}T_{iK} = \mathrm{tr}\mathring{J}_{iK}\dot{J}_{iK} = \mathrm{tr}F_{iK}^{-1}\dot{J}_{iK} \tag{3.11}$$

in the notation (2.47).

Now let us write the singular matrix J_{iK} in the form

$$J_{iK} = P\tilde{l}rP,$$

or

$$\mathring{J}_{iK} = F_{iK}^{-1} = -\tilde{r}l,$$

thereby defining the row vector l. [Here P is the matrix defined by Eq. (2.38).] Thus, l is such that

$$lJ_{iK} = 0. \tag{3.12}$$

According to Eq. (2.40) this means

$$W(lP\widetilde{\chi}_{iK}, \varphi_{iK}) = 0.$$

Since g_1 and g_2 are linearly independent, it follows that

$$lP\widetilde{\chi}_{iK} = l_1 f_2 - l_2 f_1 = 0,$$

or

$$c = l_2/l_1. \tag{3.13}$$

Therefore,

$$(1,c) = (1/l_1)l = \left(\frac{2c}{lQ\tilde{l}}\right)^{1/2} l$$

and Eq. (3.10) becomes

$$n^2 = \frac{2}{lQ\tilde{l}}[il\dot{J}_{iK}\tilde{r}]^2(-i\,\mathrm{Res}T_{iK}). \tag{3.14}$$

Note that this expression does not depend on the normalization of l. If J_k is known, then so are J_{iK} and $\dot{J}_{iK}$ and hence by Eq. (3.11) all the ingredients of the right-hand side of Eq. (3.14) are known. For later use we shall define the matrix

$$M = \tilde{r}r/n^2. \tag{3.15}$$

The bound states giving rise to poles of the diagonal elements of the S matrix in the upper half-plane, they are determined if S_k is known. Their number, in particular, is directly related to the phase of $\det S$. If we define δ_k to be continuous for $0 < k < \infty$ and such that

$$\det S_k = e^{2i\delta_k}, \tag{3.16}$$

then Eq. (2.30) tells us that

$$T = |T|e^{i\delta}. \tag{3.17}$$

Hence, the *Levinson theorem*

$$\delta(0) - \delta(\infty) = \pi n \tag{3.18}$$

follows in the usual manner[20] if T_k does not vanish at $k = 0$. Hence, n is the number of poles of T_k in the upper half-plane, which equals the number of bound states with negative eigenvalues. If there is a half-bound state at $k = 0$, then

$$\delta(0) - \delta(\infty) = \pi(n + \tfrac{1}{2}). \tag{3.18'}$$

Note that Eq. (2.41') shows that $-\delta$ is the phase of $\det J$:

$$\det J = |\det J|e^{-i\delta}. \tag{3.19}$$

Let us now "remove" a bound state from J_k. That is to say, let us define a new Jost matrix J_k' that has the same analyticity properties as J_k, that also satisfies Eqs. (2.42), (2.46), and (2.48), and that is not singular (in the matrix sense) at $k = iK$, where $\det J_{iK} = 0$. To this end we define

$$J_k' = J_k \Gamma_k, \tag{3.20}$$

with

$$\Gamma_k = \mathbb{1} - B + B\frac{k+iK}{k-iK}. \tag{3.21}$$

If B is chosen to be the singular matrix

$$B = \tilde{r}rI/rI\tilde{r} \tag{3.22}$$

with

$$I = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.23}$$

then B is a (nonorthogonal) projection

$$B^2 = B$$

and the constraints (2.46) and (2.48) are satisfied.[21] Thus, by Eq. (3.5), B is determined by the logarithmic derivative of the eigenfunction at the origin. Clearly, J'_k defined by Eq. (3.16) obeys Eq. (2.42) and it is not singular at $k = iK$.

Note that

$$\Gamma_k^{-1} = \Gamma_k^* = \Gamma_{-k} \tag{3.24}$$

and

$$\mathbb{1} - B = QBQ. \tag{3.25}$$

Therefore,

$$\Gamma_{-k} Q \Gamma_k^{-1} Q = \frac{k - iK}{k + iK} \mathbb{1},$$

and hence, if we define

$$S'_K = QJ'_{-k} Q J'^{-1}_k, \tag{3.26}$$

then Eq. (2.44) shows that

$$S_k = \frac{k + iK}{k - iK} S'_k. \tag{3.27}$$

Thus, S'_k is an S matrix that belongs to a system without a bound state of eigenvalue $-K^2$.

We remark that if

$$J'_k = IJ_k I, \tag{3.28}$$

then

$$S'_k = IS_k I, \tag{3.29}$$

in which the transmission amplitude is unchanged but both reflection amplitudes have changed sign. We may therefore alternatively remove a bound state by defining

$$J'_k = IJ_k \Gamma_k I \tag{3.20'}$$

instead of Eq. (3.20), and

$$S'_k = \frac{k + iK}{k - iK} IS_k I \tag{3.27'}$$

instead of Eq. (3.27). This is the method used by D&T.

In this fashion all the bound states may be removed and the reduced S matrix defined by

$$S_k^{\text{red}} = \prod_n \frac{k - iK_n}{k + iK_n} I^n S_k I^n \tag{3.30}$$

has the decomposition

$$S_k^{\text{red}} = QJ^{\text{red}}_{-k} Q (J_k^{\text{red}})^{-1}. \tag{3.31}$$

The Jost matrix J_k is then defined in terms of J_k^{red} by

$$J_k = I^n J_k^{\text{red}} \prod_n (\Gamma^{(n)}_{-k} I). \tag{3.32}$$

The matrices $\Gamma^{(n)}_{-k}$ depend on the order in which the bound states are removed, because each such removal changes the null spaces at the remaining singularities. D&T give explicitly[22] the change in the potential that is associated with the removal of a bound state by the method of Eq. (3.27').

4. THE SPECTRAL FUNCTION

It is well known[4] that under the assumption (2.4) on the potential the spectrum of Eq. (2.1) is complete. Thus, we have the Parseval relation for any $f \in L^2$:

$$\int_{-\infty}^{\infty} dx |f|^2 = (1/2\pi) \int_{-\infty}^{\infty} dk \left| \int_{-\infty}^{\infty} dx f(x) u_{1k}(x) \right|^2 + \sum_N \left| \int_{-\infty}^{\infty} dx f u_N \right|^2, \tag{4.1}$$

where $u_{1K}(x)$ is the first component of the solution of Eq. (2.12), and u_N are the normalized bound-state eigenfunctions. The same relation holds for the second component $u_{2k}(x)$. Let us take the sum of the two Parseval relations for u_1 and u_2 and divide by 2. The first term on the right-hand side of Eq. (4.1) may then be written

$$(1/2\pi) \int_0^{\infty} dk \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f(x) f^*(y) \psi_k(x) \psi_k^\dagger(y),$$

where ψ_k is the row-vector solution of Eq. (2.12). However, because by Eqs. (2.15), (2.37), and (2.39)

$$\varphi_k = \psi_k J_k,$$

we may write this as

$$(1/2\pi) \int_0^{\infty} dk \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f(x) f^*(y) \varphi_k(x) \times (J_k^\dagger J_k)^{-1} \varphi_k^\dagger(y).$$

The bound-state terms in Eq. (4.1) can be rewritten by means of Eqs. (3.9) and (3.15) as

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f(x) f^*(y) \varphi_{iK}(x) M_K \varphi_{iK}^\dagger(y).$$

As a result one may write Eq. (4.1), with $E = k^2$, as

$$\int_{-\infty}^{\infty} dx |f|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f(x) f^*(y) \varphi_E(x) \, d\rho(E) \varphi_E^\dagger(y), \tag{4.2}$$

where the spectral function is given by

$$d\rho/dE = \begin{cases} (1/4\pi k)(J_k^\dagger J_k)^{-1}, & \text{for } E > 0, \\ \sum_b M_b \delta(E - E_b), & \text{for } E < 0, \end{cases} \tag{4.3}$$

if E_b are the bound-state eigenvalues. Equation (2.46) implies that ρ is a symmetric matrix.

Formally, Eq. (4.2) may be written

$$\int \varphi_E \, d\rho_E \varphi_E^\dagger(y) = \delta(x - y). \tag{4.4}$$

5. THE REPRESENTATION OF φ_k

The first component of Eq. (2.36) may be written

$$g_1(x) e^{-ikx} = 1 + (1/2ik) \int_0^x dy (1 - e^{-2ik(x-y)}) \times V(y) g_1(y) e^{-iky}.$$

From this one easily shows that for $x > 0$ and $\text{Im}k \leqslant 0$, as well as for $x < 0$ and $\text{Im}k \geqslant 0$, as $|k| \to \infty$,

$$g_1(x) e^{-ik} - 1 = O(1/|k|), \tag{5.1}$$

while for $x > 0$ and $\text{Im}k > 0$, as well as for $x < 0$ and $\text{Im}k < 0$,

$$g_1(x) e^{ikx} = O(1/|k|). \tag{5.2}$$

Furthermore, for fixed x, g_1 is an entire analytic function of

k. Therefore, Eqs. (5.1) and (5.2) imply that, considered as a function of k for fixed x, $g_1 - e^{ikx}$ belongs to $L^2(-\infty,\infty)$ on the real axis, and it is an entire analytic function of exponential type x. It follows from the Paley–Wiener theorem[23] that g_1 has the Povsner–Levitan representation[24]

$$g_1(x) = e^{ikx} - \int_{-x}^{x} dy h_1(x,y) e^{iky}. \tag{5.3}$$

Similarly, we obtain

$$g_2(x) = e^{-ikx} - \int_{-x}^{x} dy h_2(x,y) e^{-iky}. \tag{5.3'}$$

Insertion of these representations in the Schrödinger equation (2.1) leads to the conclusion that h_1 and h_2 both have to satisfy the partial differential equation

$$\frac{\partial^2}{\partial x^2} h(x,y) - \frac{\partial^2}{\partial y^2} h(x,y) - V(x) h(x,y) = 0, \tag{5.4}$$

and the boundary conditions on characteristics

$$-2 \frac{d}{dx} h(x,x) = V(x), \tag{5.5}$$

$$\frac{d}{dx} h(x,-x) = 0. \tag{5.6}$$

Since these equations uniquely determine[25] h, it follows that h_1 and h_2 are equal:

$$h_1(x,y) = h_2(x,y) = h(x,y).$$

In addition, V being real, h is real. Consequently, the two-component row vector φ_k possesses the Povsner–Levitan representation

$$\varphi_k(x) = \psi_k^0(x) - \int_{-x}^{x} dy h(x,y) \psi_k^0(y), \tag{5.7}$$

in which it should be noted that h is a *scalar*. If we define

$$\bar{h}(x,y) = h(x,y)\theta(|x| - |y|)$$

and

$$\omega(x,y) = \delta(x-y) - \bar{h}(x,y),$$

then Eq. (5.7) may be written in the form

$$\varphi_k(x) = \int dy \omega(x,y) \psi_k^0(y). \tag{5.8}$$

6. THE HILBERT PROBLEM

The problem to be considered in this section is that of constructing the Jost matrix from a given S matrix. We assume that we are given a 2×2 matrix valued function of k in the s class [defined below Eq. (2.34)]. The problem is to find an analytic (matrix-valued) function J_k that is regular in the upper half-plane, has the asymptotic form (2.42), and such that Eq. (2.44), i.e.,

$$QS_kJ_kQ = J_{-k}, \tag{6.1}$$

holds. If the given S_k is associated with bound states, then we first remove them from S_k by the procedure given at the end of Sec. 3, i.e., if there are n bound states, their eigenvalues are determined by the positions of the poles of the diagonal elements T_k of S_k in the upper half-plane. We then define the reduced S matrix as in Eq. (3.30) and wish to construct a reduced Jost matrix as in Eq. (3.31). Once we have constructed J_k^{red}, the actual Jost matrix is defined by Eq. (3.32). For each bound state this introduces a free parameter (r_2/r_1), the significance of which is exhibited in Eq. (3.5) as the logarithmic derivative of the bound-state eigenfunction at $x = 0$.

Our problem, then, is to solve Eq. (6.1) with a function that is regular in the upper half-plane, nowhere singular there (in the matrix sense), and approaches unity as $|k| \to \infty$. We shall call such a function a "1 solution" of Eq. (6.1). One possible procedure for solving this Hilbert problem is analogous to that of Plemelj.[26] We shall proceed in another way. We note that since we have assumed that S_k is in the s class, for which the existence of an underlying potential obeying Eq. (2.4) is assured, the existence of a 1 solution of Eq. (6.1) is also assured.

Our second observation is that since the given function S_k is the reduced S matrix, it satisfies Levinson's theorem (3.18) with $n = 0$, which we shall refer to as Eq. (3.18)°. Let us set

$$s_k = \det S_k, \quad \xi_k = \det J_k.$$

Then Eq. (6.1) leads to the equation

$$s_k \xi_k = \xi_{-k}, \tag{6.1'}$$

where s_k is continuous on the real axis and, as $k \to \pm\infty$,

$$s_k = 1 + O(1/k),$$

while ξ_k is to be an analytic function, regular and without zeros in the upper half-plane, with

$$\lim_{|k|\to\infty} \xi_k = 1.$$

A necessary and sufficient condition for the existence of a solution of Eq. (6.1') with these properties is Eq. (3.18)°. In fact, in that case we have explicitly

$$\xi_k = \exp\left(-\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dk' \delta_{k'}}{k' - k} \right),$$

the integral being meant as Cauchy's principal value. Since we assumed Eq. (3.18)° to hold, such a solution exists. If there also existed a solution of Eq. (6.1') that vanished at $k = i\beta_m$, $m = 1,\ldots,n$, then

$$\xi'_k = \xi_k \prod_m \frac{k + i\beta_m}{k - i\beta_m},$$

would solve Eq. (6.1') with

$$s'_k = s_k \prod_m \left(\frac{k - i\beta_m}{k + i\beta_m} \right)^2,$$

and ξ'_k would have no zeros. However, the phase of s'_k would not satisfy Eq. (3.18)°. Hence, a solution of Eq. (6.1') with zeros in the upper half-plane cannot exist. Consequently, the determinant of a 1 solution of Eq. (6.1) cannot vanish in the upper half-plane. We may therefore ignore this requirement.

The third observation to be made is that the solution of Eq. (6.1) is *unique*. Suppose there were two solutions, say $J^{(1)}$ and $J^{(2)}$,

$$QS_kJ_k^{(1)}Q = J_{-k}^{(1)},$$

$$QS_kJ_k^{(2)}Q = J_{-k}^{(2)}.$$

Eliminating S_k, we get

$$J^{(2)}_{-k}{}^{-1} J^{(1)}_{-k} Q = Q J^{(2)}_k{}^{-1} J^{(1)}_k.$$

Since we have shown that no solution of Eq. (6.1) can be singular in the upper half-plane, the right-hand side is regular in the upper half-plane and the left-hand side in the lower. Consequently, both sides must be entire analytic functions. Since, furthermore, both $J^{(1)}_k$ and $J^{(2)}_k$ approach the unit matrix as $|k| \to \infty$, it follows by Liouville's theorem that

$$J^{(2)-1} J_{(1)} = \mathbb{1},$$

and the two are identical. The 1 solution of Eq. (6.1) being unique, it follows that the only "0 solution" of Eq. (6.1) is the trivial one $J_k \equiv 0$.

We are thus left with the problem of constructing the 1 solution of Eq. (6.1). The function J_k being analytic in the upper half-plane, continuous on the real axis, and obeying Eq. (2.42) asymptotically, it satisfies the "dispersion relation"

$$\begin{aligned} J_k - \mathbb{1} &= \frac{1}{2\pi i} \int_{-\infty}^{\infty} dk' \frac{J_{k'} - \mathbb{1}}{k' - k - i\epsilon} \\ &= \frac{i}{2\pi} \int_{-\infty}^{\infty} dk' \frac{J_{-k'} - \mathbb{1}}{k' + k + i\epsilon} \end{aligned} \tag{6.2}$$

and hence by Eq. (6.1)

$$\begin{aligned} J_k = \mathbb{1} &+ \frac{i}{2\pi} \int_{-\infty}^{\infty} dk' \frac{Q(S_{k'} - \mathbb{1}) J_{k'} Q}{k' + k + i\epsilon} \\ &+ \frac{i}{2\pi} Q \int_{-\infty}^{\infty} dk' \frac{J_{k'} - \mathbb{1}}{k' + k + i\epsilon} Q. \end{aligned} \tag{6.3}$$

However, because of the analyticity of J_k in the upper half-plane, the last integral vanishes. Therefore, J_k must satisfy the integral equation

$$J_k = \mathbb{1} + \frac{i}{2\pi} \int_{-\infty}^{\infty} dk' \frac{Q(S_{k'} - \mathbb{1}) J_{k'} Q}{k' + k + i\epsilon}. \tag{6.4}$$

Let us insert in Eq. (6.4) the Fourier integral

$$(k' + k + i\epsilon)^{-1} = -i \int_0^{\infty} d\alpha \, e^{i\alpha(k + k')},$$

in which the right-hand side is understood to be the limit as k approaches the real axis from above. We then obtain the representation

$$J_k = \mathbb{1} + \int_0^{\infty} d\alpha L(\alpha) e^{ik\alpha}, \tag{6.5}$$

where

$$L(\alpha) = (1/2\pi) \int_{-\infty}^{\infty} dk Q(S_k - \mathbb{1}) J_k Q e^{ik\alpha}. \tag{6.6}$$

[The inversion of the order of the two integrations is justified because (for $\epsilon > 0$) both integrals converge absolutely.] Inserting Eq. (6.5) in (6.6), one obtains a linear integral equation for $L(\alpha)$ as

$$L(\alpha) = K(\alpha) + \int_0^{\infty} d\beta K(\alpha + \beta) Q L(\beta) Q, \tag{6.7}$$

in which $K(\alpha)$ is given by

$$K(\alpha) = (1/2\pi) Q \int_{-\infty}^{\infty} dk (S_k - \mathbb{1}) e^{ik\alpha} Q. \tag{6.8}$$

Equation (6.7) has the drawback that it contains L on the right-hand side multiplied by matrices both on the right and the left. We eliminate this by replacing the 2×2 matrix L by a four-component column vector

$$L = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix} \to \begin{pmatrix} L_{11} \\ L_{21} \\ L_{12} \\ L_{22} \end{pmatrix} \equiv \mathbf{L}$$

and similarly for the inhomogeneous term

$$K \to \mathbf{K},$$

while replacing the 2×2 matrix KQ in the integral by the 4×4 matrix

$$KQ \to \begin{pmatrix} KQ & 0 \\ 0 & KQ \end{pmatrix} = \mathscr{K}.$$

Then

$$LQ \to \hat{Q} \mathbf{L},$$

where $\hat{Q}$ is the 4×4 matrix

$$\hat{Q} = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

In this language Eq. (6.7) reads

$$\mathbf{L}(\alpha) = \mathbf{K}(\alpha) + \int_0^{\infty} d\beta \mathscr{K}(\alpha + \beta) \hat{Q} \mathbf{L}(\beta). \tag{6.9}$$

To find J_k is now reduced to solving Eq. (6.9), whose structure is similar to that of the Marchenko equation. Since we assumed that S_k is in the s class, we know that a solution of Eq. (6.9) exists. All we have to worry about is its uniqueness.

Let us assume, then, that the homogeneous form of Eq. (6.9), which we shall call Eq. (6.9)°, has a nontrivial solution in L^2. Then Eqs. (6.5) and (6.6) would have a nontrivial solution without the $\mathbb{1}$ in Eq. (6.5); in other words, Eq. (6.4) without the $\mathbb{1}$ would have a nontrivial solution J^0_k. Such a solution would be analytic in the upper half-plane and tend to zero there as $|k| \to \infty$. It would follow that

$$\begin{aligned} J^0_k &= \frac{i}{2\pi} \int_{-\infty}^{\infty} dk' \frac{Q S_{k'} J^0_{k'} Q}{k' + k + i\epsilon} \\ &= \frac{i}{2\pi} \int_{-\infty}^{\infty} dk' \frac{J^0_{-k'}}{k' + k + i\epsilon}, \end{aligned}$$

and hence

$$Q S_k J^0_k Q - J^0_{-k} \equiv E_k \tag{6.10}$$

would be analytic in the upper half-plane and tend to zero there as $|k| \to \infty$. A little algebra and the use of Eqs. (2.26)–(2.28) show that then

$$(E_{-k} P) = Q S^*_k (E_k P) Q. \tag{6.11}$$

Since we known that the only 0 solution of Eq. (6.1) is the trivial one, E_k cannot be identically zero. Hence, $E_k P$ is a nontrivial 0 solution of a problem like Eq. (6.1), but with S_k replaced by S^*_k. We call this the *accompanying problem* (6.1)*.

Thus, if Eq. (6.9)° has nontrivial solutions, then the solution of Eq. (6.9) is not unique, but only one of these solutions is the 1 solution of Eq. (6.1); the others lead, via Eq.

(6.10), to a 0 solution of Eq. (6.1)*. [A solution of Eq. (6.9) will lead to the analog of Eq. (6.10) and hence Eq. (6.11), with the same properties of E_k.]

7. A GEL'FAND–LEVITAN EQUATION

An analog of the Gel'fand–Levitan equation can now be derived from the results of Secs. 4 and 5. Let us write Eq. (4.4) in the abbreviated operator notation

$$\int \varphi_E \, d\rho_E \varphi_E^\dagger = \delta, \tag{7.1}$$

and similarly

$$\int \psi_E^0 \, d\rho_E^0 \psi_E^{0\dagger} = \delta, \tag{7.2}$$

where $d\rho/dE$ is given by Eq. (4.3), and

$$d\rho^0/dE = \begin{cases} 1/4\pi k, & \text{for } E>0, \\ 0, & \text{for } E<0. \end{cases}$$

We also write Eq. (5.8) as

$$\varphi_E = \omega \psi_E^0. \tag{7.3}$$

Then Eq. (7.1) gives

$$\omega^{\dagger -1} = \omega \int \psi_E^0 \, d\rho_E \psi_E^{0\dagger},$$

and Eq. (7.2) gives

$$\omega = \omega \int \psi_E^0 \, d\rho_E^0 \psi_E^{0\dagger}.$$

Substracting these two equations, we get

$$\omega^{\dagger -1} - \omega = \omega \int \psi_E^0 d(\rho_E - \rho_E^0) \psi_E^{0\dagger}. \tag{7.4}$$

Now consider the kernel of ω^{-1}:

$$(\omega^{-1})(x,y) = \delta(x-y) + l(x,y),$$

where $l(x,y)$ is subject to the Volterra equation

$$l(x,y) = \bar{h}(x,y) + \int_{-x}^{x} dz h(x,z) l(z,y),$$

whose inhomogeneity vanishes for $|x| < |y|$. Consequently, $l(x,y) = 0$ for $|x| < |y|$. This means that for $|x| > |y|$

$$(\omega^{\dagger -1} - \omega)(x,y) = h(x,y).$$

Therefore, Eq. (7.4) becomes, for $|x| \geqslant |y|$,[27]

$$h(x,y) = h_0(x,y) - \int_{-x}^{x} dz h(x,z) h_0(z,y), \tag{7.5}$$

where

$$h_0(x,y) = \int \psi_E^0(x) d(\rho_E - \rho_E^0) \psi_E^{0\dagger}(y), \tag{7.6}$$

$$d(\rho_E - \rho_E^0)/dE = \begin{cases} (1/4\pi k)[(J_k^\dagger J_k)^{-1} - 1], & \text{for } E>0, \\ \sum_b M_b \delta(E - E_b), & \text{for } E<0. \end{cases} \tag{7.7}$$

Note that h and h_0 are both *scalars*.

Let us write

$$m_k = (J_k^\dagger J_k)^{-1} - 1.$$

Then it follows from Eq. (2.45) that

$$m_{-k} = m_k^* = \tilde{m}_k, \tag{7.8}$$

and from Eq. (2.46) that

$$m_k^{(11)} = m_k^{(22)}. \tag{7.9}$$

As a result we may write

$$h_0(x,y) = \sum_b \xi_b(x) \xi_b(y) + (1/2\pi) \int_{-\infty}^{\infty} dk \, (m_k^{(11)} e^{ikx} + m_k^{(12)} e^{-ikx}) e^{iky}, \tag{7.6'}$$

where

$$\xi_b(x) = (r_1 e^{-Kx} + r_2 e^{Kx})/n$$

and h_0 is manifestly real.

Equation (7.5) is an integral equation of the Fredholm type, with a trace-class kernel. That its homogeneous form has no nontrivial solution can be shown in the standard fashion.[28] It therefore has a unique solution $h(x,y)$ in L^1. That solution is connected with the potential in Eq. (2.1) by Eq. (5.5), and with the solution of Eq. (2.1) by Eq. (5.7).

If we define

$$\eta(x,y) = \frac{\partial^2}{\partial x^2} h(x,y) - \frac{\partial^2}{\partial y^2} h(x,y) + 2 \left[\frac{\partial}{\partial x} h(x,x) \right] h(x,y), \tag{7.10}$$

then it is a matter of straightforward differentiation and integration by parts to show from Eq. (7.5) that, for $|x| \geqslant |y|$, η satisfies the equation

$$\eta(x,y) = -2 \left[\frac{\partial}{\partial x} h(x,-x) \right] h_0(-x,y) - \int_{-x}^{x} dz \eta(x,z) h_0(z,y). \tag{7.11}$$

Therefore, if the solution $h(x,y)$ of Eq. (7.5) satisfies (5.6), then η satisfies the homogeneous form of Eq. (7.5) and hence vanishes. It then follows from Eq. (7.10) that h satisfies Eq. (5.4) with V defined as in Eq. (5.5). If we then define $\varphi_k(x)$ by Eq. (5.7), it satisfies the Schrödinger equation (2.1) with this potential. Inserting Eq. (7.6) in (7.5) and using Eq. (5.7), we obtain the expansion for $|x| \geqslant |y|$ as

$$h(x,y) = \int \varphi_E(x) \, d(\rho_E - \rho_E^0) \psi_E^{0\dagger}(y). \tag{7.12}$$

Furthermore, the functions $\varphi_k(x)$ can then be shown to form a complete set as in Eq. (4.4) with the spectral function given by Eq. (4.3) in terms of Eq. (7.7). Thus, we are led from the starting point (7.6) to the solution of the Schrödinger equation with the spectral function that entered into Eq. (7.6).

8. A MARCHENKO EQUATION

Consider the row-matrix function

$$\gamma_k(x) = \chi_k(x) e^{-iIkx}, \tag{8.1}$$

χ being defined by Eq. (2.10) and the 2×2 matrix I by Eq. (3.23). The results of Sec. 2 imply that γ_k is an analytic function of k regular in the upper half-plane, continuous on the real axis, with the asymptotic behavior as $|k|\to\infty$ for $\mathrm{Im}k\geqslant 0$:

$$\gamma_k = \hat{1} + O(1/|k|). \tag{8.2}$$

Here, $\hat{1}$ is the row vector

$$\hat{1} = (1,1).$$

It follows that γ_k satisfies the dispersion relation

$$\gamma_k - \hat{1} = \frac{1}{2\pi i}\int_{-\infty}^{\infty} dk' \frac{\gamma_{k'} - \hat{1}}{k' - k - i\epsilon} = \frac{i}{2\pi}\int_{-\infty}^{\infty} dk' \frac{\gamma_{-k'} - \hat{1}}{k' + k + i\epsilon}. \tag{8.3}$$

Now according to Eqs. (2.15), (2.19), and (2.30), in the notation defined by Eqs. (2.47) and (2.17),

$$\chi_{-k} = \chi_k \mathring{S}_k Q,$$

and therefore

$$\gamma_{-k} = \gamma_k \mathscr{S}_k Q, \tag{8.4}$$

where

$$\mathscr{S}_k(x) = e^{ilkx}\mathring{S}_k e^{-ilkx}.$$

We are looking for a solution γ_k of Eq. (8.4) that is analytic in the upper half-plane and obeys Eq. (8.2) there. Such a solution will be called a 1 solution. This resembles a Hilbert problem.

It follows from Eq. (2.31) that $\mathscr{S}_k$ is continuous and, as $k\to\pm\infty$,

$$\mathscr{S}_k = \mathbb{1} + O(1/k). \tag{8.5}$$

Inserting Eq. (8.4) in (8.3), we obtain

$$\gamma_k - \hat{1} = \frac{i}{2\pi}\int_{-\infty}^{\infty} dk' \frac{\gamma_{k'}(\mathscr{S}_{k'} - \mathbb{1})Q}{k' + k + i\epsilon} + \frac{i}{2\pi}\int_{-\infty}^{\infty} dk' \frac{(\gamma_{k'} - \hat{1})Q}{k' + k + i\epsilon}. \tag{8.6}$$

Because of the analyticity of γ_k in the upper half-plane and Eq. (8.2), the last integral vanishes. Consequently γ_k obeys the integral equation

$$\gamma_k = \hat{1} + \frac{i}{2\pi}\int_{-\infty}^{\infty} dk' \frac{\gamma_{k'}(\mathscr{S}_{k'} - \mathbb{1})Q}{k' + k + i\epsilon}. \tag{8.7}$$

Use of the Fourier integral given below Eq. (6.4) leads to the representation

$$\gamma_k = \hat{1} + \int_0^{\infty} d\alpha \eta(\alpha) e^{ik\alpha}, \tag{8.8}$$

in which

$$\eta(\alpha) = (1/2\pi)\int_{-\infty}^{\infty} dk \gamma_k(\mathscr{S}_k - \mathbb{1})Q\, e^{ik\alpha}. \tag{8.9}$$

(Inversion of the orders of the k and α integrations is justified by the absolute convergence of the integrals before $\epsilon\to 0$.)

Note that the Schrödinger equation (2.1), written in terms of $\gamma_k(x)$, reads

$$\Pi_k = V\gamma_k, \tag{8.10}$$

if Π_k is defined as

$$\Pi_k = \frac{\partial^2}{\partial x^2}\gamma_k + 2ik\frac{\partial}{\partial x}\gamma_k I. \tag{8.11}$$

Insertion of the representation (8.8) in this leads to the partial differential equation

$$\frac{\partial^2}{\partial x^2}\eta - 2\frac{\partial^2}{\partial x\,\partial\alpha}\eta I = V(x)\eta \tag{8.12}$$

and the boundary condition on a characteristic

$$-2\frac{\partial}{\partial x}\eta(\alpha = 0)I = V\hat{1}. \tag{8.13}$$

If we write

$$\eta_1(x,\alpha) = \mu_1(x, x+\alpha),$$

$$\eta_2(x,\alpha) = \mu_2(x, x-\alpha),$$

then Eq. (8.12) reads

$$\frac{\partial^2}{\partial x^2}\mu_i(x,y) - \frac{\partial^2}{\partial y^2}\mu_i(x,y) = V(x)\mu_i(x,y), \tag{8.12'}$$

for $i = 1,2$, while Eq. (8.13) becomes

$$-2\frac{\partial}{\partial x}\mu_1(x,x) = 2\frac{\partial}{\partial x}\mu_2(x,x) = V(x). \tag{8.13'}$$

In addition,

$$\lim_{y\to+\infty}\mu_1(x,y) = \lim_{y\to-\infty}\mu_2(x,y) = 0. \tag{8.14}$$

In this language Eq. (8.8) becomes

$$f_1(x) = e^{ikx} + \int_x^{\infty} dy\, \mu_1(x,y)e^{iky},$$

$$f_2(x) = e^{-ikx} + \int_{-\infty}^{x} dy\, \mu_2(x,y)e^{-iky}, \tag{8.8'}$$

which are the representations given by F.

We now insert Eq. (8.8) in (8.9) and obtain the integral equation

$$\eta(\alpha) = \hat{1}G(\alpha) + \int_0^{\infty} d\beta\eta(\beta)G(\alpha+\beta), \tag{8.15}$$

where

$$G(\alpha) = (1/2\pi)\int_{-\infty}^{\infty} dk\,(\mathscr{S}_k - \mathbb{1})e^{ik\alpha}Q. \tag{8.16}$$

Note that since $\mathscr{S}_k$ depends on x, so do G and η.

We are again in a position of knowing *a priori* that a solution to Eq. (8.15) exists if S_k is in the s class. If this solution is unique, then it must, via Eq. (8.8), lead to the 1 solution of Eq. (8.4), and to the potential via Eq. (8.13). Note that from the point of view of solving Eq. (8.15), the fact that Eq. (8.13) comes out as a scalar multiple of the row vector $\hat{1}$ may seem like a miracle. However, if the homogeneous equation (8.15)° has no nontrivial solution and S_k is in the s class, this miracle is certain to happen.

Suppose that $\eta(\alpha)$ is a solution of Eq. (8.15). Then it can be written in the form (8.9) with γ_k given by Eq. (8.8). Insertion of Eq. (8.9) in (8.8) then shows that γ_k satisfies Eq. (8.7). By straightforward differentiation and a bit of algebra one then finds that the function Π_k defined in Eq. (8.11) satisfies Eq. (8.7) with the $\hat{1}$ replaced by

$$v(x) = \frac{1}{\pi}\frac{\partial}{\partial x}\int_{-\infty}^{\infty} dk\, \gamma_k(\mathscr{S}_k - 1)P, \tag{8.17}$$

P being defined by Eq. (2.38). If this row vector is of the form

$$v = V\hat{1}, \tag{8.18}$$

where $V(x)$ is a scalar, i.e., if $v_1 = v_2$, and if Eq. (8.15) has a unique solution, then Eq. (8.10) follows. However, we cannot draw this conclusion if the solution of Eq. (8.15) is not unique. Thus, in such a case we are not assured that a given solution leads to a solution of the Schrödinger equation.

Equation (8.7) implies that γ_k is regular in the upper half-plane and approaches $\hat{1}$ as $|k| \to \infty$. It therefore satisfies both Eqs. (8.6) and (8.3), subtraction of which leads to

$$\int_{-\infty}^{\infty} dk' \frac{\beta_{k'}}{k' + k + i\epsilon} = 0,$$

where

$$\beta_k = \gamma_k \mathscr{S}_k Q - \gamma_{-k}. \tag{8.19}$$

It follows that β_k is regular in the upper half-plane and tends to zero there as $|k| \to \infty$. Now one readily shows that β_k satisfies the equation

$$-\beta_k \mathscr{S}_{-k} Q = \beta_{-k}, \tag{8.4)*}$$

which we shall call the *accompanying* problem to Eq. (8.4). It is obtained from Eq. (8.4) by the replacement $\mathscr{S}_k \to -\mathscr{S}_k^*$.

We thus have found that if γ_k solves Eq. (8.7) [or it is obtained by Eq. (8.8) from a solution of Eq. (8.15)], then *either* it is a 1 solution of Eq. (8.4), which by Eq. (8.19) would mean $\beta_k = 0$, *or* the function β_k defined by Eq. (8.19) is a 0 solution of Eq. (8.4)*. If Eq. (8.15)° has no nontrivial solution, then the former must be the case, and furthermore the potential is given by Eqs. (8.17) and (8.18).

On the other hand, we know that if the original potential produces n bound states, then there exists an n-parameter family of potentials all of which are associated with the same S matrix (and hence also the same bound-state eigenvalues, since the latter are determined by T). Consequently, in the presence of n bound states the homogeneous equations (8.15)° and (8.7)° must have at least n linearly independent solutions. However, in that case the conclusion (8.10) cannot be drawn, and there is no known systematic way of choosing the correct solutions γ_k of Eq. (8.7) that guarantee both the miracle (8.18) and (8.10). Therefore, the method of this section is applicable only to an S matrix that has no associated bound states. This S matrix may be either that of a system known to be without bound states, or it may be a reduced S matrix the bound states of which have been removed by Eq. (3.30). Such a reduced S matrix is then known to have an underlying potential in class (2.4) if the original one was. After reconstructing the potential that goes with S_k^{red} one may then add back the bound states by the method[29] of D&T.

9. THE INVERSION PROCEDURES

We are now ready to give the inversion procedures based on the equations of Secs. 6 and 7 or on those of Sec. 8. We shall also compare them with those previously known.

We assume that a 2×2 S matrix in the s class is given. This means that both reflection amplitudes and the transmission amplitude are given for all k, $0 < k < \infty$ [and hence by Eq. (2.26) for all $-\infty < k < \infty$]. The eigenvalues of the bound states, if any, are therefore also determined. This is because T_k is the boundary value of an analytic function that is meromorphic in the upper half-plane and satisfies Eq. (2.32). It has simple poles at $k = iK_n$ if $-K_n^2$ are the eigenvalues of Eq. (2.1). Hence, for $x > 0$,

$$\frac{1}{2\pi i}\int_{-\infty}^{\infty} dk\,(T_k - 1)\, e^{ikx} = \sum_n e^{-K_n x}\,\text{Res}T_{iK_n}, \tag{9.1}$$

from which we may obtain both K_n and $\text{Res}T_{iK_n}$.

We now remove the bound states and define the reduced S matrix by Eq. (3.30). In the Gel'fand–Levitan method one then proceeds as follows:

The Hilbert problem for S_k^{red} is solved to find J_k^{red} by means of the Fredholm equations (6.9) and (6.5). Since an underlying potential exists, this will work if Eq. (6.9)° has no nontrivial solution. [If Eq. (6.9)° has nontrivial solutions, then one and only one of the functions (6.5) solves Eq. (6.1).] One then has three options.

Method 1: Having obtained J_k^{red}, one constructs the spectral functions (4.3) for the reduced problem (i.e., with the bound states removed). Thus, the kernel h_0 of Eq. (7.6) is known. One then solves the Gel'fand–Levitan equation (7.5). Since an underlying potential exists, the unique solution of Eq. (7.5) must give it via Eq. (5.5). One then adds the bound states back by the method of D&T. This yields an n-parameter family of potentials, the parameters being the norming constants of the bound states.

Method 2: Having obtained J_k^{red}, one forms J_k by Eq. (3.32). This yields an n-parameter family of Jost matrices, the parameters being the logarithmic derivatives of the eigenfunctions at the origin [the r in Eq. (3.22)]. Though those two parameters, the norming constants and the logarithmic derivatives, are different, they are determined by one another. [Once J_{iK} is known, so is l, and hence Eqs. (3.13) and (3.8) determine N, $\text{Res}T_{iK}$ being known.] Next Eq. (7.5) is solved for each of these J_k and the potential is determined from Eq. (5.5).

Method 3: In this procedure a Jost matrix is not needed. Instead, we first remove the bound states as before, via Eq. (3.30). We then use the reduced S matrix in Eq. (8.16) and solve the Marchenko equation (8.15). If Eq. (8.15)° has nontrivial solutions, then solutions of Eq. (8.15) must still exist and one of them must, via Eq. (8.8), lead to $\beta_k = 0$ in Eq. (8.19). This solution then yields the potential by Eq. (8.13). We then add the bound states by the method of D&T.

The virtue of method 3 over the first two is that it requires the solution of only *one* (matrix) integral equation, though for each x the integral in Eq. (8.15) extends over an infinite interval. Methods 1 and 2, on the other hand, require the solution of *two* (matrix) integral equations (6.9) and (7.5). However, Eq. (6.9) is independent of x and the integration in Eq. (7.5) extends over a finite range only.

Let us compare the procedures of this paper with those of F and of D&T. While their methods differ from one another, they have in common that the assumed data consist of one reflection amplitude as a function of k, $0 < k < \infty$ [and

hence by Eq. (2.26) for $-\infty < k < \infty$], and, independently, the bound-state eigenvalues, plus one norming constant for each bound state. In the methods of this paper the assumed data are the S matrix for $0 < k < \infty$, and a set of as many real constants as there are bound states. (These may be either norming constants or logarithmic derivatives.) However, the restrictions on S_k implied by being in the s class are not enforcable as easily as are the needed restrictions on R in the methods of F and D&T. The way to do it here is to assign a function T_k that is meromorphic in the upper half-plane and whose modulus remains less than 1 everywhere on the real axis. In addition, the phase of either R_k^l or R_l^r may be independently assigned. Its modulus is then determined by Eq. (2.28'), and the other reflection amplitude is given by Eq. (2.28''). One would then have to check whether the so constructed reflection amplitudes obey Eq. (2.33). Whereas this procedure is "natural" for the present methods, one may alternatively start with the data assumed for F and D&T and then construct the transmission amplitude and the other reflection amplitude. The thus constructed S matrix may then be used as the starting point. Similarly, one may start with the S matrix, then construct the eigenvalues by Eq. (9.1), and then proceed with the method of F or of D&T.

ACKNOWLEDGMENTS

Most of the work reported in this paper was done while the author was on sabbatical leave at the Institute for Advanced Study, Princeton, NJ, and at the Physics Department of Princeton University. The hospitality of both institutions is gratefully acknowledged. Stimulating discussions with Martin Kruskal and Eugene Trubowitz are also acknowledged with pleasure.

This work was supported in part by the National Science Foundation under Grant No. PHY77–25337.

APPENDIX A. PROOF OF EQ. (2.49)

The Fredholm determinant Δ is given by

$$\Delta = \exp \operatorname{tr} \ln(1 - G_0^+ V), \tag{A1}$$

in which the kernel of the Green's function is given by

$$G_0^+(x,y) = -(i/2k)e^{ik|x-y|}.$$

Because of Eq. (2.8) the complete Green's function is given by

$$G^+(x,y) = \begin{cases} -(i/2k)Tf_1(x)f_2(y), & \text{if } x > y, \\ -(i/2k)Tf_1(y)f_2(x), & \text{if } x < y. \end{cases} \tag{A2}$$

Let us now replace V by γV and differentiate Eq. (A1) with respect to γ:

$$(d\Delta/d\gamma)/\Delta = -\operatorname{tr}(1 - \gamma G_0^+ V)^{-1} G_0^+ V = -\operatorname{tr} G^+ V$$
$$= -(i/2k)T\int_{-\infty}^{\infty} dx V f_1 f_2,$$

by Eq. (A2). On the other hand, differentiate Eq. (2.6):

$$d(1/T)/d\gamma = -(dT/d\gamma)/T^2$$
$$= (i/2k)\int_{-\infty}^{\infty} dx\, e^{-ikx}(f_1 + \gamma df_1/d\gamma).$$

If we write G_0 for the operator whose kernel is

$$G_0(x,y) = -(1/k)\sin k(x-y)\theta(x-y).$$

where $\theta(x)$ is the Heaviside function, then Eq. (2.3) reads

$$f_1 = e^{ikx} + \gamma G_0 V f_1,$$

and hence

$$df_1/d\gamma = G_0 V f_1 + \gamma G_0 V df_1/d\gamma.$$

Consequently,

$$\gamma df_1/d\gamma = (1 - \gamma G_0 V)^{-1}\gamma G_0 V f_1 = (1 - \gamma G_0 V)^{-1} f_1 - f_1,$$

and by using Eq. (2.3'),

$$(dT/d\gamma)/T = -(i/2k)T\int_{-\infty}^{\infty} dx\, e^{-ikx} V(1 - \gamma G_0 V)^{-1} f_1$$
$$= -(i/2k)T\int_{-\infty}^{\infty} dx\, e^{-ikx}(1 - \gamma V G_0)^{-1} V f_1$$
$$= -(i/2k)T\int_{-\infty}^{\infty} dx V f_1 f_2.$$

We have therefore found that

$$d\ln(T\Delta)/d\gamma = 0.$$

Since for $\gamma = 0$, $T = \Delta = 1$, it follows that for all γ, $\Delta = 1/T$, and Eq. (2.49) follows from Eq. (2.41').

APPENDIX B. TRANSLATIONS

Suppose that the potential is shifted by y as

$$V^y(x) = V(x + y).$$

Then it follows from Eq. (2.12) that the corresponding solution ψ_k^y is given by

$$\psi_k^y(x) = \psi_k(x + y)e^{-iIky}$$

where I is the matrix defined in Eq. (3.23). Consequently, by Eq. (2.19),

$$S_k^y = e^{iIky} S_k e^{-iIky}.$$

This means that

$$R^{ly} = e^{-2iky} R^l,$$
$$R^{ry} = e^{2iky} R^r.$$

Similarly, from Eqs. (2.3), (2.3'), and (2.36),

$$\chi_k^y(x) = \chi_k(x + y)e^{-iIky},$$
$$\varphi_k^y(x) = \varphi_k(x + y)e^{-iIky},$$

and hence from Eq. (2.40)

$$J_k^y = e^{iIky} J_k e^{-iIky}.$$

Notice that if we choose $y = \pi/2k$, then

$$S^y = \begin{pmatrix} T & -R^r \\ -R^l & T \end{pmatrix},$$

so that a change of sign of both reflection amplitudes is obtained by the energy dependent potential

$$V_k(x) = V(x + \tfrac{1}{2}\pi k^{-1}).$$

We also note that for the shifted potential the matrix F_k of Eq. (2.37) is given by

$$F_k^x = e^{iIkx} F_x e^{-iIkx}$$

and we find from Eq. (2.39) that

$$(F_{-k}^x)^{-1} = Q(F_k^x)^{-1}\mathscr{S}_k(x)Q,$$

where $\mathscr{S}_k(x)$ is the matrix defined below Eq. (8.4). Therefore, the row vector

$$\hat{1}(F_k^x)^{-1}$$

satisfies the same equation (8.4) as γ_k and has the same analyticity properties. However, it does not satisfy the asymptotic requirement (8.2).

[1]See K. Chadan and P.C. Sabatier, *Inverse Problems in Quantum Scattering Theory* (Springer, New York, 1977).
[2]I. Kay and H.E. Moses, Nuovo Cimento **3**, 276 (1956).
[3]L.D. Faddeev, Dokl. Akad. Nauk SSSR **121**, 63 (1958) [Math. Rev. **20**, 773 (1959)].
[4]L.D. Faddeev, Trudy Mat. Inst. Steklov **73**, 314 (1964) [Am. Math. Soc. Transl. **2**, 139 (1964).] We shall refer to this paper as F.
[5]L.D. Faddeev, Itogi Nauk. Tekh. Sov. Probl. Mat. **3**, 93 (1974) [J. Sov. Math. **5**, 334 (1976)].
[6]I. Kay, Commun. Pure Appl. Math. **13**, 371 (1960).
[7]P. Deift and E. Trubowitz, Commun. Pure Appl. Math. **32**, 121 (1979). We shall refer to this paper as D&T.
[8]I. Kay and H.E. Moses, Nuovo Cimento **22**, 683 (1961); and Commun. Pure Appl. Math. **14**, 435 (1961).
[9]R.G. Newton, in *Scattering Theory in Mathematical Physics*, edited by J.A. Lavita and J.-P. Marchand (Reidel, Dordrecht, 1974), pp. 193–235. A slightly different version was given in lectures at the *1974 Summer Seminar on Inverse Problems* American Mathematical Society, U.C.L.A. (unpublished).
[10]R.G. Newton, Phys. Rev. Lett. **43**, 541 (1979).
[11]H.E. Moses, Studies Appl. Math. **58**, 187 (1978).
[12]The product decompositions given by H.E. Moses, J. Math. Phys. **5**, 833 (1964), and on p. 353 of Ref. 5, may look similar but they are not "canonical" in that the factors lack analyticity properties.
[13]F uses the first absolute moment in place of the second. It was pointed out both in Ref. 1 and by D&T that the second moment is needed. The place in the analysis where the second moment is required is at $k = 0$ when there is a "half-bound" state.
[14]We use * for complex conjugation, ˜ for matrix transition, and † for Hermitian conjugation: $^\dagger = \tilde{}^*$.
[15]See p. 156 of Ref. 4 and p. 203 of Ref. 6.
[16]Reference 6, p. 210. This is the generic case.
[17]Reference 6, p. 212.
[18]See Sec. 6.
[19]The argument to the contrary in the three-dimensional case, given on p. 229 of Ref. 8, is in error. It requires the use of a differential equation with respect to rotation, as given by Faddeev, Ref. 5, p. 391, to obtain $c(\hat{k})$. There is no analog of this in one dimension.
[20]See, for example, R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw Hill, New York, 1966), pp. 355 and 356.
[21]In fact, it is readily shown that Eqs. (2.46) and (2.48) require that B be of the form (3.22).
[22]Reference 6, p. 172.
[23]See, for example, Ref. 1, p. 36.
[24]This representation was first given for a similar function by A. Sh. Bloch, Dokl. Akad. Nauk. SSSR **92**, 209 (1953).
[25]One may use the method of Z. S. Agranovich and V.A. Marchenko, *The Inverse Problem of Scattering Theory* (Gordon and Breach, New York, 1963), pp. 20ff, to prove this.
[26]J. Plemelj, Monatschr. Math. Phys. **19**, 211 (1908); see also N.I. Muskelishvili, *Singular Integral Equations* (Groningen, Holland, 1953), pp. 381ff.
[27]This equation was given by A. Sh. Bloch, Ref. 24, but without the connection of the Jost matrix to the S matrix.
[28]See, for example, Ref. 19, pp. 616 and 617.
[29]Reference 6, p. 167.

Combinatorics, partitions, and many-body physics[a)]

Wayne N. Polyzou
Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742

(Received 13 April 1979; accepted for publication 15 October 1979)

Some combinatorial techniques are presented which streamline the graphical analysis used in N-body scattering theory. The basic results are derived using properties of the lattice of partitions of N particles, which naturally arises on classifying translational symmetry properties of N-body operators. Classical cumulant expansions are recovered, previously obtained results are presented from a unified point of view, and some new theorems concerning connectivity of N-body equations are presented.

1. INTRODUCTION

In 1964, Van Winter[1] and Weinberg[2] independently developed the first N-body generalized of Faddeev's work[3] on three-particle scattering. The graphical analysis used in deriving their equation had its roots in combinatorial methods that have been widely used in many fields of physics and mathematics. In particular, we refer to the combinatorial machinery used in the construction of correlative expansions in statistical mechanics, cumulant expansions in probability theory, and truncated vacuum expectation values in quantum field theory.[4-10] As N-body physics developed an identity of its own, many of these ideas were forgotten. In this paper, we show how these ideas arise naturally in the context of the quantum mechanical N-body problem, and lead to combinatorial theorems that are indispensable tools for formulating connected integral equations. We show how translation symmetries of an N-body system can be classified by the lattice of partitions of N particles, and how this lattice structure leads to both well-known and new combinatorial results. The most striking results are Theorems 7 and 12, which are useful tools in formulating connected kernel integral equations. The utility of some of the theorems is illustrated in some familiar examples in Sec. 5. Although our considerations are restricted to the quantum mechanical N-body problem, the conbinatorial results should be useful in other applications.

Specific cases of the most important results of this paper first appeared in an unpublished paper of L'Huillier, Redish, and Tandy[11] (LRT), and later in papers by Benoist-Gueutal, L'Huillier, Redish, and Tandy,[12] and Kowalski.[13] The underlying combinatorial results in these works were extracted and extended by Redish and the author.[14] Subsequent applications[12-17] followed the work of LRT. The relation of the work of L'Huillier, Redish, and Tandy to cumulant expansions and related ideas was pointed out by Kowalski[16] and Simon.[18] The purpose of this paper is to present a unified treatment of these combinatorial results from a lattice theoretic point of view. This allows us to simplify the proofs of previously obtained results, generalized these results and obtain new results.

This paper is divided into five sections. In the Sec. 2 the lattice structure on the set of partitions is introduced as a way of classifying the subgroups of the N-body translational symmetry group. In Sec. 3 these symmetry properties are used to define the notion of connectivity and to classify symmetry properties of a certain class of linear operators. The main combinatorial results are presented in Sec. 4. Simple applications of some of the results are applied to examples from N-body quantum mechanics in Sec. 5. In the final section we discuss other uses of these results.

2. PARTITIONS AND TRANSLATIONAL INVARIANCE

In many physical systems one deals with systems of many particles involving few-particle interactions. For example, in the atomic nucleus, the force binding the nucleons has a short range and a repulsive core.[7] This indicates that the atomic nucleus can be adequately described with forces that never involve more than a few nucleons at a time. This basic property is characteristic of most N-body systems. Whenever one has such a system, it will possess a high degree of translational symmetry. In the case of the atomic nucleus the two-body interaction between particles 1 and 2 will be independent of the positions of particles 3...N. This means that this interaction operator will commute with the momentum operators associated with particles 2...N and the center of mass of the 1–2 system. For each such physical operator, there will be a many-parameter unitary group under which that operator is invariant. For the two-body interaction of our example the appropriate translation group is

$$\cup(a) = e^{-i\mathbf{a}\cdot\mathbf{P}}, \tag{1}$$

where $\mathbf{P} = (\mathbf{P}_{12}, \mathbf{P}_3, \mathbf{P}_4 \ldots \mathbf{P}_N)$, $\mathbf{P}_i$ is the momentum operator for particle i, $\mathbf{P}_{12} = \mathbf{P}_1 + \mathbf{P}_2$, and $\mathbf{a}$ is a $3(N-1)$ parameter vector. In general, an N-body theory will involve many kinds of operators with different translational symmetries. In order to treat this situation effectively, it is useful to have an efficient bookkeeping system to isolate classes of operators with different symmetries. We define the notion of a partition:

Definition 1: A *partition* a of N particles is a way of grouping the N particles into n_a disjoint groups called *clusters*. For example, a three cluster partition of seven particles is given by

$$(12)\,(3)\,(4567). \tag{2}$$

The ordering of the labeling of the clusters, or of the particles within a given cluster is immaterial. Lower case Latin letters

[a)]This work is taken from the author's Ph.D. thesis.

 0022-2488/80/030506-08$1.00

$a, b, c\cdots$ are used to denote partitions of $n_a, n_b, n_c\cdots$ clusters. The symbols $\underline{1}$ and $\underline{0}$ denote the unique 1-cluster and N-cluster partitions, respectively. It is useful to define the following sets of partitions:

Definition 2:

$$\begin{aligned}\mathscr{P} &= \{\text{All partitions of } N \text{ particles}\},\\ \mathscr{P}' &= \{a\in\mathscr{P}\,|n_a>1\} = \mathscr{P}/\{\underline{1}\},\\ \mathscr{P}'' &= \{a\in P\,|N>n_a>1\} = \mathscr{P}/\{\underline{0},\underline{1}\}.\end{aligned} \tag{3}$$

The sets $\mathscr{P}'$ and $\mathscr{P}''$ arise naturally in applications. $\mathscr{P}'$ is the set of partitions that are associated with the proper subsystem problems of our system. The one-cluster partition, $\underline{1}$, which groups all N-particles together is excluded. $\mathscr{P}''$ is the subset of $\mathscr{P}'$ which excludes the N-cluster partition $\underline{0}$. This is naturally associated with N single-particle systems which is the only trivial subsystem.

To each partition, a, we associate the $3n_a$ parameter unitary group of translations that independently transform the centers of mass of each of the clusters of a. We denote this group by $\mathscr{G}_a$.

The group, $\mathscr{G}_a$, associated with the partition a generally admits many subgroups corresponding to translations that do not change the relative positions of various sets of clusters of a. These subgroups correspond to the subgroups $\mathscr{G}_b$, where the clusters of b are composed of the particles in the clusters of a whose relative positions are invariant under the action of this subgroup. This leads to the definition:

Definition 3: If the group, $\mathscr{G}_a$, associated with a partition a is a subgroup of the group, $\mathscr{G}_b$, associated with the partition b, then one says $a\supseteq b$ (or equivalently $b\subseteq a$). In terms of the structure of partitions $a\supseteq b$ means that a is obtained by joining some (possibly 0) of the clusters of b. This is illustrated in the six-body case by the following example:

$$\left.\begin{array}{llll} a= & (12) & (345) & (6)\\ b= & (12) & (35) & (46)\\ c= & (13) & (3456) & \end{array}\right\} \Rightarrow a\supseteq b,\quad a\not\supseteq c\,. \tag{4}$$

It is easy to show that the relation $\supseteq$ is a partial ordering[19] on the set, $\mathscr{P}$, of partitions of N particles. One can define least upper bounds (unions) and greatest lower bounds (intersections) with respect to this ordering.[20] These operations give the set of partitions, $\mathscr{P}$, with the relation $\supseteq$ a lattice structure. In particular, it is a semimodular lattice[18] often referred to as the partition lattice or Birkhoff lattice. This structure provides a powerful notational and bookkeeping device:[14]

Definition 4: The *union*, $a\cup b$, of the partitions a and b is the finest partition (the one with the most clusters) satisfying both $a\cup b\supseteq a$ and $a\cup b\supseteq b$. The *intersection*, $a\cap b$, is the coarsest partition (the one with the fewest clusters) satisfying both $a\subseteq a\cap b$ and $b\subseteq a\cap b$. These operations are illustrated below:

$$\left.\begin{array}{l} a=(12)(345)(67)\\ b=(123)(45)(67)\end{array}\right\} \Rightarrow \begin{array}{l} a\cup b=(12345)(67)\\ a\cap b=(12)(3)(45)(67)\end{array}. \tag{5}$$

It is easy to show that the unions and intersections as defined above both exist and are unique. The partition $a\cup b$ is associated with the largest symmetry group, $\mathscr{G}_{a\cup b}$, that is a subgroup of both $\mathscr{G}_a$ and $\mathscr{G}_b$. The partition $a\cap b$ corresponds to the smallest symmetry group that contains both $\mathscr{G}_a$ and $\mathscr{G}_b$ as subgroups. For a finite lattice of this form there are unique minimal and maximal elements. These are the intersection of all partitions which is the partition $\underline{0}$ (all particles separate), and the union of all partitions which is the partition $\underline{1}$ (all particles together). $\mathscr{G}_{\underline{0}}$ corresponds to the full translation group that independently changes the coordinates of each particle in our system, and $\mathscr{G}_{\underline{1}}$ corresponds to the group of translations of the center of mass of the entire N-body system. We remark that this lattice structure has been used by Bencze[21] to solve the nontrivial problem of counting the Yakubovskii equations.[22]

3. OPERATORS AND PARTITIONS

The partition lattice can be used to classify translational symmetry properties of physical operators on the N-body Hilbert space. To do this we introduce the notion of connectivity. Intuitively, one thinks of an operator as having connectivity a if it is independent of the positions of the centers of mass of the clusters of a, and vanishes as the relative coordinates of any particles internal to a single cluster of a becomes large. There are many possible ways to formulate these conditions. One convenient characterization is:

Definition 5: A bounded N-body operator, A, is said to have *connectivity* a iff

(i) it commutes with the $3n_a$ parameter unitary group, $\mathscr{G}_a$, of translations of the center of mass of the clusters of a; and

(ii) for every $b\not\supseteq a$

$$s-\lim_{|\mathbf{x}_b|_{\text{rel}}\to\infty} U_b(\mathbf{x}_b)AU_b^\dagger(\mathbf{x}_b)=0\,, \tag{6}$$

where $U_b(\mathbf{x}_b)$ is the element of $\mathscr{G}_b$ corresponding to the displacement of the center of mass of the clusters of b by $\mathbf{x}_b = (\mathbf{x}_{b_1},\ldots,\mathbf{x}_{b_{n_a}})$. $s-\lim_{|\mathbf{x}_b|_{\text{rel}}\to\infty}$ indicates the strong limit as *all* of the *relative* coordinates of the centers of mass of the clusters of b approach infinity.

Condition (ii) is one possible formulation of the asymptotic condition. It is not the weakest possible choice, but is convenient for our purposes. Two-body interactions that are bounded local functions approaching 0 as the relative coordinate approaches infinity are easily seen to have $N-1$ cluster connectivity of this definition, so this condition is not unreasonable. This definition can be extended to unbounded operators with $\mathscr{G}_{\underline{0}}$ invariant domains with some modification. These analytic considerations are of secondary importance to our combinatoric considerations, and will not be discussed further.

If an operator A has connectivity a we write

$$A=[A]_a\,. \tag{7}$$

When $A=[A]_{\underline{1}}$ we say A is *connected*. In N-body scattering theory, connectivity of an operator is a prerequisite for compactness. Specifically, operators that are compact on the relative motion Hilbert space can be shown to be connected by Definition 5. Connectivity guarantees that the basic operator algebra, which is complicated for an N-body system, is done properly. It allows one to *separate* these combinatorial

considerations (which are much easier to verify) from the analytic considerations associated with actually proving compactness.

Most of the operators that we deal with in physics can be represented as a sum of parts with different connectivities.

Definition 6: The set $\mathscr{O}_{\mathscr{G}}$ of *translationally fibered operators* (TFO's)[23] is the set of bounded operators A that admit *cluster decompositions* of the form

$$A = \sum_{a\in\mathscr{P}} [A]_a , \tag{8}$$

where $[A]_a$ is a bounded operator of connectivity a.

Most operators that one deals with in the N-body problem are of this form. In all that follows we will use the following notation for sums over partitions:

$$\sum_{a\in\mathscr{P}} \to \sum_a \quad \sum_{a\in\mathscr{P}'} \to \sum_a' \quad \sum_{a\in\mathscr{P}''} \to \sum_a'' .$$

We observe[23]

Theorem 1: The cluster decomposition of a TFO is unique. To prove this assume $A\in\mathscr{O}_{\mathscr{G}}$ and that A has two-cluster decompositions:

$$A = \sum_a [A]_a = \sum_a [A']_a . \tag{9}$$

From (9) it follows that for any b

$$\sum_a U_b(\mathbf{x}_b)([A]_a - [A']_a)U_b^\dagger(\mathbf{x}_b) = 0 .$$

For $b\supseteq a$, $[A]_a$ and $[A']_a$ will be invariant under $\mathscr{G}_b$, since $\mathscr{G}_b$ is a subgroup of $\mathscr{G}_a$ in this case. This allows us to write (9) as

$$\sum_{a(\subseteq b)} ([A]_a - [A']_a) + \sum_{a(\not\subseteq b)} U_b(\mathbf{x}_b)([A]_a - [A']_a)U_b^\dagger(\mathbf{x}_b) = 0 . \tag{10}$$

Taking the strong limit as $|\mathbf{x}_b|_{\text{rel}} \to \infty$ it follows from (6) and (10) that for every $|\phi>$ our N-body Hilbert space and every $b\in\mathscr{P}$ that

$$\sum_{a(\subseteq b)} ([A]_a - [A']_a)|\phi> = 0 ,$$

or equivalently

$$\sum_{a(\subseteq b)} ([A]_a - [A']_a) = 0 .$$

We write this condition as

$$\sum_a \Delta_{b\supseteq a}([A]_a - [A']_a) = 0$$

where[24]

$$\Delta_{a\supseteq b} \equiv \begin{cases} 1 & \text{if } a\supseteq b \\ 0 & \text{otherwise} \end{cases} . \tag{11}$$

The theorem will be proved if we can show that $\Delta_{a\supseteq b}$ has an inverse as a matrix indexed by partitions.[14] This is shown in the next section when we discuss combinatorics (Theorem 3), which proves Theorem 1. □

A second interesting question concerns what happens when you multiply two TFO's. Intuition says that the connectivity must increase, and this should be reflected by our definition. The relevant result is[23]:

Theorem 2: Let $A, B\in\mathscr{O}_{\mathscr{G}}$ and assume $A = [A]_a$ and $B = [B]_b$, then $AB\in\mathscr{O}_{\mathscr{G}}$ and

$$AB = [AB]_{a\cup b} . \tag{12}$$

Proof: First we note that since $\mathscr{G}_{a\cup b}$ is the largest subgroup of both $\mathscr{G}_a$ and $\mathscr{G}_b$, the product AB is invariant under $\mathscr{G}_{a\cup b}$, and satisfies condition (i) of Def. 5 for $a\cup b$ connectivity. Next, let c be any partition satisfying $c\not\supseteq a\cup b$, then either $c\supseteq b$ and $c\not\supseteq a$ or $c\not\supseteq b$. If $c\subseteq b$ and $c\not\subseteq a$ then B is invariant under $U_c(\mathbf{x}_c)$ and

$$\|U_c(\mathbf{x}_c)ABU^\dagger(\mathbf{x}_c)|\phi>\| = \|U_c(\mathbf{x}_c)AU_c^\dagger(\mathbf{x}_c)\cdot B|\phi>\| = \|U_c(\mathbf{x}_c)AU_c^\dagger(\mathbf{x}_c)|\psi>\| ,$$

where $|\psi> = B|\phi>$. Because $c\not\supseteq a$ and B is bounded this vanishes as $|\mathbf{x}_c|_{\text{rel}} \to \infty$ by the a-connectivity of A. To treat the case $c\not\supseteq b$ note

$$\|U_c(\mathbf{x}_c)ABU_c^\dagger(\mathbf{x}_c)|\phi>\| = \|U_c(\mathbf{x}_c)AU_c^\dagger(\mathbf{x}_c)U_c(\mathbf{x}_c)BU_c^\dagger(\mathbf{x}_c)|\phi>\| \leqslant \|A\|\cdot\|U_c(\mathbf{x}_c)BU_c^\dagger(\mathbf{x}_c)|\phi>\| .$$

Again since A is bounded and $c\not\supseteq b$, this vanishes as $|\mathbf{x}_c|_{\text{rel}} \to \infty$ by the b-connectivity of B. Taken together, these conditions imply that $\forall\ c\not\supseteq a\cup b$

$$s - \lim_{|\mathbf{x}_c|_{\text{rel}}\to\infty} U_c(\mathbf{x}_c)ABU_c^\dagger(\mathbf{x}_c) = 0 .$$

Since the product AB is bounded, it follows from Def. 5 that $AB\in\mathscr{O}_{\mathscr{G}}$ and has connectivity $a\cup b$. □

Corollary: $\forall A,B\in\mathscr{O}_{\mathscr{G}}$, $AB\in\mathscr{O}_{\mathscr{G}}$ and

$$[AB]_c = \sum_{\{a,b\,|\,a\cup b = c\}} [A]_a [B]_b . \tag{13}$$

This is an immediate consequence of Theorem 2 and the cluster decomposition of TFO's.

In the proof of Theorem 1 the combination $\Sigma_a\Delta_{b\supseteq a}[A]_a$ appears. This particular combination appears frequently in applications, and it is useful to have a notation for it.

Definition 7: For $A\in\mathscr{O}_{\mathscr{G}}$

$$A_a \equiv \sum_b \Delta_{a\supseteq b}[A]_b , \tag{14}$$

This is the part of A that is invariant under $\mathscr{G}_a$, including all terms with higher symmetries. It is also useful to have a notation for the residual part.

Definition 8: For $A\in\mathscr{O}_{\mathscr{G}}$

$$A^a \equiv A - A_a . \tag{15}$$

This is the part of A that breaks the $\mathscr{G}_a$ translational symmetry. If we define

$$\bar{\Delta}_{a\supseteq b} \equiv 1 - \Delta_{a\subseteq b} , \tag{16}$$

(15) can be expressed as

$$A^a = \sum_b \bar{\Delta}_{a\supseteq b}[A]_b . \tag{17}$$

The upper-lower partition notation has become standard in N-body theory.[11-17] The decomposition (15) breaks an operator into a part that involves particles in the same clusters of

the partition a, and parts that necessarily involve particles in different clusters.

These combinations are not the only ones that arise naturally. It is also useful to define

Definition 9: For $A\in\mathscr{O}_{\mathscr{G}}$

$$A^{b_1\cdots b_m}_{a_1\cdots a_n} = \sum_c \Delta_{a_1\supseteq d}\cdots\Delta_{a_n\supseteq c}\bar{\Delta}_{b_1\supseteq c}\cdots\bar{\Delta}_{b_m\supseteq c}[A]_c\,. \quad (18)$$

This is the part of A that is invariant under the groups $\mathscr{G}_{a_1}\cdots\mathscr{G}_{a_m}$ and breaks the invariance under each of $\mathscr{G}_{b_1}\cdots\mathscr{G}_{b_m}$. This operator also satisfies a relation corresponding to (15):

$$A^{b_1\cdots b_m}_{a_1\cdots a_n} = A^{b_1\cdots b_m}_{a_1\cdots a_n,c} + A^{b_1\cdots b_m,c}_{a_1\cdots a_n}\,. \quad (19)$$

To illustrate some of these ideas we look at the example of nonrelativistic potential scattering. Let us consider the case $N=5$. If we let $G_0(z)$ be the resolvent of the five-body kinetic energy operator, and V_{ij} the two-body interaction between particles i and j, then the following operator is an example of a five-body operator of connectivity (12) (345)

$$G_0(z)V_{12}G_0(z)V_{35}G_0(z)V_{45}G_0(z)\,. \quad (20)$$

Operators of this type appear in the perturbation expansion for the full five-body resolvent. Note, $G_0(z)$ becomes unbounded as z approaches the scattering cut, however, for our purposes it is enough to consider z in the domain of analyticity of $G_0(z)$.

4. COMBINATORIAL RESULTS

In this section the main results concerning connectivity properties of operators are stated and proved. Although some of these results have appeared elsewhere,[11-14,23] we include them for the purpose of continuity. In the last section the matrices $\Delta_{a\subseteq b}$ and $\bar{\Delta}_{a\subseteq b}$ on the partition indices were introduced. These matrices relate the operators A_a and A^a to their cluster expansions. All of the results are related to simple properties of these matrices. First note:

Theorem 3: Considered as a matrix with partition indices, $\Delta_{a\subseteq b}$ has an inverse with integer elements.[14,23]

Proof: Since $\Delta_{a\supseteq b}$ is a matrix with integer elements, it is enough to show the determinant is unity. To do this, order the partitions left to right in order of increasing number of clusters. Since $a\supseteq b$ implies b is obtained from a by breaking clusters of a, it follows that $n_a\leqslant n_b$. In addition it is easy to see that $a\supseteq b$ and $n_a=n_b$ requires $a=b$. In this representation these remarks clearly require that $\Delta_{a\supseteq b}$ is an upper triangular matrix with ones on the diagonal. The determinant of such a matrix is clearly unity, providing Theorem 3. □

The inverse of $\Delta_{a\supseteq b}$ is denoted by $\Delta^{-1}_{a\supseteq b}$. The elements of $\Delta^{-1}_{a\supseteq b}$ are referred to as the *anticluster coefficients*.[25] The elements of this matrix can be constructed recursively by increasing the number of clusters on the right via

$$\Delta^{-1}_{a\supseteq a} = 1\,, \quad (21)$$

$$\sum_{b(\supseteq c)}\Delta^{-1}_{a\supseteq b} = \delta_{ac}\Rightarrow\Delta^{-1}_{a\supseteq c}\,. \quad (22)$$

Explicit expressions for these coefficients have been given by L'Huillier.[26]

The matrix $\Delta^{-1}_{a\supseteq b}$ allows the construction of some very useful quantities. First we observe that[23]

$$\delta_{ab} = \sum_c{}' \Delta^{-1}_{a\supseteq c}\Delta_{c\supseteq b}\,, \quad (23)$$

where these matrices are considered as matrices on the restricted set of partitions $\mathscr{P}'$. If we sum a over $\mathscr{P}'$ we obtain for any $b\in\mathscr{P}'$

$$1 = \sum_a{}'\delta_{ab} = \sum_{a,c}{}'\Delta^{-1}_{a\supseteq c}\Delta_{c\supseteq b} = \sum_{c(\supseteq b)}{}'\left(\sum_a{}'\Delta^{-1}_{a\supseteq c}\right). \quad (24)$$

This leads us to define the quantity

$$\mathscr{C}_b \equiv \sum_a{}'\Delta^{-1}_{a\supseteq b}\,, \quad (25)$$

from which we deduce for any $b\in\mathscr{P}'$ the following important result:[11-14]

$$\sum_{a(\supseteq b)}{}'\mathscr{C}_a = \sum_a{}'\mathscr{C}_a\Delta_{a\supseteq b} = 1\,. \quad (26)$$

This result is of primary importance in the work of L'Huillier, Redish, and Tandy[11] and in related work.[11-14] It is proved explicitly in Refs. 11–14 using conbinatorial methods where explicit expressions for the coefficients are given. In particular one can prove:

Theorem 4: The coefficients $\mathscr{C}_a$ are uniquely determined by (26) and are exactly

$$\mathscr{C}_a = (-)^{n_a}(n_a-1)! \quad (27)$$

Proof: First note that the $\mathscr{C}_a$'s are uniquely determined by (26). Assume an arbitrary set of coefficients satisfies (26). Multiplying both sides of this equation on the right by $\Delta^{-1}_{b\supseteq c}$ and summing b over $\mathscr{P}'$[27] gives

$$\sum_b{}'\Delta^{-1}_{b\supseteq c} = \sum_{b,a}{}'\mathscr{C}_a\Delta_{a\supseteq b}\Delta^{-1}_{b\supseteq c} = \sum_c{}'\mathscr{C}_a\delta_{ac} = \mathscr{C}_a\,. \quad (28)$$

From this it follows that these $\mathscr{C}_a$'s must be given by (25), proving uniqueness. It then follows that to prove the $\mathscr{C}_a$'s in (25) are given by (27), one needs only show that the coefficients (27) satisfy (26). To do this one introduces the Stirling numbers of the second kind:[28]

Definition 10: The *Stirling number of the second kind*, $\mathscr{S}^m_N$ is the number of ways N distinguishable objects can be grouped into m nonempty disjoint clusters.

The Stirling numbers are well-known combinatorial quantities and satisfy the following properties:[28]

$$\text{(i)}\quad \sum_{m=0}^{n}(-)^{n-m}m!\mathscr{S}^m_n = 1\,, \quad (29)$$

$$\text{(ii)}\quad \mathscr{S}^m_{n+1} = m\mathscr{S}^m_n + \mathscr{S}^{m-1}_n\,, \quad (30)$$

$$\text{(iii)}\quad \mathscr{S}^1_n = \mathscr{S}^n_n = 1,\quad \mathscr{S}^n_0 = \delta_{n0}\,. \quad (31)$$

It is easy to show that these properties imply[13]

$$1 = \sum_{m=2}^{n}(-)^m(m-1)!\,\mathscr{S}^m_n\,. \quad (32)$$

The coefficients (27) explicitly appear in this sum. To obtain the desired result one must write this as a sum over partitions. To do this fix a partition b and consider all partitions $a\in\mathscr{P}'$ satisfying $a\supseteq b$. These partitions are obtained by joining the clusters of b together. For the purpose of counting partitions, the clusters of b can be treated as n_b single particles. It follows that

$$\sum_{a(\supseteq b)} (-)^{n_a}(n_a-1)!$$

$$= \sum_{n_a=2}^{n_b} (-)^{n_a}(n_a-1)!$$

$$\times\{\#\text{ of } n_a \text{ cluster partitions of } n_b \text{ particles}\}, \tag{33}$$

which is by definition

$$\sum_{n_a=2}^{n_b} (-)^{n_a}(n_a-1)!\ \mathscr{S}^{n_a}_{n_b}. \tag{34}$$

Since this expression is 1 by (32), it follows that $\mathscr{C}_a = (-)^{n_a}(n_a-1)!$ satisfies (26). The uniqueness part of this proof allows these numbers to be identified with the coefficients defined in (25). □

A great deal of effort has been made to compute the value of the coefficients $\mathscr{C}_a$, The reason for this is that these numbers are involved in some very important theorems. In order to collect these results in one place, we state all of these theorems before proving any of them. The are

Theorem 5: $\forall A\in\mathscr{O}_{\mathscr{S}}$

$$[A]_{\rm DISC} \equiv \sum_a{}' [A]_a = \sum_a{}' \mathscr{C}_a A_a \quad (\text{any } A), \tag{35}$$

Theorem 6: $\forall A\in\mathscr{O}_{\mathscr{S}}$

$$[A]_{\underline{1}} = \sum_a{}' \mathscr{C}_a A^a \quad (\text{any } A), \tag{36}$$

Theorem 7: $\forall A, B\in\mathscr{O}_{\mathscr{S}}$

$$\sum_a{}' \mathscr{C}_a A_a B^a \text{ is 0 or connected}. \tag{37}$$

These theorems are the fundamental results of this paper. Theorems 5 and 7 are generalizations of results in Refs. 14 and 23 and related to results appearing in Ref. 11. They admit genrealizations corresponding to choosing the operators A and B as products of operators having several upper and lower indices. Theorem 7 also holds if the order of A and B is reversed. The proofs of these theorems utilize the cluster expansions (14), (17), and (26).

Proof of Theorem 5:

$$\sum_a{}' \mathscr{C}_a A_a = \sum_a{}' \mathscr{C}_a \sum_{b(\subseteq a)} [A]_b = \sum_b{}' [A]_b \sum_{a(\supseteq b)} \mathscr{C}_a$$

$$= \sum_b{}' [A]_b \cdot 1,$$

which is by definition [i.e., (35)] $[A]_{\rm DISC}$. □

Proof of Theorem 6: Theorem 5 and (15) yield

$$A = \sum_a{}' \mathscr{C}_a A = \sum_a{}' \mathscr{C}_a (A_a + A^a)$$

$$= [A]_{\rm DISC} + \sum_a{}' \mathscr{C}_a A^a,$$

which implies

$$[A]_{\underline{1}} = A - [A]_{\rm DISC} = \sum_a{}' \mathscr{C}_a A^a. \quad \square$$

Proof of Theorem 7: Equations (14) and (17) yield

$$\sum_a{}' \mathscr{C}_a A_a B^a$$

$$= \sum_a{}' \mathscr{C}_a \sum_b{}' \Delta_{a\supseteq b}[A]_b \sum_c \bar{\Delta}_{a\supseteq b}[B]_c$$

$$= \sum_b{}' \sum_c [A]_b [B]_c \sum_a{}' \mathscr{C}_a \Delta_{a\supseteq b}(1-\Delta_{a\supseteq b})$$

$$= \sum_b{}' \sum_c [A]_b [B]_c \sum_a{}' \mathscr{C}_a (\Delta_{a\supseteq b} - \Delta_{(b\cup c)\subseteq a}).$$

Since $n_a \geqslant 2$ by the fact that the a sum is primed, (26) implies $\Sigma_a{}' \mathscr{C}_a \Delta_{a\supseteq b} = 1$. If $b\cup c$ has at least two clusters then we also have $\Sigma_a{}' \mathscr{C}_a \Delta_{b\cup c\subseteq a} = 1$; but if $b\cup c = 1$ then there are no a's with $n_a \geqslant 2$ satisfying $b\cup c\subseteq a$ and the sum is 0. Therefore we obtain

$$\sum_a{}' \mathscr{C}_a A_a B^a = \sum_b{}' \sum_c [A]_b [B]_c \cdot [1-(1-\delta_{n_{b\cup c},1})]$$

$$= \sum_b{}' \sum_c [A]_b [B]_c \delta_{n_{b\cup c},1}.$$

This sum is 0 or connected because $A, B\in\mathscr{O}_{\mathscr{S}}$ by (12). □

This theorem is extremely useful because it allows us to construct connected operators in a few simple steps. In the next section some examples of applications of these theorems to N-body theory are given.

Recall that the existence $\Delta^{-1}_{b\supseteq a}$ means that we can freely transform between the sets of operators $\{A_a\}_{a\in\mathscr{P}}$ and $\{[A]_a\}_{a\in\mathscr{P}}$. There is a similar relation between these sets of operators and the set $\{A^a\}$ with one minor qualification.[29] First note that $[A]_0$ is contained in none of the A^a's, while $[A]_1$ is contained in all of the (nontrivial) A^a's. Thus neither $[A]_0$ nor $[A]_1$ can be recovered from the set $\{A^a\}$. However, if both $[A]_0$ and $[A]_1$ vanish, then both of the sets $\{A_a\}_{a\in\mathscr{P}''}$ and $\{[A]_a\}_{a\in\mathscr{P}''}$ can be obtained. To do this one must construct the inverse of $\bar{\Delta}_{a\supseteq b}$ on $\mathscr{P}''\times\mathscr{P}''$. One has[23]

Theorem 8: $\bar{\Delta}^{-1}_{a\supseteq b}$ exists as a matrix indexed by the set of partitions $\mathscr{P}''$.

Proof: The proof of this theorem uses the existence of $\Delta^{-1}_{a\supseteq b}$ as a matrix on the set of partitions $\mathscr{P}''$, and the explicit expressions for $\mathscr{C}_a$ given by (27). We first observe that

$$\sum_a{}'' \mathscr{C}_a = \sum_a{}' \mathscr{C}_a - \mathscr{C}_0 = 1-(-)^N(N-1)! \tag{38}$$

and

$$\Delta_{a\supseteq b} = 1 - \bar{\Delta}_{a\supseteq b}. \tag{39}$$

We multiply (39) by $\mathscr{C}_a$ and sum a over $\mathscr{P}''$ for $b\in\mathscr{P}''$ which gives

$$\sum_a{}'' \mathscr{C}_a \Delta_{a\supseteq b} = \sum_a{}'' \mathscr{C}_a - \sum_a{}'' \mathscr{C}_a \bar{\Delta}_{a\supseteq b}. \tag{40}$$

Equations (38), (26), and the fact that $b\in\mathscr{P}''$ can be used to obtain

$$1 = 1-(-)^N(N-1)! - \sum_a{}'' \mathscr{C}_a \bar{\Delta}_{a\supseteq b}$$

or

$$1 = \frac{(-)^{N-1}}{(N-1)!} \cdot \sum_a{}'' \mathscr{C}_a \bar{\Delta}_{a\supseteq b}. \tag{41}$$

If we insert (41) in (39) we obtain

$$\Delta_{a\supseteq b} = \sum_c{}'' \left(\frac{(-)^{N-1}}{(N-1)!}\mathscr{C}_c - \delta_{ac}\right)\bar{\Delta}_{c\supseteq b}. \tag{42}$$

To complete the proof we multiply (42) by $\Delta^{-1}_{d\supseteq a}$ and sum a over $\mathscr{P}''$ giving

$$\delta_{db} = \sum_c^{''}\left(\sum_a^{''}\frac{(-)^{N-1}}{(N-1)!}\Delta^{-1}_{d\supseteq a}\mathscr{C}_c - d\Delta^{-1}_{d\supseteq a}\delta_{ac}\right)\bar{\Delta}_{a\supseteq b}\,, \tag{43}$$

which immediately implies

$$\bar{\Delta}^{-1}_{a\supseteq b} = \sum_c^{''}\frac{(-)^{N-1}}{(N-1)!}\Delta^{-1}_{a\supseteq c}\mathscr{C}_b - \Delta^{-1}_{a\supseteq b}\,. \quad \square \tag{44}$$

Theorem 8 provides a means for constructing $\{[A]_a\}_{a\in\mathscr{P}''}$ and $\{A_a\}_{a\in\mathscr{P}''}$ from $\{A^a\}_{a\in\mathscr{P}''}$ when $[A]_{\underline{0}} = [A]_{\underline{1}} = 0$. In particular one may define "raising" and "lowering" operators by:

$$R^{ab} = \sum_c \bar{\Delta}_{a\supseteq c}\Delta^{-1}_{c\supseteq b}\,, \tag{45}$$

$$L_{ab} = \sum_c^{''}\Delta_{a\supseteq c}\bar{\Delta}^{-1}_{c\supseteq b}\,. \tag{46}$$

The complete set of relations is

$$A^a\sum_b\bar{\Delta}_{a\supseteq b}[A]_b = \sum_b R_{ab}A_b\,, \tag{47}$$

$$A_a = \sum_b\Delta_{a\supseteq b}[A]_b = {}^{\star}\sum_b^{''}L_{ab}A^b\,, \tag{48}$$

$$[A]_a = \sum_b\Delta^{-1}_{a\supseteq b}A_b = {}^{\star}\sum_b^{''}\bar{\Delta}^{-1}_{a\supseteq b}A^b\,. \tag{49}$$

It follows that the three sets of operators, $\{A_a\}$, $\{A^a\}$, and $\{[A]_a\}$ contain exactly the same information (*if $[A]_{\underline{0}} = [A]_{\underline{1}} = 0$).

By summing $\bar{\Delta}^{-1}_{a\supseteq b}$ over the left index one obtains coefficients analogous to the $\mathscr{C}_a$'s of (25):

$$\bar{\mathscr{C}}_a \equiv \sum_b^{''}\bar{\Delta}^{-1}_{b\supseteq a}\,. \tag{50}$$

This definition clearly leads to the sum rule, $\forall a\in\mathscr{P}''$

$$\sum_{b(\not\supseteq a)}^{''}\bar{\mathscr{C}}_b = \sum_b^{''}\bar{\mathscr{C}}_b\bar{\Delta}_{a\supseteq b} = 1\,. \tag{51}$$

The coefficients $\bar{\mathscr{C}}_a$ can be explicitly computed. The result is given by

Theorem 9:

$$\bar{\mathscr{C}}_a = -\frac{\mathscr{C}_a}{\mathscr{C}_0} = -\frac{(-)^{n_a}(n_a-1)!}{(-)^N(N-1)!}\,. \tag{52}$$

Proof: We use (45), (44), (26), and note $b,c\neq\underline{0}$ which gives

$$\begin{aligned}\bar{\mathscr{C}}_a &= \sum_c^{''}\sum_a^{'}\frac{(-)^{N-1}}{(N-1)!}\Delta^{-1}_{a\supseteq c}\mathscr{C}_b - \sum_a^{'}\Delta^{-1}_{a\supseteq b}\\ &= \sum_c^{''}\sum_a^{'}\frac{(-)^{N-1}}{(N-1)!}\Delta^{-1}_{a\supseteq c}\mathscr{C}_b - \sum_a^{'}\Delta^{-1}_{a\supseteq b}\\ &= \left(\sum_c^{''} - \frac{1}{(-)^N(N-1)!}\mathscr{C}_c\mathscr{C}_b - \mathscr{C}_b\right)\\ &= -(\mathscr{C}_b/\mathscr{C}_a)\cdot\left\{\sum_c^{'}\mathscr{C}_c - \mathscr{C}_0\right\} - \mathscr{C}_b\\ &= -\mathscr{C}_b/\mathscr{C}_a + \mathscr{C}_b - \mathscr{C}_b = -\mathscr{C}_b/\mathscr{C}_a\,. \quad \square\end{aligned}$$

The theorems corresponding to Theorems 5, 6, and 7 are

Theorem 10: For any A with $[A]_{\underline{1}} = 0$

$$\sum_a^{''}\bar{\mathscr{C}}_aA^a = A - [A]_{\underline{0}}\,, \tag{53}$$

Theorem 11: For any A with $[A]_{\underline{0}} = 0$

$$\sum_a^{''}\bar{\mathscr{C}}_aA_a = -\frac{1}{\mathscr{C}_{\underline{0}}}[A]_{\rm DISC}\,, \tag{54}$$

Theorem 12: Let $A,B\in\mathscr{O}_{\not\supseteq}$ and assume that $[A]_{\underline{0}} = 0$, then

$$\sum_a^{'}\bar{\mathscr{C}}_aA_aB^a \text{ is 0 or connected.} \tag{55}$$

Proof of Theorem 10: For $[A]_{\underline{1}} = 0$

$$\begin{aligned}&\sum_a^{''}\bar{\mathscr{C}}_aA^a\\ &= \sum_a^{''}\sum_b\bar{\mathscr{C}}_a\bar{\Delta}_{a\supseteq b}[A]_b = \sum_a^{''}\sum_b^{'}\bar{\mathscr{C}}_a\bar{\Delta}_{a\supseteq b}[A]_b\\ &= \sum_b^{''}[A]_b\cdot\left(\sum_a^{''}\bar{\mathscr{C}}_a\bar{\Delta}_{a\supseteq b}\right) + [A]_{\underline{0}}\cdot\sum_a^{''}\bar{\mathscr{C}}_a\bar{\Delta}_{a\supseteq\underline{0}}\\ &= \sum_b^{''}[A]_b\cdot 1 = A - [A]_{\underline{0}}\,. \quad \square\end{aligned}$$

Note the fact that $[A]_{\underline{1}} = 0$ has been used.

Proof of Theorem 11: For $[A]_{\underline{0}} = 0$, using Theorem 9

$$\begin{aligned}\sum_a^{''}\bar{\mathscr{C}}_aA_a &= -\frac{1}{\mathscr{C}_{\underline{0}}}\sum_a^{''}\mathscr{C}_aA_a = -\frac{1}{\mathscr{C}_{\underline{0}}}\sum_a^{'}\mathscr{C}_aA_a\\ &= -\frac{1}{\mathscr{C}_{\underline{0}}}[A]_{\rm DISC}\,,\end{aligned}$$

by Theorem 5. $\square$

Proof of Theorem 12: From Theorem 9 we have

$$\sum_a^{''}\bar{\mathscr{C}}_aA_aB^a = -\frac{1}{\mathscr{C}_{\underline{0}}}\left(\sum_a^{'}\mathscr{C}_aA_aB^a - \mathscr{C}_{\underline{0}}A_{\underline{0}}B^{\underline{0}}\right).$$

Since $A_{\underline{0}} = [A]_{\underline{0}} = 0$, and $\Sigma_a'\,\mathscr{C}_a A_a B^a$ is 0 or connected by Theorem 7, the result follows. $\square$

Under the same restrictions, these theorems also admit generalizations corresponding to multi-indexed products of operators. Theorems 2–12 are the main combinatorial result. They provide a simple means for controlling connectivity and quantities related to connectivity.

5. EXAMPLES–THE *N*-BODY PROBLEM

In this section some of these results are applied to operators of nonrelativistic N-body scattering theory to recover some well-known results.

A. Example 1: The Van Winter–Weinberg Equation

The Van Winter[1]–Weinberg[2] equation is a compact kernel equation for the resolvent of an N-body Hamiltonian H. Historically it was the first N-body generalization of Faddeev's work on three-body scattering.[3] The methods of the last section can be used to give a one-line derivation of this equation with explicit expressions for the driving term and kernel. This derivation has the additional advantage that it is valid even in the presence of multiparticle forces. To establish the notation one defines an N-body Hamiltonian of the form

$$H = K + V\,, \tag{56}$$

where K is the N-body kinetic energy operator satisfying $K = [K]_0$, and V is the sum of all interparticle interactions. We ignore the complications due to the unbounded nature of K. One defines the resolvent of H

$$G(z) = (z - H)^{-1}. \tag{57}$$

It follows that

$$H_a = K + V_a, \tag{58}$$

$$G_a(z) = (z - H_a)^{-1}, \tag{59}$$

and for any a

$$G(z) = G_a(z) + G_a(z)V^aG(z). \tag{60}$$

Equation (60) is the familiar resolvent relation in the notation of Sec. 3. We use (26) with $b = 0$ and (60) to obtain

$$\begin{aligned} G(z) &= \sum_a{}' \mathscr{C}_a G(z) \\ &= \sum_a{}' \mathscr{C}_a G_a(z) + \sum_a{}' \mathscr{C}_a G_a(z)V^aG(z). \end{aligned} \tag{61}$$

This equation is the Van Winter–Weinberg equation. The driving term is the disconnected part of the resolvent, by Theorem 5. This is a well-known property of the Van Winter–Weinberg equation for the resolvent. The kernel, $\Sigma_a' \mathscr{C}_a G_a(z) V^a$, is connected by Theorem 7. Furthermore, it is irreducible[2] in the sense that on removing the last interaction on the right (V^a) of the kernel, the resulting operator is no longer connected.

B. Example 2: The precursor Bencze–Redish–Sloan equations

In terms of the notation of the last example, the transition operators for multichannel scattering theory can be expressed as

$$T_+(a,b,z) = V^a + V^aG(z)V^b. \tag{62}$$

This can be rewritten using the resolvent relations (58) as

$$T_+(a,b,z) = V^aG(z)G_b^{-1}(z). \tag{63}$$

Assuming $[V]_1 = 0$ (no N-body forces) one can use the appropriate generalization of Theorem 5 to write

$$\begin{aligned} T_+(a,b,z) &= \sum_c{}' \mathscr{C}_c V_c^a G(z) G_b^{-1} \\ &= \sum_c{}' \mathscr{C}_c V_c^a G_c(z)[1 + V^cG(z)]G_b^{-1}(z), \\ T_+(a,b,z) &= \sum_c{}' \mathscr{C}_c V_c^a G_c(z) G_b^{-1}(z) \\ &\quad + \sum_c{}' \mathscr{C}_c V_c^a G_c(z) T_+(c,b,z). \end{aligned} \tag{64}$$

This is an equation derived by L'Huillier, Redish, and Tandy[11] (using these same methods) that was shown to be equivalent to the Bencze–Redish–Sloan equations[30–32,25] when V is a sum of two-body interactions. Equation (64) generally holds under the condition that $[V]_1 = 0$. The iterated kernal has the form

$$K^2(a,b) = \sum_c{}' \mathscr{C}_c V_c^a G(z) V_b^c \mathscr{C}_b G_b(z),$$

which is connected by the appropriate generalization of Theorem 7.

C. Example 3: Hamiltonian decomposition theorem

The application of Theorem 5 to the full Hamiltonian gives (in the absence of N-body forces, again ignoring complications arising from the unbounded nature of K)

$$H = \sum_a{}' \mathscr{C}_a H_a, \tag{65}$$

which allows one to express an N-body Hamiltonian in terms of partition Hamiltonians associated with proper subsystem problems. This result has been used by Redish and the author in developing an N-body theory admitting unitary truncations.[14,23]

6. CONCLUDING REMARKS

In the previous section it was shown how the theorems of Sec. 4 allowed us to derive well-known results with little effort. The interesting thing about this basic combinatoric framework is that many of the ideas have appeared in one form or another over the years in several disciplines. In particular, the coefficients $-\mathscr{C}_a$ appear in many applications. Typically one is given a set of operators of the type $\{A_a\}$ and one wants to construct operators of the type $\{[A]_a\}$. The operators $\{[A]_a\}$ are generally referred to as *cumulants*[7] of the operators $\{[A]_a\}$. Theorem 5 is a statement of the basic result[33]

$$[A]_1 = A - \sum_a{}' \mathscr{C}_a A_a = -\sum_a{}' \mathscr{C}_a A_a, \tag{66}$$

that is used in constructing these cumulants. The standard examples are

(i) from quantum field theory[8–10]

$\{A_a\} \longleftrightarrow \{\text{Wightman functions}\}$,

$\{[A]_a\} \longleftrightarrow \{\text{Truncated vacuum expectation values}\}$,

(ii) from statistical mechanics[4–6]

$\{A_a\} \longleftrightarrow \{\text{Mayer functions}\}$,

$\{[A]_a\} \longleftrightarrow \{\text{Ursell cluster functions}\}$.

Similar ideas have also been employed by Kowalski in N-body scattering theory.[16]

Both of these examples can be considered as simple cases of Theorem 5. The real value of these theorems is in constructing connected kernel equations for physical operators. Many evolution equations of the form

$$\begin{aligned} &\frac{d}{dt}|X(t)\rangle = -D|X(t)\rangle + |f(t)\rangle, \\ &|X(0)\rangle = |X_0\rangle, \end{aligned} \tag{67}$$

where D admits a cluster expansion of the form (8), can be expressed in terms of an equivalent time-independent integral equation for the resolvent of D or an equivalent operator (such as the transition operators of scattering theory). The theorems of Sec. 4 are indispensabel tools for developing equivalent compact kernel integral equations.[34] In general, given any integral equation whose kernel has a suitable clus-

ter expansion, the methods can be used to construct equivalent connected kernel equations.

ACKNOWLEDGMENTS

The author acknowledges the U.S. Department of Energy who supported this work. Helpful discussions and/or correspondence with the following individuals were gratefully appreciated: Gy. Bencze, K.L. Kowalski, E.F. Redish, B. Simon, and I.H. Sloan.

[1]C. Van Winter, Mat. Fys. Skr. Danske Vid. Selsk. **1**, 1 (1964).
[2]S. Weinberg, Phys. Rev. **133**, 232 (1964).
[3]L.D. Faddeev, *Mathematical Aspects of the Three-Body Problem in the Quantum Scattering Theory*, Israel Program for Scientific Translations, 1965.
[4]H.D. Ursell, Proc. Cambridge Philos. Soc. **23**, 685 (1927).
[5]J.E. Mayer and M.G. Mayer, *Statistical Mechanics* (Wiley, New York, 1941).
[6]J.K. Percus, Commun. Math. Phys. **40**, 283 (1975).
[7]H. Cramer, *Mathematical Methods of Statistics* (Princeton University, Princeton, New Jersey, 1956), p. 185.
[8]W. Zimmerman, Nuovo Cimento **13**, 503 (1959).
[9]R. Jost, *The General Theory of Quantized Fields* (A.M.S., Providence, Rhode Island, 1965), p. 66.
[10]N.N. Bogolubov, A.A. Lognunov, and I.T. Tordorov, *Introduction to Axiomatic Quantum Field Theory* (Benjamin, New York, 1975), p. 272–82.
[11]M. L'Huillier, E.F. Redish, and P.C. Tandy, unpublished.
[12]P. Benoist-Gueutal, M. L'Huillier, E.F. Redish, and P.C. Tandy, Phys. Rev. C **17**, 1924 (1978).
[13]K.L. Kowlaski, Phys. Rev. C **16**, 7 (1977).
[14]W. Polyzou and E.F. Redish, to be published in Ann. Phys. (N.Y.).
[15]Gy. Bencze and P.C. Tandy, Phys. Rev. C **16**, 564 (1977).
[16]K.L. Kowalski, to be published in Ann. Phys. (N.Y.).
[17]K.L. Kowalski, E. Siciliano, and R. Thaler, Phys. Rev. C **18**, 2017 (1978).
[18]B. Simon, private communication.
[19]G. Birkhoff, *Lattice Theory* (A.M.S. Colloquium Pub., 1949), Vol. 25.
[20]Ref. 19, p. 100.
[21]Gy. Bencze, Phys. Lett. B **72**, 155 (1977).
[22]O.A. Yakubovskii, Sov. J. Nucl. Phys. **5**, 937 (1967).
[23]W. Polyzou, University of Maryland, Ph.D. thesis, 1979.
[24]The use of the matrix Δ was introduced independently by M. L'Huillier (unpublished) and K.L. Kowalski (Ref. 13).
[25]E.F. Redish, Nucl. Phys. A **235**, 82 (1974).
[26]M. L'Huillier, unpublished.
[27]From the nature of the recurrence relation it follows that $\Delta^{-1}_{a\supseteq c}$ considered as a matrix on $\mathscr{P}'$ (respectively, $\mathscr{P}''$) is just the restriction of $\Delta^{-1}_{a\supseteq c}$ to $\mathscr{P}'\times\mathscr{P}'$ (respectively, $\mathscr{P}''\times\mathscr{P}''$). That $\Delta^{-1}_{a\supseteq c}$ exists on $\mathscr{P}'$ or $\mathscr{P}''$ follows from the proof of Theorem 1.
[28]M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions* (N.B.S., Washington, D.C., 1964), p. 824.
[29]K.L. Kowalski, private communication.
[30]I.H. Sloan, Phys. Rev. C **6**, 1945 (1972).
[31]Gy. Bencze, Nucl. Phys. A **210**, 568 (1973).
[32]E.F. Redish, Nucl. Phys. A **225**, 16 (1974).
[33]Where in this case $\mathscr{C}_1 = (-)^{-1}(0)! = -1$ in accordance with (27).
[34]T. Kato, *Perturbation Theory for Linear Operators* (Springer–Verlag, Berlin, 1966), p. 477–95.

Field equations in twistors

Asghar Qadir[a)]

International Centre for Theoretical Physics, Trieste, Italy
(Received 1 December 1978)

As part of the twistor program for the "quantization of relativity" the physical field equations have been translated into twistors.

I. INTRODUCTION

There have been numerous programs for trying to combine quantum mechanics and the general theory of relativity at various levels,[1] from the attempt to try to develop a quantum theory of gravity to the attempt to quantize the underlying space–time. A particularly hopeful approach of the latter type is the twistor program for the "quantization of relativity"[1-3] of Penrose. It is based on a search for a theory which yields quantum theory and relativity theory as special limits of a more general theory.[4,5] A particularly hopeful feature of twistors is that they appear naturally in the solution of the zero mass field equations.[6] There is further reason to hope that they will prove to be satisfactory for the purpose, as it appears that the quantum electrodynamic divergences may be avoided in this formalism when using Penrose graphs.[7-10]

In this paper, as part of the twistor "quantization" program, we shall be translating the field equations of physics from spinors into twistors. For this purpose we shall be using a set of four twistors, chosen in a particular way, which we call the "twistor quadrad." The equations are written as scalar equations with the fields themselves being expressed as scalar functions of the twistor quadrad. These functions are required to be of the type which are eigenfunctions of the Casimir operators of the Lie algebra of U(2,2).[8,11]

It is hoped that the solutions of these equations will lead to a deeper, and more thorough, understanding of Penrose graphs.[9] In particular, they should lead to a more rigorous basis for the apparently ad hoc (at present) "rules" for writing and calculating these graphs. There remains, at present, some ambiguity regarding the choice of contours for calculating these graphs, which may be eliminated by a better understanding of the fields represented by these graphs. We shall not be studying the solutions of the field equations here, however.

II. THE FIELD EQUATIONS IN SPINORS

It has been shown[6] that the source-free, zero rest-mass field equation for a spin $n/2$ field ($n\neq 0$) can be written as

$$\nabla_{AA'}\varphi^{A\cdots N}=0, \tag{2.1}$$

$A,\ldots,N$ being n indices. For a spin-zero field we have the Klein–Gordon equation

$$\nabla^{AA'}\nabla_{AA'}\varphi=0. \tag{2.2}$$

Thus, for $n=1$, Eq. (2.1) becomes the Dirac equation for massless particles, for $n=2$ it becomes the source-free Maxwell equation, and for $n=4$ the linearized Einstein free-field equations.

Contour integral solutions of these equations were found by Penrose.[4,6] He showed[6,7] that they could be expressed as contour integrals of a linear combination of two twistors U^α and V^α of the form

$$\begin{aligned} U^\alpha &= o_A e^{A\alpha} - x^{AA'} o_A e_{A'}{}^\alpha, \\ V^\alpha &= \iota_A e^{A\alpha} - i x^{AA'} \iota_A e_{A'}{}^\alpha, \end{aligned} \tag{2.3}$$

where o^A and ι^A are the spinor basis and $e_A{}^\alpha$, $e^{A'\alpha}$ are basis spinstors[8,12] (injection and projection operators from the *spin*or space to the twi*stor* space). Then Eq. (2.1) becomes

$$\frac{\partial\varphi_r}{\partial V^\alpha}=\frac{\partial\varphi_{r+1}}{\partial U^\alpha}, \tag{2.4}$$

where φ_r is the component of $\varphi^{A\cdots N}$ have r "1's" and $n-r$ "0's."

The massive field equations in spinors are more complicated. Here the Klein–Gordon equation is

$$(\nabla^{AA'}\nabla_{AA'}+m^2)\varphi=0. \tag{2.5}$$

The Dirac equations are a pair of coupled equations

$$\nabla_{AA'}\alpha^A=(m/\sqrt{2})\beta_{A'},\quad \nabla_{AA'}\beta^{A'}=(m/\sqrt{2})\alpha_A. \tag{2.6}$$

The Maxwell equations are expressed in terms of the spinor field φ^{AB} and its dual field $\tilde{\varphi}^{AB}$

$$\nabla_{AA'}\varphi^{AB}=2\pi J_{A'}{}^B,\quad \nabla_{AA'}\tilde{\varphi}^{AB}=0. \tag{2.7}$$

It should be noted that the solutions of these equations lie in a complex space–time, in general (see, for example, Ref. 13). The linearized gravitational field equations can similarly be written. However, they do not give any additional insight at this stage, and so they are not considered here.

The solutions of the massive field equations, with source terms present, generally have both unprimed and primed spinor indices—$\varphi^{A\cdots NA'\cdots M'}$. We want to write them in terms of twistors. For this purpose we must bear in mind the fact that we need to include conformal transformations in our formalism. Thus the fields are functions of position and of the conformal factor. If we are dealing with a complexified Minkowski space, as we do in twistors, the conformal factor Ω^2 will also be complex. When dealing with con-

a)On leave of absence from Mathematics Department, Quaid-i-Azam University, Islamabad, Pakistan. Present address: Physics Department, University of Texas, Austin, Texas.

 0022-2488/80/030514-07$1.00

formal transformation in spinors, we need to transform ϵ_{AB} and $\epsilon_{A'B'}$ separately by multiplying by the square root of Ω^2. Taking Ω^2 to be complex, we would multiply ϵ_{AB} and $\epsilon_{A'B'}$ by χ and $\tilde{\chi}$, respectively, such that $\chi\tilde{\chi} = \Omega^2$ and $\tilde{\chi} \neq \bar{\chi}$ in general.

To translate the fields $\varphi^{A\cdots NA\cdots M'}(x^{PP'}, \chi, \tilde{\chi})$ into twistors, we need to remove the spinor indices. One way of doing this is to convert the fields into scalars by contracting in two spinors ξ_Q, $\tilde{\xi}_{Q'}$ so as to saturate the indices ($\bar{\xi}_{Q'} \neq \tilde{\xi}_{Q'}$ in general), obtaining

$$F(x^{PP'}, \xi_Q, \tilde{\xi}_{Q'}, \chi, \tilde{\chi}) = \varphi^{A\cdots NA'\cdots M'} \xi_A \cdots \xi_N \tilde{\xi}_{A'} \cdots \tilde{\xi}_{M'}, \quad (2.8)$$

where ξ_A, $\tilde{\xi}_{A'}$ are variable spinor fields. This function will be homogeneous of degree n and m in ξ_Q and $\tilde{\xi}_{Q'}$, respectively, there being n unprimed and m primed indices. If the fields are to be dealt with under conformal transformations, they must have a "conformal weight" related to the degree of homogeneity of the field in χ and $\tilde{\chi}$. Thus we shall require that the functions be homogeneous in χ and $\tilde{\chi}$.

The field equations in terms of the spinor fields (φ's) can be converted to field equations in terms of F by differentiating it (F) with respect to ξ_Q and $\tilde{\xi}_{Q'}$. (Since φ is symmetric in the unprimed indices and in the primed indices, only one differentiation is required.) Thus if

$$\nabla_{AP'} \varphi^{A\cdots NA'\cdots M'} = 0. \quad (2.9a)$$

we shall now have

$$\nabla_{AP'} \frac{\partial}{\partial \xi_A} F = 0. \quad (2.9b)$$

Similarly, if we had

$$\nabla_{AP'} \nabla_{QA'} \varphi^{A\cdots NA'\cdots Q'} = 0, \quad (2.10a)$$

we shall now have

$$\nabla_{AP'} \nabla_{QA'} \frac{\partial}{\partial \xi_A} \frac{\partial}{\partial \tilde{\xi}_{A'}} F = 0. \quad (2.10b)$$

We are now in a position to set up the twistor formalism for writing the field equations by counting the number of variables involved

III. THE TWISTOR QUADRAD

The fields are functions of the position, $x^{PP'}$ consisting of four independent complex variables (as $x^{PP'} \neq \bar{x}^{PP'}$ in general), the spinors ξ_Q and $\tilde{\xi}_{Q'}$, consisting of four independent complex variables (two for each spinor) and the *conformal functions*, as we shall call them, χ and $\tilde{\chi}$, consisting of two complex variables (one for each function). Thus the fields are functions of ten independent complex variables. We want to write these fields in terms of twistors in a "natural" way.

Noting that $10 = 4 + 3 + 2 + 1$, we could try writing the fields as functions of four twistors W^α, X^α, Y^α, and Z^α, with Z contributing four variables, Y three, X two, and W one. This may be done by restricting the functions so that Y only occurs skewed with Z, X only occurs skewed with Y and Z, and W only occurs skewed with X, Y, and Z. We shall call this set of four twistors the *twistor quadrad*. We shall write it as

$$\eta^{a\alpha} = (W^\alpha, X^\alpha, Y^\alpha, Z^\alpha) \quad (a = 0,1,2,3). \quad (3.1)$$

In terms of the space of all twistors we may think of Z as a given point, Y as any point on a given line through Z, X as a point in a given plane passing through the line containing Y and Z, and W as a point in a given 3-hypersurface containing the plane which contains X, Y, and Z. Thus there are four independent complex variables specifying Z. Of the four independent complex variables specifying Y, one is not relevant as it picks out a particular point on the given line. Similarly, two of the four independent complex variables specifying X are not relevant as they pick a point out of a plane, and three of the four independent complex variables specifying W are not relevant as they pick a point out of a 3-space. Thus six of the 16 independent complex variables contained in the twistor quadrad are irrelevant. We thus have ten independent relevant complex variables in the twistor quadrad. This being the number of independent complex variables on which the fields depend, we should be able to write the fields as functions of the twistor quadrad subject to the above-mentioned restrictions.

To see what we are doing in terms of a complex Minkowski space, we define the *dual twistor quadrad* $\tilde{\eta}_{a\alpha}$, the inverse of $\eta^{a\alpha}$, i.e.,

$$\tilde{\eta}_{a\alpha} \eta^{b\alpha} = \delta_a^b, \quad (3.2a)$$

or, equivalently,

$$\tilde{\eta}_{a\alpha} \eta^{a\beta} = \delta_\alpha^\beta. \quad (3.2b)$$

We shall call the twistor quadrad *real* if

$$\tilde{\eta}_{a\alpha} = \bar{\eta}_{a\alpha}, \quad (3.3a)$$

which would be written explicitly as

$$\tilde{Z}_\alpha = \bar{Z}_\alpha, \quad \tilde{Y}_\alpha = \bar{Y}_\alpha, \quad \tilde{X}_\alpha = \bar{X}_\alpha, \quad \tilde{W}_\alpha = \bar{W}_\alpha. \quad (3.3b)$$

The real twistor quadrad may be thought of as four null rays in a real Minkowski space, such that each ray intersects two of the other three but not the third. Thus, Z and W do not intersect and X and Y do not intersect. In the twistor picture all four twistors of the real quadrad lie in N, the space of null twistors. The lines joining Z and Y, Z and X, Y and W, and X and W lie in N, but the lines joining Z and W and Y and X do not lie in N.

For the general twistor quadrad there is no clear picture in the twistor space because every pair of twistors can be connected by a line in the twistor space. However, we can visualize it more easily in a complex Minkowski space as merely a "complexification" of the real twistor quadrad.

We shall require that the functions of the twistor quadrad, representing the fields, be homogeneous in all four twistors, so as to ensure homogeneity of the fields in the spinors ξ_Q and $\tilde{\xi}'_Q$ and in the conformal functions χ and $\tilde{\chi}$ (as we shall see later). Writing the fields as functions of twistors as $f(\eta^{c\gamma})$, it can be shown[8] that they must satisfy the conditions

$$\eta^{a\alpha} \partial_{b\alpha} f(\eta^{c\gamma}) = 0 \quad \text{for } a > b, \quad (3.4)$$

$$h_a f(\eta^{c\gamma}) \quad \text{for } a = b,$$

where h_a is the degree of homogeneity of f in the ath twistor. Generally[8,11] for a set of n complex, n-dimensional "vectors,"

$\eta^{a\alpha}$, these functions give representations of the Lie algebra of GL(n,C), which can be regarded as finite-dimensional if the h_a's are positive integers and infinite-dimensional otherwise. For twistors, due to their dimension and their conjugation properties, the representation will be[11] of the Lie algebra of U(2,2). Using Eq. (3.4), we shall reduce the number of variables from 16 to ten, as Eq. (3.4) gives six additional constraints.

IV. TRANSLATION OF SPINOR VARIABLES INTO TWISTORS

We shall use basis spinstors,[8,12] $e_{A\alpha}$, $e_{A'}{}^{\alpha}$, $e^{A\alpha}$, $e^{A'}{}_{\alpha}$, to write twistors in terms of their component spinors and vice-versa. They are defined in component form by

$$(e^{A\alpha}) = (e_{A\alpha}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \tag{4.1a}$$

$$(e_{A'}{}^{\alpha}) = (e^{A'}{}_{\alpha}) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.1b}$$

We define the four twistors of the quadrad in terms of spinors by

$$\begin{aligned} W^{\alpha} &= \pi_A e^{A\alpha} + \sigma^{A'} e_{A'}{}^{\alpha}, \\ X^{\alpha} &= \rho_A e^{A\alpha} + \nu^{A'} e_{A'}{}^{\alpha}, \\ Y^{\alpha} &= \lambda_A e^{A\alpha} + \mu^{A'} e_{A'}{}^{\alpha}, \\ Z^{\alpha} &= \xi_A e^{A\alpha} + \eta^{A'} e_{A'}{}^{\alpha}, \end{aligned} \tag{4.2a}$$

and the four twistors of the dual quadrad by

$$\begin{aligned} \tilde{Z}^{\alpha} &= \tilde{\xi}_{A'} e^{A'}{}_{\alpha} + \tilde{\eta}^{A} e_{A\alpha}, \\ \tilde{Y}_{\alpha} &= \tilde{\lambda}_{A'} e^{A'}{}_{\alpha} + \tilde{\mu}^{A} e_{A\alpha}, \\ \tilde{X}_{\alpha} &= \tilde{\rho}_{A'} e^{A'}{}_{\alpha} + \tilde{\nu}^{A} e_{A\alpha}, \\ \tilde{W}_{\alpha} &= \tilde{\pi}_{A'} e^{A'}{}_{\alpha} + \tilde{\sigma}^{A} e_{A\alpha}, \end{aligned} \tag{4.2b}$$

where ξ_A and $\tilde{\xi}_{A'}$ are the twistors referred to earlier. [The inter-relationships of the component spinors become clearer in the subsequent calculations from Eq. (4.11) on. For a geometric interpretation of the quadrad, see Ref. 8.]

The position $x^{AA'}$ is defined by the point of interaction of Y and Z [being the position vector of the point relative to the origin which was chosen for the expansions given by Eq. (4.2)]. Thus we have

$$\eta^{A'} = -ix^{AA'}\xi_A, \tag{4.3a}$$

$$\mu^{A'} = -ix^{AA'}\lambda_A. \tag{4.3b}$$

Multiplying Eqs. (4.3a) and (4.3b) by λ_B and ξ_B, respectively, and subtracting

$$i(\lambda_A\xi_B - \lambda_B\xi_A)x^{AA'} = \xi_B\mu^{A'} - \lambda_B\eta^{A'}. \tag{4.4}$$

By using the spinstors for projection, Eq. (4.2a) gives

$$\lambda_A = Y^{\alpha}e_{A\alpha}, \quad \mu^{A'} = Y^{\alpha}e^{A'}{}_{\alpha}, \tag{4.5a}$$

$$\xi_A = Z^{\alpha}e_{A\alpha}, \quad \eta^{A'} = Z^{\alpha}e^{A'}{}_{\alpha}. \tag{4.5b}$$

Using Eq. (4.5) and Eq. (4.4) gives

$$\lambda_A\xi_B - \lambda_B\xi_A = Y^{\alpha}(e_{A\alpha}e_{B\beta} - e_{B\alpha}e_{A\beta})Z^{\beta}. \tag{4.6}$$

Using the fact that any expression skew in two spinor indices can be written as the same expression with the indices contracted times ϵ_{AB}, and the definition[2,8] of the infinity twistor

$$e_{A\alpha}\epsilon^{AB}e_{B\beta} = I_{\alpha\beta}, \tag{4.7}$$

we see that Eq. (4.6) becomes

$$\lambda_A\xi_B - \lambda_B\xi_A = Y^{\alpha}I_{\alpha\beta}Z^{\beta}\epsilon_{AB}. \tag{4.8}$$

Putting Eqs (4.5) and (4.8) into Eqs. (4.4). we obtain

$$x^{AA'} = 2i\frac{e^{A\gamma}e^{A'}{}_{\alpha}Y^{[\alpha}Z^{\beta]}I_{\beta\gamma}}{Y^{\rho}I_{\rho\pi}Z^{\pi}}. \tag{4.9a}$$

Similarly, from the dual quadrad we could obtain

$$\tilde{x}^{AA'} = -2i\frac{e^{A'}{}_{\gamma}e^{A\alpha}\tilde{Y}_{[\alpha}\tilde{Z}_{\beta]}I^{\alpha\beta}}{\tilde{Y}_{\rho}I^{\rho\pi}\tilde{Z}_{\pi}}. \tag{4.9b}$$

Then, as $Z^{\alpha}\tilde{Z}_{\alpha} = 0$ [by the definition of the dual quadrad, see Eqs. (3.2)]

$$(x^{AA'} - \tilde{x}^{AA'})\xi_A\tilde{\xi}_{A'} = 0. \tag{4.10}$$

Now we note that $x^{AA'}$ is inversely proportional to $Y^{\rho}I_{\rho\pi}Z^{\pi}$. Thus $Y^{\rho}I_{\rho\pi}Z^{\pi}$ is a convenient choice of the conformal function (which should compactify the Minkowski space). Similarly, as $\tilde{x}^{AA'}$ is inversely proportional to $\tilde{Y}_{\rho}I^{\rho\pi}\tilde{Z}_{\pi}$, $\tilde{Y}_{\rho}I^{\rho\pi}\tilde{Z}_{\pi}$ would serve as the other conformal function. Thus we define

$$\chi = Y^{\alpha}I_{\alpha\beta}Z^{\beta}, \quad \tilde{\chi} = \tilde{Y}_{\alpha}I^{\alpha\beta}\tilde{Z}_{\beta}. \tag{4.11}$$

Clearly, for a real quadrad, $\tilde{\chi} = \bar{\chi}$ and $\tilde{x}^{AA'} = \bar{x}^{AA'} = x^{AA'}$. Thus Eq. (4.10) will be trivially satisfied.

We can evaluate the twistors of the dual quadrad in terms of the quadrad by

$$\begin{aligned} \tilde{W}_{\nu} &= \epsilon_{\mu\nu\rho\pi}W^{\mu}X^{\nu}Y^{\rho}/D, \\ \tilde{X}_{\rho} &= \epsilon_{\mu\nu\rho\pi}W^{\mu}X^{\nu}Z^{\pi}/D, \\ \tilde{Y}_{\nu} &= \epsilon_{\mu\nu\rho\pi}W^{\mu}Y^{\rho}Z^{\pi}/D, \\ \tilde{Z}_{\mu} &= \epsilon_{\mu\nu\rho\pi}X^{\nu}Y^{\rho}Z^{\pi}/D, \end{aligned} \tag{4.12}$$

where

$$D = \epsilon_{\mu\nu\rho\pi}W^{\mu}X^{\nu}Y^{\rho}Z^{\pi}. \tag{4.13}$$

Now, since

$$I_{\alpha\beta}\epsilon^{\alpha\beta\gamma\delta} = 2I^{\gamma\delta} \tag{4.14}$$

and

$$\tilde{\chi} = \tilde{Y}_{\rho}I^{\rho\pi}\tilde{Z}_{\pi}, \tag{4.15}$$

putting Eq. (4.12) into Eq. (4.15), we see that

$$\tilde{\chi} = \tfrac{1}{2}\epsilon_{\mu\nu\rho\pi}Y^{\mu}Z^{\nu}I^{\rho\pi}/D, \tag{4.16}$$

which, by Eq. (4.14), gives

$$\chi = \tilde{\chi}/D. \tag{4.17}$$

Thus we see that

$$\chi/\tilde{\chi} = D. \tag{4.18}$$

It is comparatively easy to write $\partial/\partial\eta^{a\alpha}$ (which we shall denote by $\partial_{a\alpha}$) in terms of derivatives with respect to the spinor variables. However, because of the constraints on the fields [contained in Eq. (3.4)] it is not so easy to invert this so as to be able to write the spinor field equations in terms of the

twistor quadrad. Let us first translate the derivatives $\partial_{a\alpha}$ into spinor derivatives.

V. THE TRANSLATION OF THE DERIVATIVES

To translate the derivatives, we shall need a relation between $\partial_{a\alpha}$ and $\tilde{\partial}^{a\alpha}$ ($\equiv \partial/\partial\tilde{\eta}_{a\alpha}$). For this we prove the following theorem.

Theorem: The derivates with respect to the quadrad and the dual quadrad are related by

$$\partial_{a\alpha} = -\tilde{\eta}_{a\beta}\tilde{\eta}_{b\alpha}\tilde{\partial}^{b\beta}, \tag{5.1a}$$

$$\tilde{\partial}^{a\alpha} = -\eta^{a\beta}\eta^{b\alpha}\partial_{b\beta}. \tag{5.1b}$$

Proof: (i) Differentiating Eq. (3.2a) with respect to the quadrad, we see that

$$\delta^c_b\delta^\alpha_\gamma\tilde{\eta}_{a\alpha} + \eta^{b\alpha}\partial_{c\gamma}\tilde{\eta}_{a\alpha} = 0. \tag{5.2}$$

Multiplying both sides by $\tilde{\eta}_{b\beta}$, using Eq. (3.2b), and simplifying, we obtain

$$\partial_{c\gamma}\tilde{\eta}_{a\beta} = -\tilde{\eta}_{c\beta}\tilde{\eta}_{a\gamma}. \tag{5.3}$$

Now the chain rule can be written as

$$\partial_{a\alpha} = (\partial_{a\alpha}\tilde{\eta}_{b\beta})\tilde{\partial}^{b\beta}. \tag{5.4}$$

Putting Eq. (5.3) into Eq. (5.4) gives Eq. (5.1a).

(ii) Similarly, differentiating Eq. (3.2a) with respect to the dual quadrad and following the same procedure, we derive Eq. (5.1b). Hence the theorem.

For simplicity, we shall first take the fields to be conformally invariant, so that $\partial f/\partial\chi$ and $\partial f/\partial\tilde{\chi}$ give zero. Thus we do not need to consider the conformal function derivatives in our translations as yet. We shall include them later by direct extension of the earlier results. We can then write the derivatives with respect to the quadrad in terms of the derivatives with respect to ξ_A, $x^{AA'}$, and $\tilde{\xi}_{A'}$ as

$$\begin{bmatrix} \frac{\partial}{\partial W^\alpha} \\ \frac{\partial}{\partial X^\alpha} \\ \frac{\partial}{\partial Y^\alpha} \\ \frac{\partial}{\partial Z^\alpha} \end{bmatrix} = \begin{bmatrix} \frac{\partial\xi_A}{\partial W^\alpha} & \frac{\partial x^{AA'}}{\partial W^\alpha} & \frac{\partial\tilde{\xi}_{A'}}{\partial W^\alpha} \\ \frac{\partial\xi_A}{\partial X^\alpha} & \frac{\partial x^{AA'}}{\partial X^\alpha} & \frac{\partial\tilde{\xi}_{A'}}{\partial X^\alpha} \\ \frac{\partial\xi_A}{\partial Y^\alpha} & \frac{\partial x^{AA'}}{\partial Y^\alpha} & \frac{\partial\tilde{\xi}_{A'}}{\partial Y^\alpha} \\ \frac{\partial\xi_A}{\partial Z^\alpha} & \frac{\partial x^{AA'}}{\partial Z^\alpha} & \frac{\partial\tilde{\xi}_{A'}}{\partial Z^\alpha} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial\xi_A} \\ \frac{\partial}{\partial x^{AA'}} \\ \frac{\partial}{\partial\tilde{\xi}_{A'}} \end{bmatrix}. \tag{5.5}$$

We must calculate this transformation matrix using the expressions for the spinor variables in terms of twistors derived in the previous section.

Now, from Eq. (4.2a) we can evaluate the first column of the matrix, as ξ_A can be written in terms of twistors as given in Eq. (4.5b):

$$\frac{\partial\xi_A}{\partial W^\alpha} = \frac{\partial\xi_A}{\partial X^\alpha} = \frac{\partial\xi_A}{\partial Y^\alpha} = 0, \tag{5.6a}$$

$$\frac{\partial\xi_A}{\partial Z^\alpha} = 0. \tag{5.6b}$$

Similarly, to obtain the last column of the matrix, we use Eq. (4.2b) to enable us to write

$$\tilde{\xi}_{A'} = \tilde{Z}_\alpha e_{A'}{}^\alpha. \tag{5.7}$$

Thus we have

$$\frac{\partial\tilde{\xi}_{A'}}{\partial\tilde{Z}_\alpha} = e_{A'}{}^\alpha, \tag{5.8a}$$

$$\frac{\partial\tilde{\xi}_{A'}}{\partial\tilde{Y}_\alpha} = \frac{\partial\tilde{\xi}_{A'}}{\partial\tilde{X}_\alpha} = \frac{\partial\tilde{\xi}_{A'}}{\partial\tilde{W}_\alpha} = 0. \tag{5.8b}$$

Using Eq. (5.1) with Eq. (5.8), we obtain

$$\frac{\partial\tilde{\xi}_{A'}}{\partial W^\alpha} = -\tilde{Z}_\alpha\tilde{Z}_\beta e_{A'}{}^\beta,$$

$$\frac{\partial\tilde{\xi}_{A'}}{\partial X^\alpha} = -\tilde{Z}_\alpha\tilde{Y}_\beta e_{A'}{}^\beta,$$

$$\frac{\partial\tilde{\xi}_{A'}}{\partial Y^\alpha} = -\tilde{Z}_\alpha\tilde{X}_\beta e_{A'}{}^\beta,$$

$$\frac{\partial\tilde{\xi}_{A'}}{\partial Z^\alpha} = -\tilde{Z}_\alpha\tilde{W}_\beta e_{A'}{}^\beta.$$

To obtain the middle column of the matrix, we use Eq. (4.9a), which gives

$$\frac{\partial x^{AA'}}{\partial W^\alpha} = \frac{\partial x^{AA'}}{\partial X^\alpha} = 0, \tag{5.10a}$$

$$\frac{\partial x^{AA'}}{\partial Y^\alpha} = 4i\frac{e^{A\gamma}e^{A'}{}_\mu I_{\beta\gamma}Z^{[\beta}\delta^{[\mu]}_\alpha Y^{\rho]}I_{\rho\pi}Z^\pi}{(Y^\nu I_{\nu\sigma}Z^\sigma)^2}, \tag{5.10b}$$

$$\frac{\partial x^{AA'}}{\partial Z^\alpha} = 4i\frac{e^{A\gamma}e^{A'}{}_\beta Y^\rho I_{\rho\pi}Z^{[\pi}\delta^{[\mu]}_\alpha Y^{\beta]}I_{\mu\gamma}}{(Y^\nu I_{\nu\sigma}Z^\sigma)^2}. \tag{5.10c}$$

Equation (5.5) may be symbolically written as

$$(\partial_{a\alpha}) = (A_{ij})(\partial_{A,A'}), \tag{5.11}$$

where $(\partial_{a\alpha})$ represents the matrix of the derivatives with respect to the quadrad, (A_{ij}) the transformation matrix, and $(\partial_{A,A'})$ the matrix of derivatives with respect to the spinor variables. We want to "invert" the matrix (A_{ij}) subject to the weight constraints contained in Eq. (3.4) and subject to the fact that $\partial f/\partial\chi = \partial f/\partial\tilde{\chi} = 0$, to obtain

$$(\partial_{A,A'}) = (B_{ij})(\partial_{a\alpha}). \tag{5.12}$$

We now want to determine the components of (B_{ij}).

We can "invert" the transformation matrix[8] (A_{ij}) by explicitly writing out its components a_{ij} and requiring that the components of its "inverse" matrix satisfy the simultaneous equations

$$\sum_{j=0}^{16} b_{ij}a_{jk} = \delta_{i,k}, \tag{5.13}$$

where $\delta_{i,k}$ is the Kronecker delta. Writing Eq. (5.12) as

$$\begin{bmatrix} \frac{\partial}{\partial\xi_A} \\ \frac{\partial}{\partial x^{AA'}} \\ \frac{\partial}{\partial\tilde{\xi}_{A'}} \end{bmatrix} = \begin{bmatrix} p^{A\alpha} & q^{A\alpha} & r^{A\alpha} & s^{A\alpha} \\ p_{AA'}{}^\alpha & q_{AA'}{}^\alpha & r_{AA'}{}^\alpha & s_{AA'}{}^\alpha \\ p^{A'\alpha} & q^{A'\alpha} & r^{A'\alpha} & s^{A'\alpha} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial W^\alpha} \\ \frac{\partial}{\partial X^\alpha} \\ \frac{\partial}{\partial Y^\alpha} \\ \frac{\partial}{\partial Z^\alpha} \end{bmatrix},$$

it is found that Eq. (5.13) can be satisfied by[8]

$$p^{A\alpha} = q^{A\alpha} = r^{A\alpha} = 0, \tag{5.15a}$$

$$s^{A\alpha} = 2e^{A\beta}\frac{Y^{[\gamma}Z^{\alpha]}I_{\beta\gamma}}{Y^{\rho}Z^{\pi}I_{\rho\pi}}, \tag{5.15b}$$

$$\begin{aligned} p_{AA'}{}^{\alpha} &= ie_{A\beta}e_{A'}{}^{\alpha}W^{\beta}, \\ q_{AA'}{}^{\alpha} &= ie_{A\beta}e_{A'}{}^{\alpha}X^{\beta}, \\ r_{AA'}{}^{\alpha} &= ie_{A\beta}e_{A'}{}^{\alpha}Y^{\beta}, \\ s_{AA'}{}^{\alpha} &= ie_{A\beta}e_{A'}{}^{\alpha}Z^{\beta}, \end{aligned} \tag{5.16}$$

$$p^{A'\alpha} = -\frac{e^{A'}{}_{\gamma}\tilde{Y}_{\beta}W^{\alpha}I^{\beta\gamma}}{\tilde{Y}_{\rho}\tilde{Z}_{\pi}I^{\rho\pi}}, \tag{5.17a}$$

$$q^{A'\alpha} = \frac{e^{A'}{}_{\gamma}\tilde{Z}_{\beta}W^{\alpha}I^{\beta\gamma}}{\tilde{Y}_{\rho}\tilde{Z}_{\pi}I^{\rho\pi}}, \tag{5.17b}$$

$$r^{A'\alpha} = s^{A'\alpha} = 0. \tag{5.17c}$$

Thus we get the translation of the spinor derivatives into the twistor derivatives by putting Eqs. (5.15)–(5.17) into Eq. (5.14) to obtain

$$\frac{\partial}{\partial\xi_A} = 2e^{A\beta}\frac{Y^{[\alpha}Z^{\gamma]}I_{\beta\gamma}}{Y^{[\rho}Z^{\pi]}I_{\rho\pi}}\frac{\partial}{\partial Z^{\alpha}}, \tag{5.18}$$

$$\frac{\partial}{\partial x^{AA'}} = ie_{A\alpha}e_{A'}{}^{\beta}L^{\alpha}{}_{\beta}, \tag{5.19}$$

$$\frac{\partial}{\partial\tilde{\xi}_{A'}} = 2e^{A'}{}_{\beta}\frac{\tilde{Y}_{[\alpha}\tilde{Z}_{\gamma]}I^{\beta\gamma}}{\tilde{Y}_{[\rho}\tilde{Z}_{\pi]}I^{\rho\pi}}\frac{\partial}{\partial\tilde{Z}_{\alpha}}, \tag{5.20}$$

where $L^{\alpha}{}_{\beta}$ is the usual canonical generator for the Lie algebra of U(2,2) (the group for twistors) being given by[8,11]

$$L^{\alpha}{}_{\beta} = \eta^{a\alpha}\partial_{a\beta}. \tag{5.21}$$

The dependence on the conformal functions χ and $\tilde{\chi}$ can be included by considering the additional terms in the derivatives of χ and $\tilde{\chi}$ with respect to the quadrad twistors

$$\left(\frac{\partial}{\partial W^{\alpha}}\right)_{+} = -\tilde{Z}_{\alpha}\tilde{\chi}\frac{\partial}{\partial\tilde{\chi}}, \tag{5.22a}$$

$$\left(\frac{\partial}{\partial X^{\alpha}}\right)_{+} = -\tilde{Y}_{\alpha}\tilde{\chi}\frac{\partial}{\partial\tilde{\chi}}, \tag{5.22b}$$

$$\left(\frac{\partial}{\partial Y^{\alpha}}\right)_{+} = -2\tilde{X}_{\beta}\tilde{Z}_{[\alpha}\tilde{Y}_{\gamma]}I^{\beta\gamma}\frac{\partial}{\partial\tilde{\chi}} + I_{\alpha\beta}Z^{\beta}\frac{\partial}{\partial\chi}, \tag{5.22c}$$

$$\left(\frac{\partial}{\partial Z^{\alpha}}\right)_{+} = -2\tilde{W}_{\beta}\tilde{Z}_{[\alpha}\tilde{Y}_{\gamma]}I^{\beta\gamma}\frac{\partial}{\partial\tilde{\chi}} + Y^{\beta}I_{\beta\alpha}\frac{\partial}{\partial\chi} \tag{5.22d}$$

(where the + subscript signifies that the conformal functions are also varying). Multiplying the complete expressions for $\partial/\partial X^{\alpha}$ by X^{α} and $\partial/\partial Y^{\alpha}$ by Y^{α}, we obtain $\partial/\partial\chi$ and $\partial/\partial\tilde{\chi}$. The complete expressions for the derivatives with respect to the spinor variables are

$$\frac{\partial}{\partial\tilde{\chi}} = -\frac{1}{\tilde{Y}_{\rho}\tilde{Z}_{\pi}I^{\rho\pi}}X^{\alpha}\frac{\partial}{\partial X^{\alpha}}, \tag{5.23}$$

$$\frac{\partial}{\partial\tilde{\xi}_{A'}} = 2\frac{\tilde{Y}_{[\alpha}\tilde{Z}_{\gamma]}e^{A'}{}_{\beta}I^{\alpha\beta}}{\tilde{Y}_{\rho}\tilde{Z}_{\pi}I^{\rho\pi}}\left(\frac{\partial}{\partial\tilde{Z}_{\gamma}} - W^{\gamma}\tilde{Y}_{\delta}\frac{\partial}{\partial\tilde{Y}_{\delta}}\right), \tag{5.24}$$

$$\frac{\partial}{\partial x^{AA'}} = ie_{A\alpha}e_{A'}{}^{\beta}L^{\alpha}{}_{\beta}, \tag{5.25}$$

$$\frac{\partial}{\partial\xi_A} = 2\frac{Y^{[\alpha}Z^{\gamma]}e^{A\beta}I_{\alpha\beta}}{Y^{\rho}Z^{\pi}I_{\rho\pi}}\left(\frac{\partial}{\partial Z^{\gamma}} - \tilde{W}_{\gamma}Y^{\delta}\frac{\partial}{\partial Y^{\delta}}\right), \tag{5.26}$$

$$\frac{\partial}{\partial\chi} = \frac{1}{Y^{\rho}Z^{\pi}I_{\rho\pi}}Y^{\alpha}\frac{\partial}{\partial Y^{\alpha}}. \tag{5.27}$$

Equations (5.23)–(5.25) may be verified by operating on the spinor variables with them.

It is interesting to note that

$$w = -\tilde{\xi}_{A'}\frac{\partial}{\partial\tilde{\xi}_{A'}} - \tilde{\chi}\frac{\partial}{\partial\tilde{\chi}}, \tag{5.28a}$$

$$x = -\tilde{\chi}\frac{\partial}{\partial\tilde{\chi}}, \tag{5.29a}$$

$$y = \chi\frac{\partial}{\partial\chi}, \tag{5.30a}$$

$$z = \xi_A\frac{\partial}{\partial\xi_A} + \chi\frac{\partial}{\partial\chi}, \tag{5.31a}$$

where w, x, y, and z are the degrees of homogeneity of the field in W^{α}, X^{α}, Y^{α}, and Z^{α}, respectively. Considering the dual quadrad

$$\tilde{z} = \tilde{\xi}_{A'}\frac{\partial}{\partial\tilde{\xi}_{A'}} + \tilde{\chi}\frac{\partial}{\partial\tilde{\chi}} = -w, \tag{5.28b}$$

$$\tilde{y} = \tilde{\chi}\frac{\partial}{\partial\tilde{\chi}} = -x, \tag{5.29b}$$

$$\tilde{x} = -\chi\frac{\partial}{\partial\chi} = -y, \tag{5.30b}$$

$$\tilde{w} = -\xi_A\frac{\partial}{\partial\xi_A} - \chi\frac{\partial}{\partial\chi} = -z. \tag{5.31b}$$

VI. THE FIELD EQUATIONS IN TWISTORS

We have succeeded in writing the partial derivatives with respect to the spinor variables in terms of twistors of the quadrad. However, we have not taken into account the fact that the spinors, ξ_A and $\tilde{\xi}_{A'}$, and the conformal functions, χ and $\tilde{\chi}$, are themselves functions of position, i.e., they are fields themselves. This may be understood as follows. We take a fibre, defined by (ξ_A, $\tilde{\xi}_{A'}$, χ, and $\tilde{\chi}$) on the compactified complexified Minkowski space with points specified by $x^{AA'}$ to obtain in the fibre bundle with variables ($x^{AA'}$, ξ_A, $\tilde{\xi}_{A'}$, χ, $\tilde{\chi}$). We take a section of this fibre bundle with variables $y^{AA'}$, such that

$$x^{AA'}(y^{BB'}) = y^{AA'}. \tag{6.1}$$

So, whereas the position $x^{AA'}$ and the fields (ξ_A, $\tilde{\xi}_{A'}$, χ, and $\tilde{\chi}$) are independent variables, they are all functions of position, $y^{AA'}$, which happens to be equal to $x^{AA'}$. We must deal with derivatives with respect to $y^{AA'}$ and not $x^{AA'}$ only. By the chain rule we obtain that

$$\frac{\partial}{\partial y^{AA'}} = \frac{\partial}{\partial x^{AA'}} + \frac{\partial\xi_B}{\partial y^{AA'}}\frac{\partial}{\partial\xi_B} + \frac{\partial\tilde{\xi}_{B'}}{\partial y^{AA'}}\frac{\partial}{\partial\tilde{\xi}_{B'}}$$

$$+\frac{\partial\chi}{\partial y^{AA'}}\frac{\partial}{\partial\chi}+\frac{\partial\tilde{\chi}}{\partial y^{AA'}}\frac{\partial}{\partial\tilde{\chi}}. \tag{6.2}$$

Even now we have not taken into account the covariant derivative for the fields, but only the partial derivative. To obtain the covariant derivative with respect to $y^{AA'}$, ${}^{y}\nabla_{AA'}$, we note that[8]

$$ {}^{y}\nabla_{AA'}\chi=(\Upsilon_{AA'}+i\Theta_{AA'})\chi=\frac{\partial\chi}{\partial y^{AA'}}, $$

$$ {}^{y}\nabla_{AA'}\tilde{\chi}=(\tilde{\Upsilon}_{AA'}-i\tilde{\Theta}_{AA'})\tilde{\chi}=\frac{\partial\tilde{\chi}}{\partial y^{AA'}}, \tag{6.3}$$

$$ {}^{y}\nabla_{AA'}\xi_B=\frac{\partial\xi_B}{\partial y^{AA'}}-(\Upsilon_{A'B}+i\Theta_{A'B})\xi_A, \tag{6.4}$$

$$ {}^{y}\nabla_{AA'}\tilde{\xi}_{B'}=\frac{\partial\tilde{\xi}_{B'}}{\partial y^{AA'}}-(\tilde{\Upsilon}_{AB'}-i\tilde{\Theta}_{AB'})\tilde{\xi}_{A'}, \tag{6.5}$$

where $(\Upsilon_{AA'}+i\Theta_{AA'})$ is the spinor equivalent to the connection symbols in some sense,[4] Υ_{AA} and $\Theta_{AA'}$ being "real" in the sense that $\Upsilon_{AA'}=\bar{\Upsilon}_{AA'}$, $\Theta_{AA'}=\bar{\Theta}_{AA'}$. Thus we can write the covariant derivative with respect to $y^{AA'}$ in terms of the covariant derivative with respect to $x^{AA'}$ by using Eqs. (6.1)–(6.5)

$$\begin{aligned}{}^{y}\nabla_{AA'}=\nabla_{AA'}&+\left[\frac{\partial\xi_B}{\partial y^{AA'}}-(\Upsilon_{BA'}+i\Theta_{BA'})\xi_A\right]\frac{\partial}{\partial\xi_B}\\&+(\Upsilon_{AA'}+\Theta_{AA'})\chi\frac{\partial}{\partial\chi}+\left[\frac{\partial\tilde{\xi}_{B'}}{\partial y^{AA'}}-(\tilde{\Upsilon}_{AB'}\right.\\&\left.-i\tilde{\Theta}_{AB'})\tilde{\xi}_{A'}\right]\frac{\partial}{\partial\tilde{\xi}_{B'}}+(\tilde{\Upsilon}_{AA'}-i\tilde{\Theta}_{AA'})\tilde{\chi}\frac{\partial}{\partial\tilde{\chi}}.\end{aligned} \tag{6.6}$$

We shall first write the operators of the field equations with the connection symbols being zero and then include them.

The field appearing in some of the field equations can be written as scalars, as given below. Thus the Dirac fields are

$$f=\alpha^A\xi_A,\quad g=\beta^{A'}\tilde{\xi}_{A'}, \tag{6.7}$$

the Maxwell fields are

$$J=J^{BB'}\xi_B\tilde{\xi}_{B'},$$

$$\Phi=\varphi^{AB}\xi_A\xi_B,\quad\tilde{\Phi}=\tilde{\varphi}^{A'B'}\tilde{\xi}_{A'}\tilde{\xi}_{B'}, \tag{6.8}$$

and the Weyl gravitational fields are

$$\begin{aligned}\Psi&=\psi^{ABCD}\xi_A\xi_B\xi_C\xi_D,\\\tilde{\Psi}&=\bar{\psi}^{A'B'C'D'}\tilde{\xi}_{A'}\tilde{\xi}_{B'}\tilde{\xi}_{C'}\tilde{\xi}_{D'}.\end{aligned} \tag{6.9}$$

The scalar operators appearing in the field equations are

$$\Delta=\xi^A\xi^B\Box_{AB}, \tag{6.10a}$$

$$\underline{\Delta}=\tilde{\xi}^{A'}\tilde{\xi}^{B'}\Box_{A'B'}, \tag{6.10b}$$

$$D=\xi^A\tilde{\xi}^{A'}\nabla_{AA'}=\underline{D}, \tag{6.10c}$$

$$\eth=\xi^A\nabla_{AA'}\frac{\partial}{\partial\xi_A}, \tag{6.10d}$$

$$\underline{\eth}=\xi^A\nabla_{AA'}\frac{\partial}{\partial\tilde{\xi}_{A'}}, \tag{6.10e}$$

$$\Diamond=\nabla_{AA'}\frac{\partial}{\partial\xi_A}\frac{\partial}{\partial\tilde{\xi}_{A'}}=\underline{\Diamond}, \tag{6.10f}$$

where

$$\Box_{AB}=2\nabla_{(A}{}^{A'}\nabla_{B)A'}. \tag{6.11}$$

The wave operator and the scalar field remain unchanged.

Thus the Klein–Gordon equation remains

$$(\Box+m^2)\varphi=0, \tag{6.12a}$$

the Dirac equations become

$$\eth f=(m/\sqrt{2})g,\quad\eth g=-(m/\sqrt{2})f, \tag{6.12b}$$

the Maxwell equations become

$$\eth\Phi=2\pi J,\quad\underline{\eth}\Phi=0,$$

$$\eth\tilde{\Phi}=0,\quad\underline{\eth}\tilde{\Phi}=0, \tag{6.12c}$$

and the linearized gravitational field equations become

$$\eth\Psi=0=\underline{\eth}\Psi,\quad\eth\tilde{\Psi}=0=\underline{\eth}\tilde{\Psi}. \tag{6.12d}$$

We first give the operators in the special case when

$${}^{y}\nabla_{AA'}\xi_B=0, \tag{6.13a}$$

$${}^{y}\nabla_{AA'}\tilde{\xi}_{B'}=0, \tag{6.13b}$$

$${}^{y}\nabla_{AA'}\chi=0={}^{y}\nabla_{AA'}\tilde{\chi}. \tag{6.13c}$$

Then Eq. (6.10) can be expressed in twistors as

$$\begin{aligned}\Delta&=-Z^\alpha Z^\beta I_{\alpha\gamma}I_{\beta\delta}I^{\mu\nu}L^\gamma{}_\mu L^\delta{}_\nu,\\\underline{\Delta}&=-\tilde{Z}_\alpha\tilde{Z}_\beta I^{\alpha\mu}I^{\beta\nu}I_{\gamma\delta}L^\gamma{}_\mu L^\delta{}_\nu,\end{aligned} \tag{6.14}$$

as $L^\alpha_\mu=-\tilde{L}^\alpha_\mu$, which follows from Eq. (5.1),

$$D=iZ^\alpha\tilde{Z}_\beta I_{\alpha\gamma}I^{\beta\delta}L^\gamma{}_\delta=\underline{D}, \tag{6.15}$$

$$\eth=2i\tilde{Z}_\alpha\frac{I^{\alpha\beta}Y^{[\gamma}Z^{\delta]}I_{\delta\lambda}}{Y^\rho Z^\pi I_{\rho\pi}}\left(\frac{\partial}{\partial Z^\delta}-y\,\widetilde{W}_\delta\right)L^\lambda{}_\beta, \tag{6.16a}$$

$$\underline{\eth}=2iZ^\alpha I_{\alpha\lambda}\frac{\tilde{Y}_{[\gamma}\tilde{Z}_{\delta]}I^{\gamma\beta}}{\tilde{Y}_{[\mu}\tilde{Z}_{\nu]}I^{\mu\nu}}\left(\frac{\partial}{\partial\tilde{Z}_\delta}+x\,W^\delta\right)L^\lambda{}_\beta, \tag{6.16b}$$

where x and y are the degrees of homogeneity of the fields in the quadrad twistors X^α and Y^α

$$\begin{aligned}\Diamond&=4i\frac{Y^{[\alpha}Z^{\beta]}I_{\alpha\gamma}}{Y^\delta Z^\epsilon I_{\delta\epsilon}}\frac{\tilde{Y}_{[\mu}\tilde{Z}_{\nu]}I^{\mu\rho}}{\tilde{Y}_\sigma\tilde{Z}_\pi I^{\sigma\pi}}\\&\quad\times\left(\frac{\partial}{\partial Z^\beta}-y\widetilde{W}_\beta\right)\left(\frac{\partial}{\partial\tilde{Z}_\nu}+xW^\nu\right)L^\gamma{}_\rho\\&=\underline{\Diamond}.\end{aligned} \tag{6.17}$$

The wave operator is given by

$$\Box=-I^{\alpha\beta}I_{\gamma\delta}L^\gamma{}_\alpha L^\delta{}_\beta. \tag{6.18}$$

To write the complete operators when Eq. (6.13) is not satisfied, we must include the spinor equivalents of the connection symbols. For this purpose we define the *connection twistors* as

$$\Upsilon^\alpha{}_\beta=e^{A\alpha}e^{A'}{}_\beta\Upsilon_{AA'},\quad\Theta^\alpha{}_\beta=e^{A\alpha}e^{A'}{}_\beta\Theta_{AA'}. \tag{6.19}$$

Then, denoting the complete operators by putting on a "hat," $\hat{}$,

$$\hat{D} = D + Z^{\alpha}\tilde{Z}_{\beta}[y(\Upsilon^{\alpha}{}_{\beta} + i\Theta^{\alpha}{}_{\beta}) + x(\tilde{\Upsilon}_{\beta}{}^{\alpha} - i\tilde{\Theta}_{\beta}{}^{\alpha})]$$
$$= \underline{\hat{D}}, \tag{6.20}$$

$$\hat{\eth} = \eth - \bar{Z}_{\delta} I^{\delta\rho} \frac{Y^{[\alpha} Z^{\beta]} I_{\alpha\gamma}}{Y^{\mu} Z^{\nu} I_{\mu\nu}} \left(\frac{\partial}{\partial Z^{\beta}} - y\,\widetilde{W}_{\beta}\right)$$
$$\times [y(\Upsilon^{\alpha}{}_{\beta} + i\Theta^{\alpha}{}_{\beta}) - x(\tilde{\Upsilon}_{\beta}{}^{\alpha} - i\tilde{\Theta}_{\beta}{}^{\alpha})], \tag{6.21}$$

$$\hat{\underline{\eth}} = \underline{\eth} - Z^{\delta} I_{\delta\gamma} \frac{\tilde{Y}_{[\alpha}\tilde{Z}_{\beta]} I^{\alpha\rho}}{\tilde{Y}_{\mu}\tilde{Z}_{\nu} I^{\mu\nu}} \left(\frac{\partial}{\partial \tilde{Z}_{\beta}} + xW^{\beta}\right)$$
$$\times [y(\Upsilon^{\gamma}{}_{\rho} + i\Theta^{\gamma}{}_{\rho}) - x(\tilde{\Upsilon}_{\rho}{}^{\gamma} + i\tilde{\Theta}_{\rho}{}^{\gamma})], \tag{6.22}$$

$$\hat{\Diamond} = \Diamond + 4\,\frac{Y^{[\alpha} Z^{\gamma]} I_{\alpha\beta} \tilde{Y}_{[\mu}\tilde{Z}_{\rho]} I^{\mu\nu}}{Y^{\delta} Z^{\epsilon} I_{\delta\epsilon} \tilde{Y}_{\pi}\tilde{Z}_{\sigma} I^{\pi\sigma}} \left(\frac{\partial}{\partial Z^{\gamma}} - y\widetilde{W}_{\gamma}\right)$$
$$\times \left(\frac{\partial}{\partial \tilde{Z}_{\rho}} + xW^{\rho}\right)[y(\Upsilon^{\beta}{}_{\nu} + i\Theta^{\beta}{}_{\nu})$$
$$- x(\tilde{\Upsilon}_{\nu}{}^{\beta} - i\tilde{\Theta}_{\nu}{}^{\beta})]$$

$$= \underline{\hat{\Diamond}}, \tag{6.23}$$

$$\hat{\Delta} = \Delta, \tag{6.24}$$

$$\hat{\underline{\Delta}} = \underline{\Delta}, \tag{6.25}$$

$$\hat{\Box} = \Box = \underline{\Box}. \tag{6.26}$$

Equation (6.12) together with Eqs. (6.14)–(6.18) and (6.20)–(6.26) comprise the field equations in twistors.

VII. CONCLUSION

We have seen that we can write the massive field equations, with source present, in terms of twistors, using a twistor quadrad (set of four twistors) by imposing six constraints on the fields. This was done by writing the "twistor derivatives" in terms of the spinor derivatives by a transformation matrix. The matrix being nonsquare did not have a unique "inverse." The choice of the "inverse" was made by requiring consistency, internally, and with the six constraints. We thus have a "unique" inverse, in the sense that any further permissible change of the inverse matrix cannot change the transformation formulas.

It still remains to solve the twistor field equations. Some general points can be mentioned about the solutions. It is easily verified that the zero rest mass, source-free, equations are solved by fields of the type

$$f(W^{\alpha}, X^{\beta} Y^{\gamma} Z^{\delta}) = A_{\alpha\cdots\beta\gamma\cdots\delta\,\mu\cdots\nu\,\rho\cdots\pi}$$
$$\times W^{\alpha}\cdots W^{\beta} X^{\gamma}\cdots X^{\delta} Y^{\mu}\cdots Y^{\nu} Z^{\rho}\cdots Z^{\pi}, \tag{7.1}$$

where

$$A_{\alpha\cdots\beta\gamma\cdots\delta\,\mu\cdots\nu\,\rho\cdots\pi} = A_{(\alpha\cdots\beta)(\gamma\cdots\delta)(\mu\cdots\nu)(\rho\cdots\pi)}, \tag{7.2a}$$

$$A_{(\alpha\cdots\beta\gamma)\cdots\delta\,\mu\cdots\nu\,\rho\cdots\pi} = A_{(\alpha\cdots\beta\,|\gamma\cdots\delta|\,\mu)\cdots\nu\,\rho\cdots\pi}$$
$$= A_{(\alpha\cdots\beta\,|\gamma\cdots\delta\,\mu\cdots\nu|\,\rho)\cdots\pi}$$
$$= A_{\alpha\cdots\beta(\gamma\cdots\delta\,\mu)\cdots\nu\,\rho\cdots\pi}$$
$$= A_{\alpha\cdots\beta(\gamma\cdots\delta|\,\mu\cdots\nu|\,\rho)\cdots\pi}$$
$$= A_{\alpha\cdots\beta\gamma\cdots\delta(\mu\cdots\nu\,\rho)\cdots\pi} = 0, \tag{7.2b}$$

with

$$|\{\alpha,\ldots,\beta\}| \leqslant |\{\gamma,\ldots,\delta\}| \leqslant |\{\mu,\ldots,\nu\}| \leqslant |\{\rho,\ldots,\pi\}|. \tag{7.2c}$$

It would be interesting to find the most general conditions for solutions of the zero rest-mass field equations.

A field of the above type, which is only a function of Z, will be a spin-$z/2$ field. If it also depends on Y, it will be a spin $(z-y)/2$, conformal weight y field. It would be interesting to determine what all the degrees of homogeneity refer to. It is hoped that this will lead to a better understanding of Penrose graphs.[9]

Of immense importance is the solution of the massive field equations with source terms present. In particular, we could write the complete gravitational field equations (i.e., Einstein equations) in twistors.[12] This is of prime importance in the program for the "quantization of relativity" using twistors.

By regarding the fields as functions giving us a representation of the Lie algebra of U(2,2), the Casimir operators can be evaluated in terms of the degrees of homogeneity of the fields in the twistor variables.[8,11] This would have some bearing on the significance of the degrees of homogeneity. Conversely, if the significance is already understood, it would be interesting to consider the "conserved quantities" contained in the Casimir operators due to the requirement of invariance under U(2,2). It would also be of interest[14] to consider the cohomology classes of the functions representing the fields.

It should be mentioned that by bringing in the infinity twistor we are breaking conformal invariance, as infinity is "being located" in some sense. However, Poincaré invariance is being maintained. That would be broken by the introduction of the origin twistor. Thus our fields cannot have the origin twistor in them. This places a severe restriction on the fields. It is to be expected that massive fields will include the infinity twistor, as masses break conformal invariance, but the massless fields need not do so.

ACKNOWLEDGMENTS

I would like to express my gratitude to professor R. Penrose, the "father of twistors," for suggesting the problem and for his encouragement and advice. Thanks are also due to Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste.

[1] *Quantum Gravity*, edited by C.J. Isham, R. Penrose, and D. W. Sciama (Clarendon, Oxford, 1975).
[2] R. Penrose, J. Math. Phys. **8**, 345 (1967).
[3] R. Penrose, Int. J. Theoret. Phys. **1**, 61 (1968).
[4] R. Penrose, *An Analysis of the Structure of Space-Time* (Adams Prize Essay, 1967).
[5] R. Penrose, "Structure of Space-Time," in *Battelle Rencontres*, edited by B. S. DeWitt and J. A. Wheeler (Benjamin, New York, 1968).
[6] R. Penrose, J. Math. Phys. **10**, 38 (1969).
[7] R. Penrose and M. A. H. MacCallum, Phys. Rep. C **6**, 243 (1973).
[8] A. Qadir, "Twistor fields," Ph.D. Thesis, London (1971).
[9] A. Qadir, Phys. Rep. C **39**, 133 (1978).
[10] G. Sparling, in Ref. 1, pp. 408–500.
[11] A. Qadir, Int. J. Theoret. Phys. **16**, 25 (1976).
[12] K. Dighton, Int. J. Theoret. Phys. **11**, 31 (1974).
[13] F. A. E. Pirani, in *Brandeis Summer School* 1964, *Vol.* 1: *Lectures on General Relativity*, edited by F. A. E. Pirani, H. Bondi, and A. Trautmann (Prentice-Hall, Englewood Cliffs, N. J., 1965).
[14] R. Penrose, private communication.

Inhomogeneous rotating universes with closed timelike geodesics of matter

I.Damião Soares
Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil

(Received 6 November 1978; accepted for publication 9 November 1979)

We present a new class of inhomogeneous stationary cosmological solutions of Einstein–Maxwell equations, with rotating dust and electromagnetic fields. For a subclass of these models, the topology of the space-time manifold is $S^3\times R$, and the timelike geodesic lines of dust are closed; electromagnetic fields are necessary only to avoid matter singularities, while in the completely homogeneous limit they are essential for consistency of field equations; these electromagnetic fields have the structure of magnetic monopole fields, since the total magnetic flux across the unit 2-sphere is different from zero—that is, magnetic monopoles must be present as a source of the fields.

1. INTRODUCTION

Cosmological solutions of Einstein field equations, in which the matter content of the model rotates relative to the local compass of inertia, have long been known in the literature.[1-5] Although in most of them the space–time manifold contains closed timelike lines, these curves are not however geodesics, and the world lines of the matter content of the models—whose energy–momentum tensor enters the right-hand side of Einstein equations—are *open* timelike geodesic lines.

We present here a class of cosmological models, a subclass of which presents not only timelike lines of the above type, but also the timelike lines of matter are *closed* geodesics. The matter content of the models is a rotating inhomogeneous dust, since in fact rotation seems to be a necessary feature for the existence of acausal structures in fluid-filled cosmological solutions. Electromagnetic fields are eventually present. In Sec. II the line element and its local properties are discussed, as well as local calculations relative to the velocity field of matter, electromagnetic fields, and Einstein equations. In Sec. III our interest is a subclass of the models, whose space–time manifold is shown to have the topology $S^3\times R$, and the timelike geodesics of dust are closed curves over S^3. In other words, the timelike Killing vector $\partial/\partial t$ is the velocity field of matter and the orbit of a point under the group of time translations is a closed geodesic. In Sec. IV the motion of matter and geodesics of the model are analyzed.

2. THE GEOMETRY OF THE MODELS

Taking (t,r,Θ,Φ) as local coordinates on the manifold, we choose a tetrad field $e_\alpha^{(A)}(x)$ such that the line element can be reduced to[6]

$$ds^2=\eta_{AB}\Theta^A\Theta^B, \tag{2.1}$$

where the $\Theta^A=e_\alpha^{(A)}\,dx^\alpha$ are given by

$$\begin{aligned}\Theta^0&=A_0[dt+4m^2(\Theta)\,d\Phi],\\ \Theta^1&=dr,\\ \Theta^2&=B(r)K(\Theta)\,d\Theta,\\ \Theta^3&=B(r)K(\Theta)\sin\Theta\,d\Phi.\end{aligned} \tag{2.2}$$

Here A_0 is a constant, and the functions $m(\Theta)$ and $K(\Theta)$ satisfy

$$\frac{4m}{K^2\sin\Theta}\frac{dm}{d\Theta}=\lambda_1, \tag{2.3}$$

$$\frac{d^2K}{d\Theta^2}-\frac{1}{K}\left(\frac{dK}{d\Theta}\right)^2+\cot\Theta\frac{dK}{d\Theta}-K=\lambda K^3, \tag{2.4}$$

where λ_1 and λ are constant, λ being proportional to the curvature of the 2-spheres $d\Sigma^2=K^2(\Theta)\,(d\Theta^2+\sin^2\Theta\,d\Phi^2)$.

For (2.2), Ricci rotation coefficients γ_{ABC} and Ricci tensor R_{AB} are calculated to have nonnull components

$$\begin{aligned}&\gamma_{023}=-\gamma_{032}=\gamma_{320}=A_0\lambda_1/B^2,\\ &\gamma_{212}=\gamma_{313}=-B'/B,\\ &\gamma_{323}=-\frac{1}{BK^2}\left(\frac{dK}{d\Theta}+K\cot\Theta\right),\end{aligned} \tag{2.5}$$

and

$$R_{00}=2A_0^2\lambda_1^2/B^4,\quad R_{11}=-2B''/B, \tag{2.6}$$

$$R_{22}=R_{33}=-\frac{B''}{B}-\left(\frac{B'}{B}\right)^2-\frac{\lambda}{B^2}+\frac{2A_0^2\lambda_1^2}{B^4}.$$

In the local Lorentz frame determined by (2.2) we assume that an observer comoving with the fluid has 4-velocity

$$u^A=\delta_0^A, \tag{2.7}$$

and we denote by ρ the density of mass–energy of the pressureless fluid, as measured locally by the observer (2.7). The energy–momentum tensor for the fluid is then expressed

$$T_{AB}=\rho u_A u_B. \tag{2.8}$$

We remark that the velocity field of dust (2.7) determines a congruence of timelike geodesics, since $\gamma_{0A0}=0, A=1,2,3$.

For (2.1) Maxwell equations are expressed

$$e^\rho{}_{|(P)}F_{QR\,]|\rho}+2\,F_{A\,[R}\gamma^A{}_{PQ\,]}=0, \tag{2.9a}$$

$$e^\rho{}_{(P)}F^P{}_{D\,|\rho}-F_{AD}\gamma^{AP}{}_P-F_{PB}\gamma_D{}^{PB}=j_D, \tag{2.9b}$$

and we consider two cases:

(i) The components of the electromagnetic tensor F_{AB} are restricted to

$$F_{01}=-F_{10}=E(r),\quad F_{23}=-F_{32}=H(r),$$

 0022-2488/80/030521-05$1.00

and $j^A = 0$, which correspond to a parallel electric field and magnetic field relative to (2.7), locally along the direction determined by Θ^1. For (2.2) and (2.5) we have the independent solutions of (2.9),

$$E = (\Sigma/B^2)\cos 2\lambda_1 \tilde{r}, \quad H = (\Sigma/B^2)\sin 2\lambda_1 \tilde{r},$$

where Σ is a constant and the variable $\tilde{r}$ is defined by $d\tilde{r} = A_0 B^{-2}\, dr$.

(ii) We restrict F_{AB} to

$$F_{23} = -F_{32} = H(r),$$

corresponding to a pure magnetic field relative to the comoving observer (2.7), and the electric current

$$j_D = Q u_D,$$

where Q is the proper electric charge density relative to (2.7). We have the solutions of (2.9).

$$H = \Sigma/B^2, \quad Q = -2\lambda_1 A_0 \Sigma/B^4,$$

where Σ is a constant. These electromagnetic fields have the structure of magnetic monopole fields, in the sense that Eq. (2.9a) is not everywhere valid. In fact for $K(\Theta) = 1$, the total magnetic flux crossing the unit 2-sphere is easily calculated to be different from zero: In case (ii) the magnetic flux is equal to the total electromagnetic flux, with value $4\pi\Sigma$; and in case (i) the total magnetic flux is given by $4\pi\Sigma \sin\tilde{r}$. This implies that *magnetic monopoles* must be present as a source of the fields.[7]

For both cases (i) and (ii), the electromagnetic energy–momentum tensor $T_{AB} = -F_{AC}F_B{}^C + \frac{1}{4}\eta_{AB}F_{CD}F^{CD}$ is expressed by

$$T_{AB} = (\Sigma^2/2B^4)\,\mathrm{diag}(+1, -1, +1, +1). \tag{2.10}$$

The total energy–momentum tensor for the models is then

$$T_{AB}(\mathrm{total}) = \rho\delta^0_A\delta^0_B + (\Sigma^2/2B^4)\,\mathrm{diag}(+1, -1, +1, +1), \tag{2.11}$$

and the Einstein field equations

$$R_{AB} - \tfrac{1}{2}\eta_{AB}R + \Lambda\eta_{AB} = kT_{AB}(\mathrm{total}),$$

resulting in

$$2k\rho = R_{00} + 3R_{11} + 2\Lambda + k\Sigma^2/B^4, \tag{2.12}$$

$$R_{00} - R_{11} - 2\Lambda - k\Sigma^2/B^4 = 0, \tag{2.13}$$

$$\frac{B''}{B} - \left(\frac{B'}{B}\right)^2 + \frac{2A_0^2\lambda_1^2 - k\Sigma^2}{B^4} - \frac{\lambda}{B^2} = 0. \tag{2.14}$$

Denoting

$$\alpha_0 = 2A_0^2\lambda_1^2 - k\Sigma^2 \tag{2.15}$$

a first integral of (2.14) is given by

$$(B'/B)^2 = \mathscr{C} + (\alpha_0/2)B^{-4} - \lambda B^{-2}, \tag{2.16}$$

where $\mathscr{C}$ is an integration constant, provided B'/B is not identically zero. From (2.16), (2.14), and (2.6), we have from (2.12) and (2.13)

$$2k\rho = -4\mathscr{C} + (8A_0^2\lambda_1^2 - 2k\Sigma^2)B^{-4}, \quad \Lambda = \mathscr{C}. \tag{2.17}$$

From (2.17) we can see that the models have a matter singularity where $B = 0$, unless $4A_0^2\lambda_1^2 = k\Sigma^2$. A singularity in the electromagnetic fields always occurs for $B = 0$.

The following solutions of Eq. (2.16) for B are possible (here r_0 is an integration constant such that $r_0 \leqslant r$):

(1) $\mathscr{C} = 0$,

$$B^2 = -(1/\lambda)[[r - r^0]^2 - \alpha_0/2];$$

(2) $\mathscr{C} < 0$,

$$B^2 = \frac{-\lambda + (\lambda^2 + 2\alpha_0|\mathscr{C}|)^{1/2}\sin 2(|\mathscr{C}|)^{1/2}(r - r_0)}{2|\mathscr{C}|}.$$

In cases (1) and (2) for $\lambda = -1$, the singularity at $B = 0$ can be avoided only if $\alpha_0 < 0$, which implies from (2.15) that an electromagneic field must be present.

(3) $\mathscr{C} > 0$, $\lambda = 1 = 2\alpha_0\mathscr{C}$

$$B^2 = \left\{1 - \exp\left[-2\sqrt{\mathscr{C}}(r - r_0)\right]\right\}/2\mathscr{C}.$$

Other possible solutions are not considered because they imply negative values of ρ. For the completely homogenous case ($B = \mathrm{const}$) we obtain directly from (2.14)

$$B^2 = \alpha_0/\lambda,$$

$$2k\rho = 4A_0^2\lambda_1^2/\alpha_0^2,$$

$$2\Lambda = 1/\alpha_0,$$

and for $\lambda = -1$, electromagnetic fields are necessary in order to have $B^2 > 0$. In general, from the above expressions of B^2 we see that the structure of the models is very sensitive to the values of the parameters $A_0\lambda_1$ and Σ.

The isometry groups acting transitively on the $r = \mathrm{const}$ sections are of Bianchi types II, VIII, or IX,[8] respectively for $\lambda = 0, +1, -1$. For $\lambda = 0$, Eq. (2.4) has the general solution

$$K(\Theta) = \frac{1}{\sin\Theta}\left(\frac{1 - \cos\Theta}{1 + \cos\Theta}\right)^{q_0/2}, \tag{2.18}$$

where q_0 is an arbitrary integration constant. For $q_0 \neq 0$ the function $K(\Theta)$ defines a projective mapping of the sphere on the whole plane, and for $q_0 = 0$ the projected sphere has the structure of a cylinder, although in both cases the local curvature is $\lambda = 0$. Up to additive constants, we have from (2.3) and (2.18)

$$m^2(\Theta) = \frac{\lambda_1}{4q_0}\left(\frac{1 - \cos\Theta}{1 + \cos\Theta}\right)^{q_0} \quad \text{for } q_0 \neq 0$$

or

$$m^2(\Theta) = (\lambda_1/2)\ln(\tan\Theta/2) \quad \text{for } q_0 = 0.$$

Defining new coordinates

$$\bar{\Theta} = \frac{1}{q_0}\left(\frac{1 - \cos\Theta}{1 + \cos\Theta}\right)^{q_0/2}$$

with $q_0 \neq 0$ and $0 \leqslant \bar{\Theta} \leqslant +\infty$, and $\bar{\Theta} = -\ln(\tan\Theta/2)$ with $q_0 = 0$ and $-\infty \leqslant \bar{\Theta} \leqslant +\infty$, the manifolds can be extended to manifolds with line elements

$$ds^2 = A_0^2(dt - 2\lambda_1\bar{\Theta}\,d\Phi)^2 - dr^2 - B^2(r)(d\bar{\Theta}^2 + d\Phi^2) \quad \text{for } q_0 = 0,$$

$$ds^2 = A_0^2(dt + \lambda_1 q_0\bar{\Theta}^2\,d\Phi)^2 - dr^2 - B^2(r)(d\bar{\Theta}^2 + \bar{\Theta}^2 q_0^2\,d\Phi^2) \quad \text{for } q_0 \neq 0.$$

For $\lambda = 1$ we take for (2.4) the solution $K(\Theta) = \tan\Theta$, and (2.3) can be integrated to $m^2 = (\lambda_1/2)(1/\cos\Theta + \cos\Theta)$. Defining the coordinate $\bar{\Theta} = -\ln(\cos\Theta)$, $0 \leqslant \bar{\Theta} \leqslant +\infty$, the manifold can be extended to one with metric

$$ds^2 = A_0^2\,[dt + (\lambda_1/2)\cosh\bar{\Theta}\,d\Phi\,]^2 - dr^2 - B^2(r)(d\bar{\Theta}^2 + \sinh^2\bar{\Theta}\,d\Phi^2).$$

The most interesting subclass of the models is the one defined by $\lambda = -1$. For this we take $K(\Theta) = 1$ and we have from (2.3), $m^2 = -\lambda_1/2\cos\Theta$. The line element in this case assumes the form

$$ds^2 = A_0^2(\omega^0)^2 - B^2(r)[(\omega^1)^2 + (\omega^2)^2] - (\omega^3)^2, \tag{2.19}$$

where

$$\omega^0 = dt - 2\lambda_1\cos\Theta\,d\Phi,$$
$$(\omega^1)^2 + (\omega^2)^2 = d\Theta^2 + \sin^2\Theta\,d\Phi^2, \tag{2.20}$$
$$\omega^3 = dr.$$

The structure of the space–time with metric (2.19), (2.20) is discussed in the next section, and we show that the manifold has the topology $S^3\times R$, namely the timelike sections $r = \text{const}$ have the structure of the 3-sphere S^3. The t-lines (which are the trajectories of the matter content of the model) are closed geodesics on S^3—determined by the vector field $\partial/\partial t$ of the invariant basis of S^3—with the corresponding t-coordinate cyclic.

3. THE STRUCTURE OF THE $\lambda = -1$ MANIFOLDS

The methods used in this section are borrowed from Oszvath and Schücking,[5] and are presented here concisely for completeness. We construct a Lorentzian manifold which has the structure $S^3\times R$ and is isometric to the solutions of Einstein equations (2.19) and (2.20).

Let E_4 be the four-dimensional Euclidean space with Cartesian coordinates $q^\mu = (q^0,q^1,q^2,q^3)$, and the unit vectors along the Cartesian axes denotes by $\mathbf{e}_\mu$. With a multiplication law defined by

$$\mathbf{e}_\mu\mathbf{e}_0 = \mathbf{e}_0\mathbf{e}_\mu = \mathbf{e}_\mu, \quad \mathbf{e}_i\mathbf{e}_i = -\mathbf{e}_0, \quad \mathbf{e}_i\mathbf{e}_j = \sum_k \epsilon_{ijk}\mathbf{e}_k, \tag{3.1}$$

$i,j,k = 1,2,3$, E_4 is the quaternation algebra, and the vectors

$$\mathbf{q} = q^\mu\mathbf{e}_\mu = q^0\mathbf{e}_0 + \sum_i q^i\mathbf{e}_i \tag{3.2}$$

are called quaternions. Distributivity and associativity are assumed. From (3.1) we have that $\mathbf{e}_0$ is the identity of the algebra, with quaternions of the type $\mathbf{q} = q_0\mathbf{e}_0$ isomorphic to the field of real numbers, and we hence can identify $q_0\mathbf{e}_0 \sim q_0$.

For a quaternion (3.2), we define its conjugate quaternion by

$$\mathbf{q}^* = q_0\mathbf{e}_0 - \sum_{i=1}^{3} q^i\mathbf{e}_i. \tag{3.3}$$

We then have $\mathbf{q}\mathbf{q}^* = \mathbf{q}^*\mathbf{q} = \Sigma_{\mu=0}^3(q^\mu)^2 \geqslant 0$, the equality occurring for $\mathbf{q} = 0$. Since $N(q) = \Sigma(q^\mu)^2$ always exists, every quaternion $\mathbf{q}\neq 0$ has an inverse $\mathbf{q}^{-1} = (N(q))^{-1}\mathbf{q}^*$ such that $\mathbf{q}^{-1}\mathbf{q} = \mathbf{q}\mathbf{q}^{-1} = 1$.

The equation of the 3-sphere S^3 can be expressed

$$\mathbf{q}\mathbf{q}^* = 1. \tag{3.4}$$

We can now identify S^3 with the group of motions of S^3, with S^3 acting on itself by left multiplication. In fact for any given quaternion $\mathbf{v}$ of S^3 ($\mathbf{v}^*\mathbf{v} = 1$), a left motion of S^3 on itself is expressed by

$$\mathbf{q}' = \mathbf{v}\mathbf{q}, \tag{3.5}$$

and we have [using that $(\mathbf{ab})^* = \mathbf{b}^*\mathbf{a}^*$]

$$\mathbf{q}'\mathbf{q}'^* = 1 = \mathbf{q}\mathbf{q}^*.$$

S^3 is a simply transitive group since for each $\mathbf{a}\neq 0$ of S^3, there exists only one left translation $\mathbf{r}$ from $\mathbf{a}$ to a given $\mathbf{a}'$, namely $\mathbf{r} = \mathbf{a}'\mathbf{a}^*$.

S^3 acting on itself by left multiplications (3.5) is a Lie group, and the independent left invariant vector fields over S^3 generate a representation of the Lie algebra of the three-parameter rotation group. To obtain these fields we proceed as follows. Representing the unit quaternions $(\mathbf{e}_0,\mathbf{e}_i)$ by the matrices

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\quad i\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},$$
$$i\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\quad i\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},$$

we replace every quaternion $\mathbf{q}$ of S^3 by the matrix

$$A = \begin{pmatrix} q^0 + iq^3 & q^2 + iq^1\\ -q^2 + iq^1 & q^0 - iq^3\end{pmatrix},\quad \det A = 1.$$

The quaternion multiplication goes over to matrix multiplication. Introducing the Euler coordinates (x^0,x^1,x^2) on S^3 by

$$q^0 = \cos\frac{x^1}{2}\cos\frac{x^0 + x^2}{2},$$
$$q^1 = \sin\frac{x^1}{2}\cos\frac{x^2 - x^0}{2},$$
$$q^2 = -\sin\frac{x^1}{2}\sin\frac{x^2 - x^0}{2},$$
$$q^3 = \cos\frac{x^1}{2}\sin\frac{x^0 + x^2}{2},$$

where $0\leqslant x^1\leqslant\pi$, $0\leqslant x^0,x^2\leqslant 2\pi$, the left invariant 1-forms ω^μ are obtained by calculating[9]

$$\omega = A^{-1}\,dA = \omega^\mu\mathbf{e}_\mu, \tag{3.6}$$

and we have

$$\omega = \begin{pmatrix} \frac{i}{2}(\cos x^1\,dx^2 + dx^0) & \frac{i}{2}e^{-ix^0}\,dx^1 - \frac{1}{2}\sin\frac{x^1}{2}e^{-ix^0}\,dx^2\\ \frac{i}{2}e^{ix^0}\,dx^1 \sin x^1 e^{-ix^0}\,dx^2 & \frac{i}{2}(\cos x^1 dx^2 + dx^0)\end{pmatrix},$$

and the last equality (3.6) yields the following independent left-invariant 1-forms,

$$\omega^1 = \cos x^0\,dx^1 + \sin x^1\sin x^0\,dx^2,$$
$$\omega^2 = -\sin x^0\,dx^1 + \sin x^1\cos x^0\,dx^2, \tag{3.7}$$
$$\omega^3 = \cos x^1\,dx^2 + dx^0,$$

with corresponding left-invariant vector fields

$$X_1 = \cos x^0\frac{\partial}{\partial x^1} + \frac{\sin x^0}{\sin x^1}\frac{\partial}{\partial x^2} - \cot x^1\sin x^0\frac{\partial}{\partial x^0},$$
$$X_2 = -\sin x^0\frac{\partial}{\partial x^1} + \frac{\cos x^0}{\sin x^1}\frac{\partial}{\partial x^2} - \cot x^1\cos x^0\frac{\partial}{\partial x^0}, \tag{3.8}$$
$$X_3 = \frac{\partial}{\partial x^0}.$$

We construct a Lorentzian manifold $S^3\times R$ by adding to the invariant basis (3.8) the vector field $\partial/\partial r$ (defining the

congruence of r-lines) with dual 1-form dr, and defining the metric by

$$ds^2 = a_0^2(\omega^3)^2 - B^2(r)\,[(\omega^1)^2 + (\omega^2)^2] - dr^2 \qquad (3.9)$$

so that the $r = \text{const}$ sections have the structure of S^3. To recover expressions (2.19) and (2.20) we redefine $x^0 = -t/2\lambda_1$, $a_0 = 2\lambda_1 A_0$, $x^1 = \Theta$, $x^2 = \Phi$. Actually (3.9) can be understood to correspond to a deformation of the algebra of the vector fields (3.8) [cf. for instance (4.2)] and a consequent deformation of the S^3 manifold (3.4) into the ellipsoid

$$\left(\frac{1}{2\lambda_1 A_0}\right)^2 (q_1)^2 + \frac{1}{B^2(r)}\{(q_2)^2 + (q_3)^3\} + (q_0)^2 = 1. \qquad (3.10)$$

The manifold (3.10) is now parametrized with the coordinate r, with $B(r)$ given from Einstein equations in Sec. II. We can use the expression (3.10) for a description of the global (topological) structure of the $r = \text{const}$ sections, as depending on $B(r)$ and $A_0\lambda_1$. In both cases (3.10) or (3.4) we refer to the $r = \text{const}$ manifolds as S^3.

Definition (3.9) is equivalent to the choice [cf. (3.8)]

$$\begin{aligned} &g(X_3,X_3) = a_0^2, \\ &g(X_1,X_1) = g(X_2,X_2) = -B^2(r), \\ &g\left(\frac{\partial}{\partial r},\frac{\partial}{\partial r}\right) = -1, \end{aligned} \qquad (3.11)$$

other products zero.

Since the right-invariant vector fields $\{\xi^i\}$ over S^3 satisfy, by construction $\pounds_{\xi^i} X_j = 0$, $i,j = 1,2,3$, the metric (3.9) shall obviously have ξ^i as Killing vectors, which in the present coordinates are expressed

$$\xi^1 = \frac{\partial}{\partial x^2},$$

$$\xi^2 = \cos x^2 \frac{\partial}{\partial x^1} - \cot x^1 \sin x^2 \frac{\partial}{\partial x^2} + \frac{\sin x^2}{\sin x^1}\frac{\partial}{\partial x^0},$$

$$\xi^3 = -\sin x^2 \frac{\partial}{\partial x^1} - \cot x^1 \cos x^2 \frac{\partial}{\partial x^2} + \frac{\cos x^2}{\sin x^1}\frac{\partial}{\partial x^0}.$$

From (3.9) or (3.11) we see that an additional Killing vector is

$$\xi^4 = \frac{\partial}{\partial x^0}.$$

The metric (3.9) is a solution of the Einstein field equations for dust and/or eventually electromagnetic fields, for $B(r)$ given in Sec. II with $\lambda = -1$. The matter velocity field $Y = \delta_0^A e^\alpha_{(A)}(\partial/\partial x^\alpha)$ [cf. (2.7)] can be expressed, from (3.8) and (3.11),

$$Y = \frac{1}{A_0}\frac{\partial}{\partial t}. \qquad (3.12)$$

(3.12) is a geodesic field, and belongs to the left invariant basis (3.8) of S^3, so that the congruence of curves along which the fluid propagates are *closed* timelike geodesics, of proper length or period $4\pi\lambda_1 A_0$. Y generates a one-parameter subgroup of isometries, the time translations; t is a canonical cyclic coordinate on S^3 and the orbit of an arbitrary point of the manifold under Y is homeomorphic to a circle.

4. THE MOTION OF MATTER

We consider the vector fields

$$\begin{aligned} Z_0 &= \frac{1}{A_0}\frac{\partial}{\partial t}, \\ Z_1 &= \frac{1}{B(r)}\left(\sin\frac{t}{2\lambda_1}\frac{\partial}{\partial\Theta} + \frac{\cos t/2\lambda_1}{\sin\Theta}\frac{\partial}{\partial\Phi}\right. \\ &\quad \left. + 2\lambda_1 \cot\Theta\cos\frac{t}{2\lambda_1}\frac{\partial}{\partial t}\right), \\ Z_2 &= \frac{1}{B(r)}\left(\cos\frac{t}{2\lambda_1}\frac{\partial}{\partial\Theta} - \frac{\sin t/2\lambda_1}{\sin\Theta}\frac{\partial}{\partial\Phi}\right. \\ &\quad \left. - 2\lambda_1\cot\Theta\sin\frac{t}{2\lambda_1}\frac{\partial}{\partial t}\right), \\ Z_3 &= \frac{\partial}{\partial r}, \end{aligned} \qquad (4.1)$$

which define local Lorentz frames in the space–time (3.7) and (3.9). (Z_0,Z_1,Z_2) provides a basis for the algebra of $S^3(r)$, namely

$$\begin{aligned} [Z_0,Z_1] &= \frac{1}{2\lambda_1 A_0} Z_2, \\ [Z_0,Z_2] &= \frac{-1}{2\lambda_1 A_0} Z_1, \\ [Z_1,Z_2] &= \frac{2\lambda_1 A_0}{B^2(r)} Z_0. \end{aligned} \qquad (4.2)$$

The congruence of dust world lines is defined by the geodesic field Z_0, Z_0 being orthogonal to (Z_1,Z_2,Z_3), and the coordinates (r,Θ,Φ) being constant along Z_0.

Let Z be a vector orthogonal to Z_0 and connecting two neighboring fluid particles, one of them located at the origin of the Lorentz frame determined by (4.1). We can write

$$Z = \sum_{i=1}^{3} \mu^i Z_i \qquad (4.3)$$

and the motion of Z can be obtained from

$$\underset{Z_0}{\pounds} Z = 0. \qquad (4.4)$$

Denoting $\dot{\mu}^i = Z_0 \mu^i$ and using (4.2), Eq. (4.4) yields

$$\dot{\mu}^1 = \frac{1}{2A_0\lambda_1}\mu^2, \quad \dot{\mu}^2 = \frac{-1}{2A_0\lambda_1}\mu^1, \quad \dot{\mu}^3 = 0,$$

which corresponds to a rotation of period $4\pi\lambda_1 A_0$, with respect to the inertial frame Z_A, equal to the period of a closed timelike geodesics of matter. The orbit of the extremity of Z is the circle in the Z_1Z_2 plane,

$$(\mu^1)^2 + (\mu^2)^2 = \text{const}, \quad \mu^3 = \text{const}.$$

The motion of the frame (4.1) along Z_0 can be calculated by

$$\dot{Z}_A \equiv \underset{Z_0}{\nabla} Z_A = \gamma_A{}^B{}_0 Z_B.$$

Using (2.5) we have

$$\begin{aligned} &\dot{Z}_1 = (A_0\lambda_1/B^2)Z_2, \quad \dot{Z}_2 = -(A_0\lambda_1/B^2)Z_1, \\ &\dot{Z}_3 = 0 = \dot{Z}_0, \end{aligned} \qquad (4.5)$$

that is, the plane 1-2 of the frame Z_A rotates with respect to the local compass of inertia, with a circular frequency

$\omega = A_0\lambda_1/B^2$ (the axes of the local compass of inertia being determined, for instance, by gyroscopes). The rotation of matter relative to the local compass of inertia is then given by

$$\Omega = \frac{1}{2A_0\lambda_1} + \frac{A_0\lambda_1}{B^2}.$$

The rigid rotation of matter is a direct expression of the fact that the shear of the matter velocity field (2.7), $\sigma^2 = \gamma_{0(AB)}\gamma_0^{(AB)}$, is zero, according to (2.5).

Closed timelike lines, of the type present in Gödel and other rotating models, are also present here. The curves $t = \text{const}$, $\Theta = \text{const}$, $r = \text{const}$, with

$$ds^2 = [4A_0^2\lambda_1^2\cos^2\Theta_0 - B^2(r)\sin^2\Theta_0]\,d\Phi^2,$$

are timelike for values of (r,Θ_0) such that

$$4A_0^2\lambda_1^2\cos^2\Theta_0 - B^2(r)\sin^2\Theta_0 > 0.$$

5. NONSPACELIKE GEODESICS FOR THE STATIONARY CASE. FINAL REMARKS

For the metric

$$ds^2 = A_0^2(dt - 2\lambda_1\cos\Theta\,d\Phi)^2 - dr^2 - B_0^2(d\Theta^2 + \sin^2\Theta\,d\Phi^2),$$

where $B_0^2 = k\Sigma^2 - 2A_0^2\lambda_1^2$ (cf. Sec. 2), the nonspacelike geodesics are described by the tangent vector field $\dot{x}^\alpha = dx^\alpha/ds$, with

$$\dot{t} = k_0 + \lambda_1\cos\Theta\,\frac{h_0 + k_0A_0^2\lambda_1\cos\Theta}{(3A_0^2\lambda_1^2\cos^2\Theta - B_0^2\sin^2\Theta)},$$

$$\dot{r} = r_0,$$

$$\dot{\Theta}^2 = \frac{(h_0 + A_0^2k_0\lambda_1\cos\Theta)^2 + A_0^2k_0^2 - \binom{1}{0} - r_0}{B_0^2},$$

$$\dot{\Phi} = \frac{h_0 + k_0A_0^2\lambda_1\cos\Theta}{(3A_0^2\lambda_1^2\cos^2\Theta - B_0^2\sin^2\Theta)},$$

where h_0, k_0, and r_0 are arbitrary parameters. The values (1,0) are taken for timelike or null geodesics, respectively. Two trivial cases can be given by choosing $\Theta = \Theta_0 = \text{const}$.

(i) Choose r_0, h_0, k_0 such that

$A_0^2k_0^2 = 1$ (timelike geodesics),

$r_0 = 0$,

$h_0 + k_0A_0^2\lambda_1\cos\Theta_0 = 0$.

This gives $\dot{x}^\alpha = k_0\delta_0^\alpha$ or

$$V = \frac{1}{A_0}\frac{\partial}{\partial t},$$

which is the tangent field of the closed trajectories of matter. In this case the parameter s is defined on S^1, since t is a cyclic coordinate.

(ii) Choose (k_0,h_0) such that

$$h_0 + k_0A_0^2\lambda_1\cos\Theta_0 = 0 \quad \text{and} \quad A_0^2k_0^2 = \binom{1 + r_0^2}{r_0^2},$$

respectively, for timelike and null geodesics. We have $t = k_0s$, $r = r_0s$ and $dt/dr = k_0/r_0$, where s is defined on S^1.

We finally remark that from our models, the solution $\lambda = 1$, $\Sigma = 0$, $B^2 = 2A_0^2\lambda_1^2$ is locally isometric to Gödel's universe, and the solution $\lambda = -1$, $B^2 = k\Sigma^2 - 2A_0^2\lambda_1^2$ is a homogeneous model recently discussed[10] where the dust and electromagnetic field content is described as a fluid with an anisotropic pressure.

[1] W.J. van Stockum, Proc. Roy. Soc. Edinburgh **57**, 135 (1937).
[2] K. Gödel, Rev. Mod. Phys. **21**, 447 (1949).
[3] J.P. Wright, J. Math. Phys. **6**, 103 (1965).
[4] S.C. Maitra, J. Math. Phys. **7**, 1025 (1966).
[5] I. Ozsvath and E.L. Schucking, Ann. Phys. **55**, 166 (1969).

[6] Notations and conventions: Capital Latin indices are tetrad indices and run from 0 to 3; they are raised and lowered with the Minkowski metric $\eta^{AB}, \eta_{AB} = \text{diag}(+1,-1,-1,-1)$. Greek indices run from 0 to 3 and are raised and lowered with $g^{\alpha\beta}, g_{\alpha\beta}$. Square and round brackets denote antisymmetrization and symmetrization, respectively. A bar denotes partial derivative. $\pounds_X$ is Lie derivative with respect to the vector field X. Ricci rotation coefficients are defined $\gamma_{ABC} = -e^{\alpha}_{(A)\|\beta}\,e^{\beta}_{(C)}$. The Riemann tensor is defined by $V_{\alpha\|\beta\|\gamma\|} = R^{\lambda}{}_{\alpha\beta\gamma}V_\lambda$, and the Ricci tensor by $R_{\alpha\beta} = R^{\lambda}{}_{\alpha\lambda\beta}$, that implies Einstein constant $k > 0$.

[7] Expressing Eq. (2.9a) as $dF = 0$, where $F = F_{\alpha\beta}dx^\alpha \Lambda\, dx^\beta$, we have from Stokes theorem[9] that

$$\int_V dF = \int_S F.$$

For instance, for $K = 1$ and case (ii), the flux across the unit sphere S^2 is $\frac{1}{2}\int_S F = 4\pi\Sigma \neq 0$ for $\Sigma \neq 0$. We acknowledge the referee for calling our attention to this point.

[8] A.H. Taub, Ann. Math. **53**, 472 (1951).
[9] H. Flanders, *Differential Forms* (Academic, New York, 1963).
[10] M. Novello and M.J. Rebouças, Phys. Rev. D **19**, 2850 (1979).

On the role of space–time topology in quantum phenomena: Superselection of charge and emergence of nontrivial vacua[a)]

Abhay Ashtekar
Département de Physique, Université de Clermont-Fd, 63170 Aubière, France
Max Planck Institut, Föhringer Ring 6, 8 Mühchen 40, Germany,
and Enrico Fermi Institute, Chicago, Illinois 60637

Amitabha Sen
Enrico Fermi Institute, Chicago, Illinois 60637

(Received 7 December 1978; accepted for publication 11 October 1979)

Schwarzschild–Kruskal space–time admits a two-parameter family of *everywhere regular*, static, source-free Maxwell fields. It is shown that there exists a corresponding two-parameter family of unitarily inequivalent representations of the canonical commutation relations. Elements of the underlying Hilbert space may be interpreted as "quantum fluctuations of the Maxwell field off nontrivial classical vacua." The representation corresponding to the "trivial" sector—i.e., the zero classical solution— is the usual Fock representation. All others are "non-Fock." In particular, in all other sectors, the Maxwell field develops a nonzero vacuum expectation value. The parameters labelling the family can be interpreted as electric and magnetic charges. Therefore, unitary inequivalence naturally leads to superselection rules for these charges. These features arise in spite of the linearity of field equations only because the space–time topology is "nontrivial." Also, because of linearity, an exact analysis is possible at the quantum level; recourse to perturbation theory is unnecessary.

1. INTRODUCTION

The present investigation has two primary motivations.

The first has its origin in certain recent developments in field theory which suggest that nonperturbative effects may play a significant role in the quantum description of fields. Consider, for example, the $\lambda\phi^4$ or the Sine–Gordon interaction in two space–time dimensions or the coupled Yang–Mills and Higgs fields in Minkowski space–time. At the classical level, the corresponding field equations admit nonzero, everywhere regular static solutions which minimize the energy functional; solutions which have been interpreted as "nontrivial classical vacua." At the quantum level, one expects fluctuations around these vacua to influence probability amplitudes for various processes in ways which have escaped usual perturbative expansions in coupling constants. However, a complete analysis of this issue is not yet available: The detailed quantum-field-theoretic investigation is hampered by the fact that the field equations are *nonlinear*. Now, it turns out that if one considers space–times with "nontrivial" topologies, one can obtain examples of *linear* fields which admit physically interesting classical vacua.[1] In this case the linearity of field equations greatly simplifies the transition from classical to quantum description. As a result, one can obtain models of exactly soluble quantum field theories which bring out the role played, in quantum theory, by nontrivial classical vacua. Furthermore, it seems reasonable to expect that the qualitative features of the manifestations of these vacua are "universal" and will persist in more complicated, nonlinear models.

The second motivation comes from quantum aspects of the gravitational interaction. It is often stated that, since the strength of the gravitational coupling constant is so weak, inclusion of gravitational effects in quantum theory would only involve making minute corrections to various results obtained in the usual framework of Minkowskian quantum field theory. If this indeed turns out to be the case, predictions of a quantum theory of gravity would be rather uninteresting at least as far as the laboratory experiments are concerned. There is, however, another possibility: Inclusion of the gravitational interaction in the quantum picture might introduce *qualitatively new* features whose existence is quite independent of the strength of the coupling constant. Such effects, if they exist, might have ramifications already on the laboratory scale. To illustrate this point,[2] consider the transition from Newtonian framework to the special relativistic one. While it is true that most "relativistic effects" are significant only when the velocities involved are comparable to that of light, there is at least one "qualitatively new" feature which arises in this transition: The theoretical prediction that every elementary particle has an associated antiparticle! This prediction has its origin in the general structure of the Poincaré invariant equations and has nothing to do with the numerical value of the velocities of the particles involved. One would imagine that if quantum gravity has a significant role to play in laboratory physics, it would be via similar qualitatively new effects. An obvious direction to search for these effects is to analyze the role of space–time topology in quantum phenomena: The fact that nontrivial topologies are permissible is a qualitatively new aspect of general relativity, having nothing to do with the strength of the gravitational coupling constant.

The purpose of this note is to present an explicit exam-

a)Supported in part by the NSF contract PHY 76-81102 with the University of Chicago.

 0022-2488/80/030526-08$1.00

ple to illustrate the role of space–time topology by working in an external field approximation. More specifically, we shall examine the influence of "worm holes" on the quantum description of (linear, test) Maxwell fields. We shall see that this influence manifests itself via the presence of spontaneous symmetry breaking and non-Fock representations of the CCR (vacuum degeneracy) and leads to superselection rules for electric and magnetic charges. The example is rather elementary and is intended to provide only directions for more extensive analysis of "nonperturbative" phenomena with and without gravitational interaction.

2. THE BASIC IDEAS

Our model is based on the following ideas: More than ten years ago, Streater[3] pointed out, in connection with spontaneous breaking of symmetries, that for massless scalar fields in *Minkowski space* a family of non-Fock representations arises naturally and has a number of interesting features. Denote the scalar field by ϕ and consider the transformation $\phi \to \phi + k$, where k is a constant. This transformation leaves the Lagrangian density $(\nabla_a \phi \nabla^a \phi)$ invariant; it is a classical symmetry. This symmetry gives rise to a one-parameter family of automorphisms on the algebra of quantum (field) operators. However, the automorphism turns out *not* to be unitarily implementable in the standard Fock representation: The symmetry is spontaneously broken! Consequently, one obtains a one-parameter family of unitarily inequivalent representations—labeled by k—of the canonical commutation relations. This example has been very useful as a simple and exact mathematical model for illustrating theorems concerning symmetry breaking. Can one extend these ideas to higher spins? Let us begin with Maxwell fields. In this case, $A_a \to A_a + K_a$ is the corresponding transformation, where K_a is a constant vector field. However, the symmetry is now relegated to potentials; the fields $F_{ab} = \nabla_{[a} A_{b]}$ are left untouched by these transformation. Hence, if one works with F_{ab}'s—the directly observable quantities—rather than with A_a's, all the interesting phenomena just disappear![4] This state of affairs continues for higher spins as well.

However, if the restriction to Minkowski space is dropped, the situation is quite different owing to the availability of nontrivial global topologies: In this case, the extension to Maxwell fields *can* lead to phemomena of physical interest. Consider, for example, the Schwarzschild–Kruskal space–time. In this case, there is available a translation Killing field t^a (timelike at infinity) which is everywhere regular, and satifies, as a consequence of Einstein's vacuum equation, source-free Maxwell equations. Furthermore, the Maxwell fields $\mathring{F}_{ab} = \nabla_{[a} t_{b]}$ and its dual $^*\mathring{F}_{ab} = \epsilon_{abcd} \mathring{F}^{cd}$ are the only everywhere regular, stationary, linearly independent solutions of source-free Maxwell equations on this space–time. (It also turns out that the total energy—but *not* the energy density—associated with these solutions vanishes identically.) *Finally, although they satisfy source-free Maxwell equations everywhere, the net electric charge associated with* $\mathring{F}_{ab}$ *and the net magnetic charge associated with* $^*\mathring{F}_{ab}$ *fail to vanish*! This is a manifestation of the fact that the space–time topology is now nontrivial.[5] (Note that there is nothing "pathological" about the global structure of this space–time. In particular, the metric is analytic everywhere, the space–time is globally hyperbolic, and the Cauchy problem is well posed for test Maxwell fields.) Consider, next, transformations $F_{ab} \to F_{ab} + q\mathring{F}_{ab} + g^*\mathring{F}_{ab}$, where q and g are any real numbers. Under these transformations, the Lagrangian density changes only by the addition of a divergence of a vector field, and the Hamiltonian is left invariant. In this sense, the transformations represent classical symmetries. (Note that the transformation $A_a \to A_a + qt_a$ causes the field F_{ab} to change by $F_{ab} \to F_{ab} + q\mathring{F}_{ab}$. Since t^a is a translation Killing field, it is the closest analog of the constant vector fields in Minkowski space. In this sense, the present transformations are natural extensions of the ones noted above in the Minkowskian context.) These transformations do give rise to automorphisms on the usual algebra of field operators. It turns out that the automorphisms are *not* unitarily implemented in the Fock representation! The result is a two-parameter family of representations of the CCR. The elements of the Hilbert space corresponding to parameter values (q,g) describe the "quantum fluctuations of the Maxwell field off the nontrivial background $q\mathring{F}_{ab} + g^*\mathring{F}_{ab}$"; in this Hilbert space, the vacuum expectation value of the Maxwell field operator (-valued distribution) is precisely $q\mathring{F}_{ab} + g^*\mathring{F}_{ab}$. Unitary inequivalence of these representations implies that operators representing physical observables can not "mix" states labeled by different values of electric and magnetic charges. In this sense, charge superselection is built in! To summarize, because Schwarzschild–Kruskal space–time admits a nontrivial topology, the extension of Streater's ideas to electromagnetic fields leads to interesting quantum phenomena. In the resulting model, the space of classical solutions is naturally divided into various sectors labeled by values of electric and magnetic charges such that, in the quantum description, each sector is endowed with a distinct representation of the CCR.

The model is discussed in the next section. Although the general procedure followed is analogous to that followed by Streater, there is, nonetheless, an important difference: While, in a certain sense,[6] Streater's model involves nontrivial representations of the CCR only for potentials, the present model yields such representations for fields as well. This difference reflects itself in the mathematical techniques required. In Streater's model, delicate functional-analytic issues arose in the proof of unitary inequivalence of the representations based on distinct classical vacua; introduction of the Fock representation was straightforward. In the present case, on the other hand, these delicate issues arise in the construction of the Fock representation; the proof of unitary inequivalence uses only elementary results from topology and differential geometry. Thus, although both models exhibit symmetry breaking and vacuum degeneracy, the "origin" of these phemomena is quite distinct in the two cases.

Our analysis is closely related also to Sorkin's work[7] on quantization of Maxwell fields in multiply connected space–times. However, the goals are different. Specifically, Sorkin considers a wide class of multiply connected space–times and obtains superselection rules by examining the algebra of quantum operators; he does not analyze the problem of con-

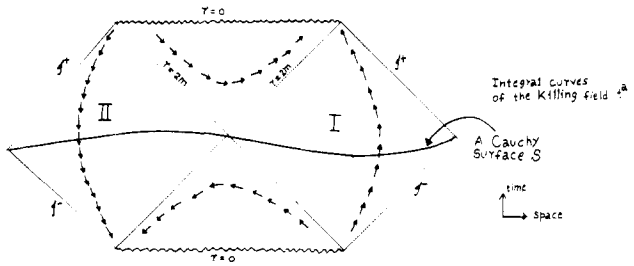

FIG. 1. The Penrose diagram of the Schwarzschild–Kruskal space-time. There are two asymptotic regions I and II. The translational Killing field t^a has been chosen to be future directed in region I and past directed in region II.

structing a Hilbert space of states. In this paper, on the other hand, we restrict our detailed analysis to the Schwarzschild–Kruskal space–time but construct the Hilbert space in detail. It is this construction that enables us to demonstrate that interesting "nonperturbative" phemomena can occur ever for linear fields provided the underlying space–time topology is nontrivial.

3. THE MODEL

This section is divided into four parts: In the first, we introduce the nontrivial classical vacua; in the second, the algebra of quantum operators satisfying the CCR; in the third, the Fock representation; and in the fourth, the two parameter family of unitarily inequivalent representations.

A. Nontrivial classical vacua

Let $(\mathbf{M}, g_{ab})$ be the Schwarzschild–Kruskal space–time.[8] Let t^a denote the translational Killing field which is future directed, timelike in region I, and is normalized such that $\lambda := -t_a t^a$ tends to 1 at infinity (Fig. 1). (Note, by continuity, t^a must be past directed in region II.) Set $\mathring{F}_{ab} = \nabla_a t_b$. Then, $\nabla_{[a}\mathring{F}_{bc]} = 0$ by inspection and $\nabla_a \mathring{F}^{ab} = \nabla_a \nabla^a t^b = -R^{ab} t_b = 0$, where we have used the affine collineation equation on t_a (satisfied by any affine Killing field) and the fact that g_{ab} satisfies Einsteins's vacuum equation. Thus, $\mathring{F}_{ab}$—and hence also its dual $^*\mathring{F}_{ab} = \epsilon_{abcd}\mathring{F}^{cd}$—is a source-free Maxwell field. Next, let us compute the electric and magnetic charges associated with these fields. Given any 2-sphere S_2, homotopic to the 2 sphere $r = $ const. in region I, we have

$$\int_{S_2} {}^*\mathring{F}_{ab}\, dS^{ab} = \int_{S_2} \epsilon_{abcd} t^d \lambda^{-1} \nabla^c \lambda dS^{ab} = 8\pi M\,, \qquad (1)$$

where M is the Schwarzschild mass of the underlying space–time $(\mathbf{M}, g_{ab})$. [Note that the integral in Eq.(1) is just the Komar scalar[9] associated with the Killing field t^a.] Therefore, although $\mathring{F}_{ab}$ and $^*\mathring{F}_{ab}$ satisfy source-free equations *everywhere*, the electric charge associated with $\mathring{F}_{ab}$ and the magnetic charge associated with $^*\mathring{F}_{ab}$ fail to vanish! These electromagnetic charges are of "topological origin"; this is a particular—and in a sense simplest—example of Wheeler's "charge without charge"[5] (see Fig. 2). While $\mathring{F}_{ab}$ has a *natural* vector potential t_a, as one might expect, $^*\mathring{F}_{ab}$ can not admit *any* globally well behaved vector potential.

In addition to satisfying source-free Maxwell equations, $\mathring{F}_{ab}$ and $^*\mathring{F}_{ab}$ share a number of interesting properties. First, it is obvious from their definitions that they are static, everywhere regular, and spherically symmetric. Second, because the Killing field t^a is future directed in region I and past directed in region II, it follows that the total energy of *both* these fields vanishes identically. Finally, one can show that these two are the only linearly independent, everywhere regular, static, source-free solutions to Maxwell's equations on $(\mathbf{M}, g_{ab})$. This last property is nontrivial and follows either from the explicit calculations[10] on stability of the Schwarzschild black hole against electromagnetic perturbations or from an extension[11] of Hodge theory to asymptotically flat spaces.

These properties imply that the space of all source-free, regular, i.e., say of finite energy separately in regions I and II, solutions to Maxwell's equations can be divided into a two-parameter family of sectors; each sector admits one and only one static solution; and the parameters labelling these sectors have the natural interpretation of electric and magnetic charges. Thus, for all real number q and g, $q\mathring{F}_{ab} + g^*\mathring{F}_{ab}$ represents a "classical vacuum state" of the electromagnetic field.

B. Quantization: Preliminaries

1. *Formulation of the problem*

The problem of quantizing Maxwell fields on $(\mathbf{M}, g_{ab})$ may be formulated as follows: Find a Hilbert space $\mathscr{H}$ and a representation thereon of the operator valued distribution $\mathbf{F}_{ab}(x)$ satisfying Maxwell's equations

$$\nabla_{[a}\mathbf{F}_{bc]} = 0 \quad \text{and} \quad \nabla_a \mathbf{F}^{ab} = 0\,, \qquad (2)$$

and the CCR

$$[\mathbf{F}_{ab}(x), \mathbf{F}_{a'b'}(x')] = \hbar/iG_{aba'b'}(x,x')\,\mathbf{I}\,, \qquad (3)$$

where $G_{aba'b'}(x,x')$ is the Lichnerowicz propagator.[12] (Here, indices a and b refer to the tangent space at x and a', b' to that at x'.) Choose, as usual, the space of test fields to be the space τ of real, second-rank, skew, C^∞ tensor fields t^{ab} of compact support on $\mathbf{M}$. Then, the problem is that of constructing a Hilbert space $\mathscr{H}$ and of introducing, for each t^{ab} in τ, a self-adjoint operator $\mathbf{F}(t)$ on $\mathscr{H}$ satisfying (i) $\mathbf{F}(t) + \mathbf{F}(\bar{t}) = \mathbf{F}(t + \bar{t})$; (ii) $\mathbf{F}(t) = 0$ if t^{ab} belongs to the subspace K of

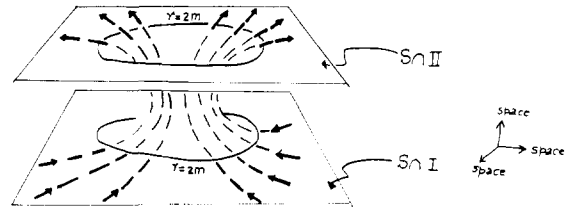

FIG. 2. A pictorial "explanation" of charge without charge. Every static Cauchy surface in $(\mathbf{M}, g_{ab})$ admits two distinct asymptotic regions $S\cap$I and $S\cap$II, connected to each other at the throat via an identification. S does not intersect the region $r < 2M$. Lines of force of a static electric field with charge q in region $S\cap$II and $-q$ in $S\cap$I have been sketched. It is because the topology is $S^2 \times \mathbb{R}$ rather than $\mathbb{R}^3$ that regular nonzero, source-free electric fields can exist: Lines of force can converge in from infinity in $S\cap$I, pass through the throat, and diverge to infinity in $S\cap$II.

τ consisting of fields of the type $\nabla^{[a}t^{b]}$ or $\nabla_c t^{abc}$, where t^a is a C^∞ vector field and t^{abc}, a C^∞ totally skew tensor field, of compact support on **M**; and (iii) $[\mathbf{F}(t),\mathbf{F}(\bar{t})] = \hbar/i\,[\int_{\mathbf{M}}\int_{\mathbf{M}} G_{aba'b'}(x,x')t^{ab}(x)\bar{t}^{a'b'}(x')]\,\mathbf{I}$. The $\mathbf{F}(t)$'s are the smeared out field operators.

2. *The abstract * algebra of field operators*

It will be convenient to have at one's disposal an abstract * algebra $\mathscr{A}$ generated by field operators which is not tied down to any specific Hilbert space.[13] This algebra $\mathscr{A}$ can be constructed as follows: Consider first the tensor algebra (over complexes) $\hat{\mathscr{A}}$ generated by the vector space τ: $\hat{\mathscr{A}} = \oplus_{k=0}^{\infty}\tau^{\otimes k}$. Thus, a typical element A of $\hat{\mathscr{A}}$ is a string with only a finite number of nonzero entries $A = (t_0, t_1, t_2, \ldots, t_k, 0,0,\ldots)$, where t_0 is any complex number and t_k any k^{th} rank tensor over the vector space τ. Sums and products of these strings can be defined in an obvious way; $\hat{\mathscr{A}}$ is an associative (non-Abelian) algebra. Next, we define a * operation on $\hat{\mathscr{A}}$. We demand that the action of * on complexified τ be the complex conjugation and extend its action on all of $\hat{\mathscr{A}}$ by requiring that it be complex antilinear, i.e., satisfy $(A+\lambda B)^* = A^* + \lambda^* B^*$, and obey the product rule $(A\cdot B)^* = B^*\cdot A^*$, for all A and B in $\hat{\mathscr{A}}$ and complex numbers λ, where λ^* is the complex conjugate of λ. Then, we automatically have $(A^*)^* = A$; the * operation is an involution. Thus, $\hat{\mathscr{A}}$ is now a * algebra. To impose Maxwell's equations and the CCR, we proceed as follows: Consider the * ideal $\hat{\mathscr{I}}$ of $\hat{\mathscr{A}}$ generated by elements of $\hat{\mathscr{A}}$ of the type $(\hbar/i\,G(t,\bar{t}),\hat{t}$, $t\otimes\bar{t} - \bar{t}\otimes t, 0,0,\ldots 0,\ldots)$, where t and $\bar{t}$ are arbitrary elements of τ, $G(t,\bar{t}) = \int_{\mathbf{M}}\int_{\mathbf{M}} G_{aba'b'}(x,x')\,t^{ab}(x)\,\bar{t}^{a'b'}(x')$, and where $\hat{t}$ belongs to the subspace K of τ. Set $\mathscr{A} = \hat{\mathscr{A}}/\hat{\mathscr{I}}$. This $\mathscr{A}$ is the required * algebra. It is easy to check that each element $\{t^{ab}\}$ of τ/K gives rise, unambiguously, to an element of $\mathscr{A}$ which we shall denote by $\mathbf{F}(t)$. (Here, following the usual convention in the literature, we have dropped the symbol $\{\ \}$ denoting "equivalence class" for simplicity.) These $\mathbf{F}(t)$'s are, of course, the smeared-out field operators. One can regard $\mathscr{A}$ as being "generated" by these $\mathbf{F}(t)$'s.

3. *Relation between test fields and classical solutions*

Denote by Γ the space of C^∞ solutions to Maxwell's equations which induce, on any Cauchy surface, initial data of compact support. There is a natural linear mapping from the space τ/K onto Γ which plays an important role in the quantum description: Since $G_{aba'b'}(x,x')$ satisfies Maxwell's equations, it follows that for all elements $\{t^{ab}\}$ of τ/K, $F^t_{ab}(x) := \int_{\mathbf{M}} G_{aba'b'}(x,x')\,t^{a'b'}(x')$ is a Maxwell field in Γ which is independent of the particular choice t^{ab} of the test field from the equivalence class $\{t^{ab}\}$ in τ/K, used in the integral. It turns out, however, that this mapping is *not* 1–1: We shall see in Sec. 3D that $G_{aba'b'}(x,x')$ has a nontrivial kernel. (that the mapping is onto follows from a theorem due to Fourès-Bruhat.[14] The action of G on $\tau\times\tau$ gives rise to a linear mapping Ω from $\Gamma\times\Gamma$ to the reals:

$$\Omega(F^t,F^{\bar{t}}) := G(t,\bar{t})\,. \tag{4}$$

Since G is skew symmetric,[12] so is Ω. Next, it follows from the definition of $F^t_{ab}(x)$ that $\Omega(F^t,F^{\bar{t}}) = 0$ for all F^t in Γ if and only if the field $F^t_{ab}(x) = 0$. Thus, unlike G, Ω has a *trivial* kernel. Hence, Ω is a weakly nondegenerate sympletic structure on Γ. Finally, using various properties of G, it is possible to re-express Ω in a way which makes no reference at all to test fields: $\Omega(F,\bar{F}) = \int_S (F_{ab}\bar{A}^b - \bar{F}_{ab}A^b)\,dS^a$, where S is any Cauchy surface in $(\mathbf{M}, g_{ab})$ and A and $\bar{A}$ are, respectively, vector potentials for F and $\bar{F}$. (The existence of A and $\bar{A}$ follows from the fact F and $\bar{F}$ have initial data of compact support,[11] while the surface independence and gauge invariance of the integral follows from Maxwell's equations).

C. The Fock representation

As remarked in Sec. 2, construction of the Fock representation [singled out by the isometry group of $(\mathbf{M}, g_{ab})$] turns out not to be straightforward: Since the Killing field t^a fails to be everywhere future directed timelike, one can not directly apply various theorems available for quantization of test fields on stationary backgrounds. For Maxwell fields, the situation is further complicated due to the presence of nontrivial classical vacua: Since these solutions are everywhere regular and represent zero frequency modes, one is faced with ambiguities when one attempts to decompose fields into positive and negative frequency parts. Therefore, before going on to discuss symmetry breaking and the associated non-Fock representations, we shall introduce the Fock representation in detail.

Introduce on Γ the "energy inner product" $(\ ,\ )$ as follows:

$$(F,\widetilde{F}) := \int_{S_{\mathrm{II}}} - \int_{S_{\mathrm{I}}} (F_{m(a}\widetilde{F}_{b)}{}^{m} + {}^*F_{m(a}{}^*\widetilde{F}_{b)}{}^{m})t^a\,dS^b, \tag{5}$$

where S_{I} and S_{II} are, respectively, the restrictions to regions I and II of any smooth Cauchy surface S in $(\mathbf{M}, g_{ab})$, passing through the "throat" (see Fig. 1). [That the integral is independent of S follows from Maxwell's equations. Next, since t^a is future directed in region I and past directed in region II, we have $(F,F)\geqslant 0$, the equality holding if and only if $F = 0$; the norm induced on Γ by this inner product is positive definite.] Let $\bar{\Gamma}$ denote the Cauchy completion of $(\Gamma, (,))$. Note that although static solutions $\overset{\circ}{F}_{ab}$ and ${}^*\overset{\circ}{F}_{ab}$ have finite energy norm, they do *not* belong to $\bar{\Gamma}$: It is easy to verify that $\overset{\circ}{F}$ and ${}^*\overset{\circ}{F}$ are orthogonal to $\bar{\Gamma}$ in the Hilbert space of all solutions with finite energy norm.[15] Introduce on $\bar{\Gamma}$ an operator T with domain $D(T) = \Gamma$: $T\cdot F \equiv \mathscr{L}_t F_{ab}$ for all F in Γ. This T has the following properties: (i) $T\subset -T^*$; $(T, D(T) = \Gamma)$ is a skew-symmetric operator on $\bar{\Gamma}$;[16] (ii) the range $R(T)$ of T is dense in $\bar{\Gamma}$, or, equivalently, the kernel of T^* contains only the zero vector (we repeat that static solutions do not belong to $\bar{\Gamma}$!); and (iii) the symplectic structure Ω on Γ admits an extension $\bar{\Omega}(\ ,\ )$ which is a linear mapping from $\bar{\Gamma}\times R(T)$ into reals which is continuous in the first variable. In order to obtain the Fock representation of the CCR in which the time-translation isometry is unitarily implemented, we must introduce on $\bar{\Gamma}$ a complex structure J, i.e., a continuous linear operator J on $\bar{\Gamma}$ satisfying $J^2 = -I$, which is compatible with $\bar{\Omega}$ in the following sense: $\bar{\Omega}(JF,\tilde{F}) = \bar{\Omega}(J\tilde{F},F)$, and $\bar{\Omega}(JF,F)\leqslant 0$, the equality holding if and only if $F = 0$, for all $F,\tilde{F}$ in $R(T)$. (For details, see, for example, Ref. 17.)

We now introduce the required J. Since T is skew symmetric with $D(T)$ dense in $\bar{\Gamma}$, it follows that T is closable.[18] Denote the closure by $\bar{T}$. Since $T \subset \bar{T}$ and since $R(T)$ is dense in $\bar{\Gamma}$, $R(\bar{T})$ is also dense in $\bar{\Gamma}$. Let $\hat{J}$ denote the isometry part of the polar decomposition of $\bar{T}$[19]: Set $\hat{J} = (\bar{T}^*\bar{T})^{1/2}\,\bar{T}^{-1}$ on $R(\bar{T})$ and extend it to $\bar{\Gamma}$ by continuity. This $\hat{J}$ is a complex structure on $\bar{\Gamma}$. Furthermore, it commutes with $\exp(\lambda\,\bar{T})$ for all λ in $\mathbb{R}$, which is the action on $\bar{\Gamma}$ of the isometry group generated[16] by t^a on $(\mathbf{M}, g_{ab})$. However, this $\hat{J}$ is *not* compatible with $\bar{\Omega}$: Because the Killing field is past directed in region II, $-\bar{\Omega}(\hat{J}F,F)$ fails to be positive definite on $R(T)$. We therefore proceed as follows. Introduce on $\bar{\Gamma}$ an operator ϵ: For all F in Γ, ϵF is the element of $\bar{\Gamma}$ which coincides with F_{ab} in region I and with $-F_{ab}$ in region II. Since I and II are causally disjoint, ϵ is well defined. (Note that although ϵF is discontinuous at the throat, it does belong to $\bar{\Gamma}$.[20]) Furthermore, ϵ is a continuous operator with respect to the energy norm: $(\epsilon F,\epsilon F) = (F,F)$ for all F in Γ. Hence, it admits a continuous extension to $\bar{\Gamma}$ which we denote also by ϵ. Next, we note certain properties of this operator: (i) $\epsilon^2 = I$; (ii) ϵ commutes with $\exp(\lambda\bar{T})$ on $\bar{\Gamma}$, and hence also with J, which is expressible in terms of the projector family of $\exp(\lambda\bar{T})$; and (iii) $\bar{\Omega}(\epsilon F, T\cdot\tilde{F}) = (F,\tilde{F})$ for all F in $\bar{\Gamma}$ and $\tilde{F}$ in Γ. Set $J = \epsilon\cdot\hat{J}$. Then J is a complex structure on $\bar{\Gamma}$ which is compatible with $\bar{\Omega}$ and which commutes with $\exp(\lambda\bar{T})$. This J is the required complex structure. Finally, by a simple extension of the arguments given in Ref. 17, one can show that J is the unique complex structure on $\bar{\Gamma}$ with these properties.

Using $(\bar{\Gamma}, \bar{\Omega}, J)$, we can now obtain the required Fock representation. Introduce on $R(\bar{T})$ the following Hermitian inner product:

$$\langle F,\tilde{F}\rangle = (1/2\hbar)\bar{\Omega}(JF,\tilde{F}) + (i/2\hbar)\,\bar{\Omega}(F,\tilde{F})$$
$$\equiv (1/2\hbar)\,(F,(\bar{T}^*\bar{T})^{-1/2}\cdot\tilde{F}) + (i/2\hbar)\,\bar{\Omega}(F,\tilde{F}).$$

Then, $(R(T), J, \langle,\rangle)$ is a complex pre-Hilbert space.[21] Denote the Cauchy completion by H. Let $\mathscr{H}$ be the symmetric Fock space based on this H. Set $\pi\cdot[\mathbf{F}(t)] = \hbar\mathbf{C}(F^t) + \hbar\mathbf{A}(F^t)$, where $\mathbf{C}(F^t)$ and $\mathbf{A}(F^t)$ are, respectively, the creation and annihilation operators on $\mathscr{H}$ associated with F^t in H. Using the commutation relations between $\mathbf{C}$'s and $\mathbf{A}$'s—which are determined by the Fock space structure of $\mathscr{H}$—one has $[\pi\cdot\mathbf{F}(t), \pi\cdot\mathbf{F}(\tilde{t})] = \hbar^2[\mathbf{C}(F^t),\mathbf{A}(F^t)] + \hbar^2[\mathbf{A}(F^t),\mathbf{C}(F^t)] \subset \hbar/i\Omega(F^t,F^t)\mathbf{I} = \hbar/iG(t,\tilde{t})\mathbf{I} = \pi\cdot[\mathbf{F}(t),\mathbf{F}(\tilde{t})]$. Hence, the mapping π from the space of smeared-out field operators into the space of operators on $\mathscr{H}$ admits an extension to all of $\mathscr{A}$; $(\mathscr{H},\pi)$ is a representation of the CCR. This is a Fock representation: $\mathscr{H}$ is a symmetric Fock space based on the Hilbert space H of suitably regular solutions to the classical field equations and the smeared-out field operators are represented by sums of creation and annihilation operators on $\mathscr{H}$. Using the definition of J and the fact that the Killing field t^a commutes with all Killing fields on $(\mathbf{M}, g_{ab})$, one can show that the entire four-parameter group of isometries is unitarily implemented on $(\mathscr{H},\pi)$. (See Ref. 17, Sec. 3.)

Remarks: (1) The above construction is a natural generalization of the procedure used to obtain the standard Fock representation of massless Bose fields in stationary space-times.[17] The key idea was to decompose the fields into positive and negative frequency parts and to build the one particle Hilbert space H from solutions which are of positive frequency in region I and of negative frequency in region II. (The use of complex structure $\hat{J}$ in place of J would have corresponded to using positive frequences in both regions; the space of one-particle states would then have inherited an indefinite metric.) For massless scalar fields, the analogous representations have been obtained by Israel[22] and Unruh.[23] Our construction rigorizes their work and extends it to the case of Maxwell fields.

(2) What is the status of the zero-frequency modes in the final quantum picture? We have seen that these modes do not belong to the Hilbert space $\bar{\Gamma}$. Since the one-particle Hilbert space was obtained by the Cauchy completion of $R(\bar{T}) \subset \bar{\Gamma}$ with respect to the norm $\langle F,F\rangle = (F,(\bar{T}^*\bar{T})^{-1/2}\cdot F)$, and since the zero frequency modes are in the kernel of $\mathscr{L}_t$, it follows that they do *not* belong to H either. This is just what one would expect on intuitive grounds: One expects photon states to correspond only to "radiative" modes of the Maxwell fields while the zero-frequency fields correspond to the "longitudinal" modes.

(3) Note that the representation $(\mathscr{H},\pi)$ is not faithful: The element $\mathbf{F}(t)$ of $\mathscr{A}$ is sent to the zero operator by π if the element $\{t^{ab}\}$ in τ/K is in the kernel of the Lichnerowicz propagator $G_{aba'b'}(x,x')$. We shall show in the next subsection that the kernel is in fact nontrivial.

D. Symmetry breaking and "non-Fock" representations

Consider the two parameter family of transformations $\alpha_{(q,g)}$ on the space of classical solutions to Maxwell's equations, defined by $\alpha_{(q,g)}\cdot F_{ab} = q\mathring{F}_{ab} + g^*\mathring{F}_{ab} + F_{ab}$. Then, $\alpha_{(q,g)}$ sends the sector of classical solutions labeled by $(\tilde{q},\tilde{g})$ to that labeled by $(q+\tilde{q}, g+\tilde{g})$. Denote by $\Delta_{(q,g)}\mathscr{L}$ the change in the Lagrangian density under $\alpha_{(q,g)}$. Then, using the fact that every Maxwell field admits, locally, a vector potential, it follows that given any point of $\mathbf{M}$, there exists a neighborhood of this point such that $\Delta_{(q,g)}\mathscr{L}$ can be expressed as the divergence of a vector field. (If we restrict ourselves to sectors and transformations with zero g and $\tilde{g}$, then $\Delta_{(q,g)}\mathscr{L}$ can be expressed as a divergence *globally*; the restriction to local neighborhoods is required, as usual, only when the magnetic charges involved are nonzero.) Furthermore, a simple calculation shows that the *total* Hamiltonian is left invariant by these $\alpha_{(q,g)}$'s. In this sense, these transformations represent classical symmetries.

Can these symmetries be promoted to the quantum level? In terms of the operator-valued distribution $\mathbf{F}_{ab}(x)$, the corresponding transformations are clearly $\mathbf{F}_{ab}(x) \to \mathbf{F}_{ab}(x) + q\mathring{F}_{ab}\mathbf{I} + g^*\mathring{F}_{ab}\mathbf{I} \equiv \Lambda_{(q,g)}\cdot\mathbf{F}_{ab}(x)$, say. Clearly, the transformed distribution $\Lambda_{(q,g)}\cdot\mathbf{F}_{ab}(x)$ again satisfies Maxwell's equations (2) and the CCR (3). Hence, $\Lambda_{(q,g)}$ extends to an automorphism on the * algebra $\mathscr{A}$. In particular, for smeared-out field operators $\mathbf{F}(t)$ we have

$$\Lambda_{(q,g)}\cdot\mathbf{F}(t) = \mathbf{F}(t) + \left(\int_{\mathbf{M}} (q\mathring{F}_{ab} + g^*\mathring{F}_{ab})t^{ab}\right)\mathbf{I}. \qquad (6)$$

[Note that Eq. (6) is a well-defined mapping: Because F_{ab} and $^*\mathring{F}_{ab}$ satisfy Maxwell's equations, if t^{ab} and $\tilde{t}^{ab}$ belong to the same equivalence class in τ/K, i.e., if $\mathbf{F}(t) = \mathbf{F}(\tilde{t})$ in $\mathscr{A}$,

then $\Lambda_{(q,g)}\cdot\mathbf{F}(t) = \Lambda_{(q,g)}\cdot\mathbf{F}(t)$ in $\mathscr{A}$ for all real numbers q and g.] Thus, classical symmetries $\alpha_{(q,g)}$ do indeed extend to quantum operators at the algebraic level. The obvious question now is whether (in the Fock representation) they also extend to quantum states, i.e., for any real numbers (q,g), does there exist a unitary mapping $U_{(q,g)}$ on the Fock space $\mathscr{H}$ of quantum states such that $U^{-1}_{(q,g)}\cdot(\boldsymbol{\pi}\cdot\mathbf{F}(t))\cdot U_{(q,g)}$ $\equiv U^{-1}_{(q,g)}\cdot(\mathbf{C}(F^t)+\mathbf{A}(F^t))\cdot U_{(q,g)}$ equals $\boldsymbol{\pi}(\Lambda_{(q,g)}\cdot\mathbf{F}(t))$ $\equiv \mathbf{C}(F^t)+\mathbf{A}(F^t)+(\int_M(q\mathring{F}_{ab}+g^*\mathring{F}_{ab})t^{ab})\,\mathbf{I}$?

Theorem: $\Lambda_{(q,g)}$ can not be unitarily implemented on $(\mathscr{H},\boldsymbol{\pi})$ unless $q=g=0$.

Proof: We first exhibit test fields $\hat{t}^{ab}$ which are in the kernel of the Lichnerowicz propagator $G_{aba'b'}(x,x')$ *and* satisfy $\int_M \mathring{F}_{ab}(x)\hat{t}^{ab}(x)\neq 0$. Set $\hat{t}^{ab} = f(r)h(t)\,\nabla^{[a}r\nabla^{b]}t$, where r and t are the usual Schwarzschild coordinates and where $f(r)$ and $h(t)$ are any C^∞ functions of compact support in their respective arguments such that the support of $\hat{t}^{ab}$ is contained strictly in region I (Fig. 1), satisfying $\int_{2m}^{\infty} f(r)dr = \int_{-\infty}^{\infty} h(t)dt = 1$. Then, given any ($C^1$) Maxwell field F_{ab}, we have

$$\int_M F_{ab}\hat{t}^{ab} = Q\,, \tag{7}$$

where Q is the electric charge associated with the given F_{ab} in region I. Recall that given any test field t^{ab} in τ, the corresponding solution F^t_{ab} to Maxwell's equations has initial data of compact support on any Cauchy surface, and hence zero electric charge. Therefore, for all test fields t^{ab} in τ we have

$$\begin{aligned} G(t,\hat{t}) &= \int_M\int_M G_{aba'b'}(x,x')t^{ab}(x)\hat{t}^{a'b'}(x') \\ &= \int_M F^t_{ab}(x')\hat{t}^{a'b'}(x') \\ &= 0\,. \end{aligned} \tag{8}$$

Hence, $\hat{t}^{ab}$ is in the kernel of the propagator $G_{aba'b'}(x,x')$. Next, using Eq. (7), we have $\int_M q\mathring{F}_{ab}(x)\hat{t}^{ab}(x) = 8\pi Mq$, the electric charge of $q\mathring{F}_{ab}$. Thus, $\hat{t}^{ab}$ has the desired properties.

Now, let us suppose that the automorphism $\Lambda_{(q,g)}$ is unitarily implementable on $(\mathscr{H},\boldsymbol{\pi})$. Then there exists a unitary operator $U_{(q,g)}$ on $\mathscr{H}$ which satisfies, in particular,

$$\begin{aligned} U^{-1}_{(q,g)}\cdot(\mathbf{F}(\hat{t}))\cdot U_{(q,g)} &= \boldsymbol{\pi}(\Lambda_{(q,g)}\cdot\mathbf{F}(\hat{t})) \\ &= \boldsymbol{\pi}\Big(\mathbf{F}(\hat{t}) + \int_M q\mathring{F}_{ab}\hat{t}^{ab}\Big)\mathbf{I}\,. \end{aligned} \tag{9}$$

However, since $\hat{t}^{ab}$ is in the kernel of $G_{aba'b'}(x,x')$, we have $\boldsymbol{\pi}\cdot(\mathbf{F}(\hat{t})) \equiv \mathbf{C}(F^{\hat{t}}) + \mathbf{A}(F^{\hat{t}}) = 0$. Hence, Eqs. (7) and (9) yield

$$0 = (8\pi Mq)\,\mathbf{I} \tag{10}$$

on the Hilbert space $\mathscr{H}$, implying $q=0$. Thus, $\Lambda_{(q,0)}$ is unitarily implementable in the Fock representation $(\mathscr{H},\boldsymbol{\pi})$ if and only if $q=0$. An identical argument using $\hat{t}^{ab} + \epsilon^{ab}{}_{cd}\hat{t}^{cd}$ in place of $\hat{t}^{ab}$ shows that $\Lambda_{(q,g)}$ is unitarily implementable on $(\mathscr{H},\boldsymbol{\pi})$ if and only if $q=g=0$. □

Remark: The essential facts used in the proof are that in the Fock representation $(\mathscr{H},\boldsymbol{\pi})$, the mapping $\boldsymbol{\pi}$ from $\mathscr{A}$ into the space of operators on $\mathscr{H}$ has a nontrivial kernel and that it fails to be invariant under the action of automorphisms $\Lambda_{(q,g)}$ on $\mathscr{A}$ unless $q=g=0$. The existence of the kernel itself has nothing to do with the nontrivial topology: The test field $\hat{t}^{ab}$ continues to be in the kernel of the Lichnerowicz propagator even in Minkowski space–time. The topology plays a role through the existence of the automorphisms which arise due to the presence of nontrivial classical vacua.

The above theorem implies that classical symmetries $\alpha_{(q,g)}$ are spontaneously broken in the Fock representation. This symmetry breaking leads, as usual,[3] to a two parameter family $(\mathscr{H}_{(q,g)},\boldsymbol{\pi}_{(q,g)})$ of unitarily inequivalent representations of the CCR. Perhaps the easiest way to exhibit these representations is the following: Choose for $\mathscr{H}_{(q,g)}$ a copy of $\mathscr{H}$ and set $\boldsymbol{\pi}_{(q,g)}\cdot\mathbf{F}(t) = \boldsymbol{\pi}\cdot\Lambda_{(q,g)}\cdot\mathbf{F}(t) = \mathbf{C}(F^t) + \mathbf{A}(F^t)$ $+\int_M(q\mathring{F}_{ab}+g^*\mathring{F}_{ab})\,t^{ab}$, for all smeared out field operators $\mathbf{F}(t)$ in $\mathscr{A}$. Clearly, $(\mathscr{H}_{(0,0)},\boldsymbol{\pi}_{(0,0)})$ is the Fock representation. All others are non-Fock: Using the result of the theorem above, it is easy to check that no two of the two parameter family $(\mathscr{H}_{(q,g)},\boldsymbol{\pi}_{(q,g)})$ of representations of the CCR are unitarily equivalent.

Remark: There is another way of exhibiting these representations which, in spite of being more complicated from a technical viewpoint, brings out the physical interpretation more directly. Recall that the Fock space $\mathscr{H}$ may be introduced as the space of all entire holomorphic functions ψ on the one-particle Hilbert space H for which the integral $\int_H|\psi|^2 d\mu$ converges, where $d\mu$ is the Gaussian promeasure invariantly associated with the Hilbert space H. It turns out than *any* element of $\oplus_{q,g}\mathscr{H}_{(q,g)}$ can be similarly represented by holomorphic functions but now on the space of *all* classical solutions with finite energy norm, including the static ones. In this description, elements of $\mathscr{H}_{(q,g)}$, for a fixed value of q and g, are described by holomorphic functions with support only on the sector of classical solutions with electric charge q and magnetic charge g. Hence, one can think of elements of $\mathscr{H}_{(q,g)}$ as describing "quantum fluctuations of the Maxwell field off the nontrivial classical background $q\mathring{F}_{ab}+g^*\mathring{F}_{ab}$." The vacuum state in $\mathscr{H}_{(q,g)}$ itself arises as a holomorphic function representing a coherent state peaked at the classical vacuum $q\mathring{F}_{ab}+g^*\mathring{F}_{ab}$. Finally, it is easy to check that *any* quantum state ψ with support on the (q,g) sector, i.e., any element of $\mathscr{H}_{(q,g)}$, has the property that the expectation values of electric and magnetic charge operators equal, respectively, $8\pi Mq$ and $8\pi Mg$.

4. DISCUSSION

(1) Of the two-parameter family of systems $(\mathscr{H}_{(q,g)},\boldsymbol{\pi}_{(q,g)})$ obtained in the previous section, only $(\mathscr{H}_{(0,0)},\boldsymbol{\pi}_{(0,0)})$ is the Fock representation of the CCR: While the vacuum state in $\mathscr{H}_{(0,0)}$ is analogous to the standard vacuum state in Minkowski space, other vacua have no counterparts at all anywhere in the Fock space of Minkowskian photons. We wish to emphasize, however, that, at lease *a priori*, there is nothing unphysical about these non-Fock representations. It is easy to check, for example, that the four parameter family of isometries of the underlying space–time are unitarily implemented in any of these representations. Furthermore, these unitary transformations leave the vacuum state invariant in all sectors. How does one reconcile this situation with the

theorems[24,17] on the uniqueness of the vacuum state for linear Bose systems? There do indeed exist *mathematical* conditions which single out the Fock representation from all others. However, these include a continuity requirement on the vacuum expectation values of Weyl operators $\exp[\, i\, \mathbf{F}(t\,)]$, a requirement which has no direct physical motivation but which simplifies the mathematical analysis considerably. (For details, see, for example, Refs. 3 and 17.) It turns out that the non-Fock representations obtained here—as well as those obtained by Streater[3]—satisfy all conditions of the theorems characterizing the Fock representation, except this continuity requirement. From a physical viewpoint, therefore, there is no principle which can discriminate against these new quantum vacua.

(2) Although, for simplicity, we have restricted ourselves to the Schwarzschild–Kruskal space–times, extension of the present analysis to globally hyperbolic, stationary, topologically nontrivial space–times is rather straightforward. If Einstein's vacuum equations are not satisfied, one can not construct the nontrivial vacua from the Killing field itself. Nonetheless, if Cauchy surfaces are either compact or suitably asymptotically flat, results from Hodge theory ensure the existence of these vacua.[7] Using the presence of the time-like Killing field, one can construct the Fock representation, introduce classical symmetries analogous to $\alpha_{(q,g)}$, show that they are spontaneously broken, and obtain new representations of the CCR. In the nonstationary contexts, on the other hand, a "standard" Fock representation is simply not available. Nonetheless, one can construct ("observer dependent") representations relative to Cauchy surfaces. If the Cauchy surface has a suitably nontrivial topology, one can again repeat the analysis (although, in this case, the classical vacua will, of course, not be stationary). Thus, the essential features of the model are really of "topological origin"; they do not refer to the detailed properties, e.g., the spherical symmetry, the presence of a horizon, the satisfaction of Einstein's vacuum equation, etc., of the Schwarzschild–Kruskal space–time.

(3) In the finished picture, the presence of a nontrivial topology does indeed lead to a "qualitatively new" prediction. To see this, we must first physically interpret particles whose quantum states are described by $(\mathscr{H}_{(q,g)}, \boldsymbol{\pi}_{(q,g)})$. Perhaps the easiest way to get a feel for the nature of these particles is to couple them to other fields and analyze the resulting interaction. For simplicity, let us restrict ourselves to the $(q,0)$ sector and do quantum electrodynamics replacing the usual photon states by elements of $\mathscr{H}_{(q,0)}$. Then the resulting interaction is *different* from that in the usual QED: Effectively, the interaction Hamiltonian has an additional term $q\mathbf{J}_a t^{\,a}$, where $\mathbf{J}_a$ is the usual current involving the quantized Dirac fields. Thus, the amplitudes for Compton scattering in Born approximation, for example, get additional contributions involving potential scattering due to $q\mathring{F}_{ab}$. In other words, although the particles described by $\mathscr{H}_{(q,0)}$ have zero rest mass, they have an effective *longitudinal* mode! They are therefore physically distinguishable from photons. Thus, q and g serve as *new quantum numbers*—analogous to spin and mass in Minkowskian quantum theory—for labeling physical particles, quantum numbers with origin in space–time topology. Note that the classical field equations corresponding to all these particles are the same: The usual Maxwell equations. Thus, we have an exactly soluble *quantum* model which confirms the suggestion,[25] made from semi-classical considerations, that a wave equation admitting nontrivial classical vacua may, in the quantum regime, simultaneously describe a variety of distinct particles.

(4) What is the status of the Hawking effect for these new particles? Perhaps the easiest way to analyze this issue is in terms of the Feynman propagators.[26,27] Using the fact that this propagator is expressible in terms of the vacuum expectation values of the time-ordered products of field operators, and the definition of the representation map $\boldsymbol{\pi}_{(q,g)}$, it follows that

$$ {}_{(q,g)}G^{F}_{aba'b'}(x,x') = G^{F}_{aba'b'}(x,x') + (q\mathring{F}_{ab} + g^{*}\mathring{F}_{ab})(x) \cdot (q\mathring{F}_{ab} + g^{*}\mathring{F}_{ab})(x')\,, \tag{11} $$

where G^F is the usual Feynman propagator for photons and ${}_{(q,g)}G^F$, that for the particles with quantum numbers (q,g).[28] (Here, as in the derivations of the usual Hawking effect in the Schwarzschild–Kruskal space–time,[23,27] we use those boundary conditions on the past horizon—or, equivalently, that "past vacuum state"—which implicitly incorporate the collapsing situation in static Kruskal picture.) From the definition of $\mathring{F}_{ab}$ it now follows that the corrective term in Eq. (11) has the same periodicity in the imaginary time as G^F. Hence, it follows that an evaporating black hole must emit the new particles again with a thermal spectrum.

(5) The model raises a number of questions concerning nonperturbative effects in general. The nontrivial classical vacua considered here—and hence also the automorphisms which led to symmetry breaking—are intimately related to the (Maxwell analog of the) Einstein–Yang–Mills instanton of Duff and Madore.[29] Is this a generic feature? That is, is there a class of instantons whose presence always signals symmetry breaking at the *quantum* level? It should not be difficult to extend the model to linearized gravitational fields on the Schwarzschild–Kruskal background. There again, the nontrivial classical vacua, i.e., the zero frequency modes, would lead to inequivalent representations of the CCR, labels (q,g) being replaced by mass and angular momentum. Is this interplay between the zero frequency modes and inequivalent *quantum* vacua generic? How do the non-Fock representations affect the one loop contributions? For Maxwell fields in the Schwarzschild–Kruskal space–time, we can obtain additional non-Fock representations as follows: Consider the (distributional, C number) solution to Maxwell's equation $\hat{F}_{ab}$ which coincides with $q\mathring{F}_{ab} + g^{*}\mathring{F}_{ab}$ in region I and with $q'\mathring{F}_{ab} + g'^{*}\mathring{F}_{ab}$ in region II and exploit the fact that the automorphism on $\mathscr{A}$ induced by the transformation $\mathbf{F}_{ab} \to \mathbf{F}_{ab} + \hat{F}_{ab}$ fails to be unitarily implemented in the Fock representation. Since $\hat{F}_{ab}$ is Lie derived by $t^{\,a}$ only in the (open) regions I and II, in the resulting representation there is no state which is invariant under the action of time translations. Nonetheless, these representations may be of physical interest: They are analogous to the representations describing Bloch walls and dislocations in spin systems, the role of the Bloch wall being played by the horizon. Can one exploit one's experience in magnetism to predict new electromag-

netic phenomena of topological origin in the quantum domain? More generally, non-Fock representations feature prominently in the descriptions of phase transitions in statistical mechanics. Can the representations of topological origin give us new insight into the "critical phenomena" that presumably occur at Planck length due to the quantum fluctuations of the gravitational field?

ACKNOWLEDGMENTS

We wish to thank John Friedmann, John Madore, Anne Magnon-Ashtekar, Rafael Sorkin, and Walter Thirring for discussions, and an anonymous referee for suggesting improvements in the presentation of the material. One of us (A.A) thanks the Relativity group in Munich for hospitality.

[1]In the context of nonlinear fields, Segal pointed out, as early as 1960, that the space–time topology may play an important role quantum theory. I.E. Segal, J. Math. Phys. **1**, 474 (1960).

[2]We are grateful to Kamesh Wali for suggesting this example.

[3]R.F. Streater, Proc. R. Soc. (London) Ser. A **287**, 510 (1965).

[4]The transformation $F_{ab} \to F_{ab} + K_{ab}$ with K_{ab} a constant skew tensor field is not permissible since, for example, the total energy associated with K_{ab} is infinite.

[5]C.W. Misner and J.A. Wheeler, Ann. Phys. (N.Y.) **2**, 525 (1962).

[6]Since the observables associated with the scalar field always involve derivatives, one might regard ϕ as a potential and $\nabla_a\phi$ as the field. Streater himself adopts this viewpoint in his article in *Recent Developments in Mathematical Physics*, edited by P. Urban (Springer, Wien, New York, 1973). For the "field" $\nabla_a\phi$, Streater's model of Ref. 3 gives only *one* representation of the CCR; there is no vacuum degeneracy.

[7]R. Sorkin, J. Phys. A **12**, 403 (1979).

[8]Our conventions are the following: g_{ab} has signature $(-+++)$ and the curvature tensors are defined by $\nabla_{[a}\nabla_{b]}V_c = 1/2R_{abc}{}^dV_d$; $R_{ab} = R_{am}{}^{bm}$; and $R = R_a{}^a$.

[9]A. Komar, Phy. Rev. **113**, 934 (1959).

[10]See, for example, W.H. Press and S.A. Teukolsky, Astrophys. J. **185**, 649 (1973).

[11]See the Appendix in Ref. 7.

[12]A. Lichnerowicz, *Propagateurs et commutateurs en relativité générale* (IHES, Paris, 1961), No. 10 pp. 38,39.

[13]From the mathematical point of view, a more satisfactory alternative would be the algebra generated by the Weyl operators $\exp[i\mathbf{F}(t)]$, which are represented by bounded operators on the Hilbert space of quantum states. However, their use makes the discussion of symmetry breaking more cumbersome and the basic ideas underlying the model less transparent.

[14]Foures-Bruhat, C.R. Acad. Sci. **251**, 29 (1960).

[15]Proof: For all F in Γ, we have $(q\mathring{F} + g^*\mathring{F}, F) = 0$ by a direct calculation involving an integration by parts. Since Γ is dense in $\bar{\Gamma}$ it follows that the static solutions are orthogonal to $\bar{\Gamma}$ in the Hilbert space of all solutions with finite energy.

[16]In fact, using Leray's results on evolution of initial data, one can show that $(T, D(T) = \Gamma)$ is essentially self-adjoint on $\bar{\Gamma}$.

[17]A. Ashtekar and A. Sen, "Quantum Fields in Curved Space–times: Selection and Uniqueness of the Vacuum State," preprint.

[18]M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Vol. I* (Academic, New York, 1972), p. 253.

[19]Reference 18, pp. 297, 298.

[20]For all F in Γ, set $E_a = F_{ab}n^b$ and $B_a = {}^*F_{ab}n^b$, where n^a is the unit future-directed normal to a given static slice S. By Maxwell's equations, $D_aE^a = D_aB^a = 0$, where D is the natural derivative operator on S. Hence $\bar{\Gamma} = \overline{\delta\mathscr{D}_2} \oplus \overline{\delta\mathscr{D}_2}$, where $\mathscr{D}_2$ is the space of all C^∞ 2 forms of compact support on S, δ the cocurl operator, and $\overline{\delta\mathscr{D}_2}$ the Cauchy completion of $\delta\mathscr{D}_2$ in $L_2(S,\lambda^{-1/2}\,dV)$, with dV the natural volume element on S. By Kodiara decomposition theorem, we have $L_2(S,\lambda^{1/2}\,dV) = d\mathscr{D}_0 \oplus \overline{\delta\mathscr{D}_2} \oplus h$, where h is the one-dimensional vector space of harmonic forms on S. A direct calculation shows that for all v in $\delta\mathscr{D}_2$, $\epsilon\cdot v$ is orthogonal to $d\mathscr{D}_0 \oplus h$. Hence, $\epsilon\cdot v$ is $\overline{\delta\mathscr{D}_2}$, whence the result.

[21]It is easy to verify that $R(\bar{T})$ is left invariant by J and that $\Gamma \subset H$.

[22]W. Israel, Phys. Lett. A **57**, 107 (1975).

[23]W.H. Unruh, Phys. Rev. D **14**, 870 (1976). We are refering to Unruh's "ξ vacuum."

[24]I.E. Segal, Ill. J. Math. **6**, 500 (1962).

[25]See, for example, R. Jackiw, Rev. Mod. Phys. **49**, 681 (1977).

[26]J. Hartle and S.W. Hawking, Phys. Rev. D **13**, 2188 (1976).

[27]G.W. Gibbons and M.J. Perry, Proc. R. Soc. (London) Ser. A **358**, 467 (1978).

[28]The propagators ${}_{(q,g)}G^F$ satisfy the same differential equation and the same boundary conditions as the usual propagator G^F, the only difference being that if q or g fails to vanish, ${}_{(q,g)}G^F$ goes to zero more slowly as x tends to infinity, x' being fixed, with x and x' on the same static slice in the Euclidean space. Thus, with the nontrivial topologies, one has to be careful in formulating uniqueness statements for Feynman propagators.

[29]M. Duff and J. Madore, Phys. Rev. D **18**, 2788 (1978).

Analytic properties of the critical dynamics ϵ expansion

Jayanta K. Bhattacharjee and Richard A. Ferrell

Institute for Physical Science and Technology and Center for Theoretical Physics, Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742

(Received 24 July 1979; accepted for publication 15 October 1979)

The ϵ-expansion technique for evaluating certain integrals in critical dynamics is studied in detail. The analytic properties in the complex ϵ plane are exploited to get improved answers at the physical value of ϵ. A different scheme—the $\bar{\epsilon}$ expansion—is suggested. This has a simpler analytic structure and thus gives more reliable results.

I. INTRODUCTION

In this note we discuss analytic properties and different methods of perturbative evaluation of a certain class of integrals which appear in the critical dynamics[1] of various physical systems such as liquid helium,[2-4] the antiferromagnet,[5] the Heisenberg ferromagnet,[6,7] etc. The task of critical dynamics is to calculate the wave number and frequency-dependent relaxation rates of the order parameter [an n-vector field with O(n) symmetry] of the transition as well as of any secondary variable that may couple to it through the equations of motion. It is generally convenient to single out for study a particular Fourier component of the fluctuating variable under consideration. The problem of critical dynamics is nonlinear in that each Fourier component is influenced by all the other components via a trilinear coupling. Consequently, to obtain the relaxation rate of a particular component, it is necessary to perform a sum over all the Fourier components. In other words, we have to carry out a sum in wave number or "momentum" space. (We use the terms interchangeably.) The intermediate modes are more effective the more slowly they damp out. Consequently, the corresponding relaxation rates appear in the denominators of the sums that have to be effected. The relaxation rates satisfy certain dynamic scaling laws,[8,9] requiring them to depend upon momentum according to some critical exponent, which does not in general have an integer value. It is this feature which makes it impossible to evaluate the critical dynamics integrals exactly.

Here we shall specifically discuss the integrals that have to be evaluated to obtain the order-parametr (ψ) and the entropy (S) relaxation rates in liquid Helium.[4] To one-loop order in the language of Feynman graphs, these are

$$I_\psi = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^{D/2}} \tag{1.1}$$

and

$$I_S = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2}\frac{(p^2-p'^2)^2}{p^{D/2}+p'^{D/2}}, \tag{1.2}$$

where momentum conservation requires $\mathbf{p}+\mathbf{p}'=\mathbf{k}$. The external momentum is scaled to unity, $|\mathbf{k}|=1$. The area of the unit sphere in D-dimensional space is

$$C_D = \frac{2\pi^{D/2}}{\Gamma(D/2)}. \tag{1.3}$$

Integrals such as I_ψ and I_S cannot in general be evaluated analytically at $D=3$, the physical dimension. I_ψ is an exception and can be evaluated at arbitrary dimensionality D. It therefore provides a good testing ground for the different approximation schemes to be discussed below. As far as getting a number for the integral at $D=3$ is concerned, it might be argued that numerical integration techniques are sufficient. However, these are multidimensional integrals, which get progressively more complicated as higher loops are considered. They are consequently not very suitable for numerical work. The occurrence of singularities in the integrands is a further hindrance to a numerical integration. As these are integrable singularities, they do not pose any difficulty to an analytic approach. It is therefore more attractive to study the required integrals as continuous functions of a suitably defined small paramete ϵ, related to D. Expanding in powers of ϵ gives then a systematic means of obtaining approximate values for the integrals. Consequently, from now on, we shall refer to the integrals as the function $I_\psi(\epsilon)$ and $I_S(\epsilon)$. The aim of the work will thus be to develop a kind of perturbative evaluation of the integrals in the form of a small number of terms.

It needs to be emphasized that there are two sources of inaccuracy in applying the ϵ-expansion to three-dimensional ($\epsilon=1$) critical dynamics problems. The first is the error incurred by using a truncated series expansion for the integrals involved. The purpose of the present paper is to study the nature of this error and to show how it can be relatively easily avoided. The second source of error is the truncation of the loop expansion itself. We have reason to think that under some circumstances this error is not as serious as the first. This difficult question requires, however, much further study and will not be addressed here.

The conventional means of attacking the integrals in Eqs. (1.1) and (1.2) is the so-called "ϵ-expansion."[2,10] This is analogous to the dimensional regularization in quantum field theory,[11] where renormalization factors are calculated by evaluating the loop integrals as Laurent series in ϵ and extracting the coefficient of the $1/\epsilon$ term. The ϵ expansion exploits the ultraviolet divergence of these integrals at $D=4$ and therefore defines the expansion parameter as $\epsilon = 4-D$. Until now, only a two-term expansion has been achieved. The suitability of such an expansion and some ways of im-

[a]This work constitutes a portion of the dissertation of J.K. Bhattacharjee submitted to the faculty of the University of Maryland in partial fulfillment of the requirements for the Ph.D. degree (1979).

 0022-2488/80/030534-09$1.00

proving it will be the subject of the following sections. In Sec. II we present the ϵ-expansion results for the integrals and compare them at $\epsilon = 1$ with the exact answers for $I_\psi(1)$ and $I_S(1)$. The accuracy of the two-term expansion is found to be poor. This shortcoming can be traced to a special feature of the analytic properties of Eqs. (1.1) and (1.2), which reveal an infrared singularity at $D = 2$, or $\epsilon = 2$. The neglect of this singularity is a gross omission which impairs the usefulness of the ϵ-expansion as far away as $\epsilon = 1$. We remedy this deficiency by means of a double expansion based on the simple poles of the functions which correspond to their ultraviolet and infrared singularities. This improvement is carried further in Sec. III, by taking into account more of the poles. In Sec. IV we present an alternative expansion scheme for evaluating such integrals a $D = 3$. This scheme, the "$\bar{\epsilon}$-expansion," manages to remove or, at least, delay the onset of the infrared singularity. It thus provides a two-term expansion based on the ultraviolet divergence which yields more accurate answers. Section V deals with further applications while Sec. VI constitutes a brief summary.

II. THE ϵ-EXPANSION AND THE INFRARED SINGULARITY

We shall first treat the integral I_ψ. The ultraviolet divergence of Eq. (1.1) can be extracted as follows,

$$\begin{aligned} I_\psi(\epsilon) &= \frac{1}{C_D}\int_{p>1}\frac{d^Dp}{p^2\cdot p^{D/2}} + I_v(\epsilon) \\ &= \int_{p>1}\frac{dp}{p^{3-D/2}} + I_v(\epsilon) \\ &= 2/\epsilon + I_v(\epsilon) \sim 2/\epsilon, \end{aligned} \tag{2.1}$$

where the ultraviolet-regularized integral

$$I_v(\epsilon) = \frac{1}{C_D}\int\frac{d^Dp}{p^2}\left(\frac{1}{p'^{D/2}} - \frac{\theta(p-1)}{p^{D/2}}\right) \tag{2.2}$$

remains finite and is therefore asymptotically negligible as $\epsilon \to 0$. $\theta(x)$ is the step function

$$\theta(x) = \begin{cases} 0, & x<0 \\ 1, & x>0. \end{cases} \tag{2.3}$$

The first correction to the asymptotic formula of Eq. (2.1) is found by evaluating Eq. (2.2) at $\epsilon = 0$. This yields $I_v(0) = \frac{1}{2}$, giving a two-term expansion which at $\epsilon = 1$ becomes

$$I_\psi(1) = [2/\epsilon + \tfrac{1}{2} + O(\epsilon)]_{\epsilon=0} = 2.5 + O(1). \tag{2.4}$$

On the other hand, for $2 < D < 4$, I_ψ can be exactly evaluated to yield

$$I_\psi(\epsilon) = \frac{2}{\epsilon}\cdot\frac{\Gamma(2-\epsilon/2)\Gamma(1-\epsilon/2)\Gamma(1+\epsilon/4)}{\Gamma(2-3\epsilon/4)}. \tag{2.5}$$

Expanding the gamma function to the lowest order in ϵ yields Eq. (2.4) once more. By means of Eq. (2.5) we can compare Eq. (2.4) with the exact value

$$I_\psi(1) = \pi = 3.14\cdots. \tag{2.6}$$

Equation (2.4) is thus a significant underestimate of the integral.

To understand this difference we can study Eq. (2.5) for the analytic properties of the integral. We note that at $\epsilon = 2$ there is also a singularity. Looking back at Eq. (1.1) we can identify this as the infrared singularity. The infrared divergence can be extracted by the subtraction

$$I_\psi(\epsilon) = \frac{1}{C_D}\int\frac{d_p^D}{p^2} + I_r(\epsilon), \tag{2.7}$$

where the infrared regularized integral is

$$I_r(\epsilon) = \frac{1}{C_D}\int\frac{d^Dp}{p^2}\left(\frac{1}{p'^{D/2}} - \theta(p-1)\right). \tag{2.8}$$

Performing the angular averages, we find for $D = \epsilon = 2$

$$\begin{aligned} I_r(2) &= \frac{2}{\pi}\int_0^\infty\frac{dp}{p(1+p)}\left[K\left(\frac{2\sqrt{p}}{1+p}\right) - \frac{\pi}{2}\right] \\ &= \sum_{n=1}^{\infty}\frac{1}{2^{2n}}\frac{(2n)!}{(n!)^2}\cdot\frac{1}{n}. \end{aligned} \tag{2.9}$$

$K(x)$ is the elliptic integral of the first kind. To sum the series, we substitute

$$\frac{1}{n} = \int_0^1 dt\cdot t^{n-1} \tag{2.10}$$

into Eq. (2.9) and interchange the order of summation and integration to obtain

$$\begin{aligned} I_r(2) &= \int_0^1 dt\sum_{n=1}^{\infty}\frac{1}{2^{2n}}\frac{(2n)!}{(n!)^2}t^{n-1} \\ &= \int_0^1\frac{dt}{t}\sum_{n=1}^{\infty}\binom{-\frac{1}{2}}{n}(-t)^n \\ &= \int_0^1\frac{dt}{t}[-1+(1-t)^{-1/2}] = 2\ln 2 \simeq 1.38. \end{aligned} \tag{2.11}$$

It is instructive to obtain this answer in another form. Instead of evaluating the integral in the last step of Eq. (2.11), we can express it as

$$\begin{aligned} I_r(2) &= \lim_{a\to 1}\left[-\int_0^1 dt\,t^{-a} + \int_0^1 dt\,t^{-a}(1-t)^{-1/2}\right] \\ &= \lim_{a\to 1}[-\beta(1-a,1) + \beta(1-a,\tfrac{1}{2})] \\ &= \lim_{a\to 1}\left(-\frac{\Gamma(1-a)\Gamma(1)}{\Gamma(2-a)} + \frac{\Gamma(1-a)\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})-a}\right) \\ &= \lim_{\delta\to 0}\Gamma(\delta)\left(\frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}+\delta)} - \frac{\Gamma(1)}{\Gamma(1+\delta)}\right) \\ &= \psi(1) - \psi(\tfrac{1}{2}), \end{aligned} \tag{2.12}$$

where $\psi(x)$ is the usual digamma function defined as $\psi(x) = (d/dx)\ln\Gamma(x)$. This is the same as the answer in Eq. (2.11) since $\psi(1) = -\gamma$ and $\psi(\frac{1}{2}) = -\gamma - 2\ln 2$, where γ is Euler's constant.

Thus, expansion about the infrared singularity yields

$$I_\psi(\epsilon) = 1/(D-2) + 1.38 + O(D-2). \tag{2.13}$$

Equation (2.13) could also have been obtained from Eq. (2.5) by expanding in the variable $D-2$. This leads to

$$\begin{aligned} I_\psi(\epsilon) = \frac{1}{D-2} &+ \tfrac{1}{2}[1 + 2\psi(1) - \tfrac{1}{2}\psi(\tfrac{3}{2}) - \tfrac{3}{2}\psi(\tfrac{1}{2})] \\ &+ O(D-2), \end{aligned} \tag{2.14}$$

where $\psi(x)$ is the digamma function. Using the identity

$$\psi(\tfrac{3}{2}) - \psi(\tfrac{1}{2}) = 2, \tag{2.15}$$

we can put Eq. (2.14) in the form

$$I_\psi(\epsilon) = \frac{1}{D-2} + \psi(1) - \psi(\tfrac{1}{2}) + O(D-2), \tag{2.16}$$

the same result as before.

We now use the information contained in Eqs. (2.4) and (2.13) to obtain an improved representation of the function $I_\psi(\epsilon)$ in the range $0 \leqslant \epsilon \leqslant 2$. $I_\psi(\epsilon)$ has no other singularities in this region apart from the ones discussed above and therefore, when the poles are subtracted, we get a function

$$F_\psi(\epsilon) = I_\psi(\epsilon) - 2/\epsilon - 1/(D-2), \tag{2.17}$$

which is completely analytic in the region. Consequently it can be represented as a power series in ϵ,

$$F_\psi(\epsilon) = \sum_{n=0}^{\infty} c_n \epsilon^n. \tag{2.18}$$

The information regarding the correction to the leading singularities is sufficient to fix the first two coefficients c_0 and c_1. From Eqs. (2.14) and (2.3), $F_\psi(\epsilon) = 0$ at $\epsilon = 0$. Hence $c_0 = 0$. From Eqs. (2.14) and (2.11), $F_\psi(0) = 0.38$ at $\epsilon = 2$. This gives $c_1 = 0.19$. The resulting expression for $I_\psi(\epsilon)$ is

$$I_\psi(\epsilon) = 2/\epsilon + 1/(2-\epsilon) + 0.19\epsilon, \tag{2.19}$$

which, evaluated at $\epsilon = 1$, yields

$$I_\epsilon(1) = 3.19, \tag{2.20}$$

a considerable improvement over the previous estimate.

We now turn to the integral I_S defined by Eq. (1.2). The ultraviolet divergence occurs at $D = 4$ according to

$$\begin{aligned} I_S(\epsilon) &= \frac{2}{DC_D}\int_{p\geqslant 1} \frac{d^D p}{p^{2+D/2}} + J_v(\epsilon) \\ &= 4/D\epsilon + J_v(\epsilon) \sim 1/\epsilon, \end{aligned} \tag{2.21}$$

where the ultraviolet regularized integral is

$$J_v(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2}\left(\frac{(p^2-p'^2)^2}{p^{D/2}+p'^{D/2}} - \frac{2}{D}\frac{p'^2}{p^{D/2}}\theta(p-1)\right), \tag{2.22}$$

and is asymptotically negligible in the limit $\epsilon \to 0$. Evaluated at this point,[4]

$$J_v(0) = \tfrac{1}{2}\ln 2 - \tfrac{1}{8}, \tag{2.23}$$

which yields the two-term expansion

$$\begin{aligned} I_S(\epsilon) &= 1/\epsilon + \tfrac{1}{4} + J_v(0) + O(\epsilon) \\ &= 1/\epsilon + 0.47 + O(\epsilon). \end{aligned} \tag{2.24}$$

At $\epsilon = 1$ this becomes $I_S(1) = 1.47$, as compared to the value 2.25 ± 0.02 obtained from numerical integration. Once again this severe underestimate results from the neglect of the infrared divergence at $D = 2$. $I_S(\epsilon)$ is analytic for $0 \leqslant \epsilon \leqslant 2$ except for the ultraviolet and infrared singularities at $\epsilon = 0$ and 2, respectively.

To study the behavior of $I_S(\epsilon)$ near $\epsilon = 2$, we write Eq. (1.4) as

$$I_S(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2} + J_r(\epsilon), \tag{2.25}$$

where the infrared regularized integral is

$$J_r(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2}\left[\frac{(p^2-p'^2)^2}{p^{D/2}+p'^{D/2}} - 1\right]. \tag{2.26}$$

The infrared singularity is contained in the first integral on the right-hand side of Eq. (2.25). For a two-term expansion about this singularity, we need

$$\begin{aligned} J_r(2) &= \frac{1}{C_2}\int \frac{d^2 p}{p^2 p'^2}\left(\frac{(p^2-p'^2)^2}{p+p'} - 1\right) \\ &= \frac{1}{\pi}\int_0^\infty \frac{dp}{p}\int_0^\pi \frac{d\theta}{p'^2}\left(\frac{(p^2-p'^2)^2}{p+p'} - 1\right). \end{aligned} \tag{2.27}$$

Noting that

$$p\, d\theta\, dp = \frac{p'\, dp'\, dp}{\sin\theta} = \frac{2p\, p'\, dp\, dp'}{\sqrt{4p^2 - (1+p^2-p'^2)^2}}, \tag{2.28}$$

and introducing the elliptic coordinates $x = p + p'$ and $y = p - p'$ with $1 \leqslant x \leqslant \infty$ and $-1 \leqslant y \leqslant 1$, we find

$$\begin{aligned} J_r(2) &= \frac{4}{\pi}\int_1^\infty dx \int_{-1}^{1} dy \frac{(xy^2-1)}{(x^2-y^2)[(x^2-1)(1-y^2)\,]^{1/2}} \\ &= 4(\ln 2 - 1) \\ &\simeq -1.20. \end{aligned} \tag{2.29}$$

We also note that

$$\begin{aligned} &\frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2} \\ &\quad = \frac{[\Gamma(D/2-1)\,]^2 \Gamma(2-D/2)\Gamma(D/2)}{2\Gamma(D-2)} \\ &\quad = 2/(D-2) + O(D-2). \end{aligned} \tag{2.30}$$

[The $O(1)$ term is easily seen to vanish.] Using Eqs. (2.25), (2.30), and (2.29), we obtain the infrared approximant

$$I_S(\epsilon) = 2/(2-\epsilon) - 1.20 + O(2-\epsilon). \tag{2.31}$$

We can now define the analytic function

$$F_S(\epsilon) = I_S(\epsilon) - 1/\epsilon - 2/(2-\epsilon), \tag{2.32}$$

and obtain, by arguments similar to the case for $F_\psi(\epsilon)$

$$F_S(\epsilon) = -0.53 - 0.58\epsilon + O(\epsilon^2). \tag{2.33}$$

Thus the four-term approximant for $I_S(\epsilon)$ is

$$I_S(\epsilon) = 1/\epsilon - 2/(2-\epsilon) - 0.53 - 0.58\epsilon, \tag{2.34}$$

yielding

$$I_S(1) = 1.89, \tag{2.35}$$

a considerable improvement on the two-term result.

III. FURTHER SINGULARITIES

In the preceding section we confined our attention to the physical region $0 \leqslant \epsilon \leqslant 2$, bounded by the ultraviolet pole at $\epsilon = 0$ and the infrared pole at $\epsilon = 2$. We now consider ϵ to be a complex variable and analytically continue the functions $I_\psi(\epsilon)$ and $I_S(\epsilon)$ beyond the "physical" region. This is readily accomplished by means of the regularized integrals $I_{u,r}$ and $J_{u,r}$, which are finite at the "physical" poles. What we find are more singularities: In addition to the "physical" ones at $\epsilon = 0$ and 2, there are both high-momentum and low-momentum singularities in the integrals for other values of ϵ. We call these the "unphysical" ultraviolet and infrared poles. A general representation of a function of this sort is a Mittag-Leffler[12] expansion, which is a sum over the principal parts associated with the various poles plus an entire

function. In this section we improve upon our work in Sec. II by means of a truncated Mittag–Leffler expansion based upon the two physical poles and the two nearest unphysical poles—one ultraviolet and the other infrared.

Before illustrating the procedure in detail, it is useful to estimate the improvement that can be expected by including the unphysical poles. The contribution of one of these, at, say $\epsilon = \epsilon_1$, to the Taylor's series at $\epsilon = 0$ is of the form

$$\frac{a_i}{\epsilon - \epsilon_i} = -\frac{a_i}{\epsilon_i} + \frac{a_i}{\epsilon_i^2}\epsilon - \frac{a_i}{\epsilon_i^3}\epsilon^2 + \cdots. \tag{3.1}$$

Now the first two terms are effectively taken into account by the linear fitting procedure of Sec. II, which fixes the first two coefficients of an expansion such as Eq. (3.1). This is done by means of the two pieces of information provided by the regularized integrals evaluated at the two physical poles. It is the third and higher terms of Eq. (3.1) which are neglected in this procedure. The error incurred in the neglect of such terms can therefore be expected *a priori* to fall off as the inverse third power of the distance of the nearest unphysical pole not included in the truncated Mittag–Leffler expansion. For example, assume for definiteness that the unphysical poles occur at roughly equal intervals along the real ϵ-axis (not always true, as will be seen below). Then including the nearest ones in the Mittag–Leffler expansion can be expected to reduce the error by roughly 2^{-3}, or one order of magnitude.

We demonstrate this approach first for I_ψ. The subtraction in the integrand of Eq. (2.2) depends upon the difference between p' and p. This difference is small for $p \gg 1$ and after angle averaging introduces into the integrand an additional factor proportional to p^{-2}. The denominator of the integrand is therefore proportional to $p^{4+D/2}$, indicating an ultraviolet divergence at $D = 4 + D/2$, or $D = 8$. More precisely, in the vicinity of this unphysical pole we find for the square brackets of the integrand

$$\begin{aligned}
&\frac{1}{p'^{D/2}} - \frac{1}{p^{D/2}} \\
&\quad \approx \frac{1}{p'^4} - \frac{1}{p^4} = \frac{1}{(p^2+1-2\mu p)^2} - \frac{1}{p^4} \\
&\quad = \frac{1}{p^4}\left[\left(1 + \frac{1}{p^2} - \frac{2\mu}{p}\right)^{-2} - 1\right] \\
&\quad = \frac{1}{p^4}\left[-2\left(\frac{1}{p^2} - \frac{2\mu}{p}\right) + 12\,\frac{\mu^2}{p^2} + \cdots\right] \\
&\Rightarrow -\frac{1}{2}\frac{1}{p^6} \approx -\frac{1}{2}\frac{1}{p^{2+D/2}},
\end{aligned} \tag{3.2}$$

where μ is the cosine of the angle between $\mathbf{p}$ and $\mathbf{k}$ and the last line results from angle averaging. Thus the factor of p^{-2} referred to above carries a numerical coefficient of $-\frac{1}{2}$. Substituting Eq. (3.2) into Eq. (2.2) gives the asymptotic formula

$$\begin{aligned}
I_v(\epsilon) &\sim -\frac{1}{2}\int_{p\gg 1} p^{D/2-5}\,dp = -\frac{1}{2}\frac{1}{4-D/2} \\
&= -1/(8-D) = -1/(\epsilon+4),
\end{aligned} \tag{3.3}$$

for $|\epsilon + 4| \ll 1$. In other words, the principal part associated with the unphysical ultraviolet pole at $\epsilon = -4$ is the right-hand member of Eq. (3.3).

As the computation for the unphysical infrared pole follows similar lines, we sketch here only the salient points. First we observe that it is convenient not to include with $I_\psi(\epsilon)$ the prefactor C_D^{-1} because of the zeros of C_D in the infrared range (at D equal to the negative even integers). Thus we will be dealing with $C_D I_\psi(\epsilon)$, and instead of angle averages we will need the solid angle integrals

$$\int d^D\Omega \mu^l = 2\pi^{(D-1)/2}\Gamma\left(\frac{l+1}{2}\right)\Gamma\left(\frac{D+l}{2}\right), \tag{3.4}$$

for l equal to the positive even integers. Equations (3.4) reduces to Eq. (1.3) for $l = 0$ and vanishes, by symmetry, for l equal to an odd integer. As before, we expand the factor $p'^{D/2} - 1$ of the integrand of $I_r(\epsilon)$ in powers p, where now $p \ll 1$. After integrating over angle, only even powers of p remain, so the unphysical infrared poles can be expected to occur at $D = 0$, -2, -4, etc. $p^{-D/2}|_{D=0} - 1$, however, vanishes identically. Therefore, $D = 0$ is not actually a pole and $D = -2$ is the location of the closest unphysical pole. In this case Eq. (3.4) vanishes for $l = 0$ and 2. For $l = 4$ it yields

$$\int d^{-2}\Omega\,\mu^4 = 3/2\pi. \tag{3.5}$$

Therefore, regularizing at the physical infrared pole, we find

$$\begin{aligned}
&\int d^{-2}\Omega\,(p'^{-D/2}|_{D=-2} - 1) \\
&\quad = \int d^{-2}\Omega\,[(1-2\mu p + p^2)^{1/2} - 1\ ] \\
&\quad = \int d^{-2}\Omega \sum_{\nu=1}^{\infty}\binom{\frac{1}{2}}{\nu}(-2\mu p + p^2)^\nu \\
&\quad = \binom{\frac{1}{2}}{\nu}(2p)^4\int d^{-2}\Omega\,\mu^4 + \cdots \\
&\quad \approx -\frac{15}{16\pi}p^4,
\end{aligned} \tag{3.6}$$

where we have substituted Eq. (3.5). The neglected terms are of higher order in p^2. The terms of lower order than p^4 vanish by virtue of Eq. (3.4) for $D = -2$. For $D \neq -2$ the p^2 term is removed by an additional regularization so that the radial integration is asymptotically

$$\int dp\,p^{D+1} \sim \frac{1}{D+2} = \frac{1}{6-\epsilon}. \tag{3.7}$$

Thus

$$C_D I_r \sim \frac{15}{16\pi}\frac{1}{\epsilon - 6} \tag{3.8}$$

is the principal part associated with the unphysical infrared pole at $\epsilon = 6$.

From Eq. (3.3) the unphysical ultraviolet pole in $C_D I_\psi$ has strength $|-C_D| = C_8 = \pi^4/3$ or $16\pi^5/45 \approx 109$ times greater than the coefficient in Eq. (3.8). This disparity results from the skewing effect of the numerator of Eq. (1.3), i.e., $2\pi^{D/2}$, which accounts for a ratio of π^5. Such skewing is undesirable for interpolation purposes. Therefore we remove this factor and study $(C_D/2\pi^{D/2})I_\psi(\epsilon)$ $= \Gamma^{-1}(D/2)I_\psi(\epsilon)$, which has the more natural ratio of

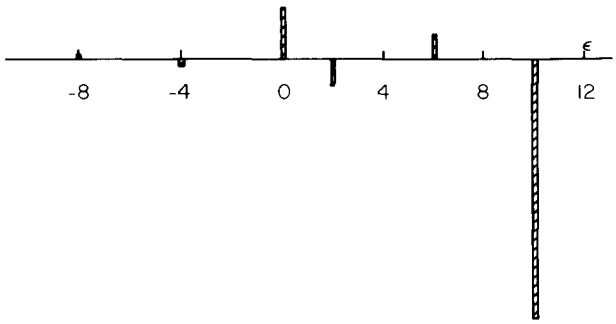

FIG. 1. The positions ϵ_i and strengths a_i of the first few poles of $I_\psi(\epsilon)/\Gamma(D/2)$ vs ϵ. The principle parts associated with the poles are of the form $a_i/(\epsilon-\epsilon_i)$. The strength at $\epsilon=0$ is unity. The others are drawn according to this scale.

$16/45 \approx 1/3$ for its unphysical pole strengths. From Eqs. (3.8), (3.3), (2.1), and (2.13) we have a four-pole truncated Mittag–Leffler expansion

$$\frac{I_\psi(\epsilon)}{\Gamma(D/2)} = \frac{2}{\epsilon}\frac{1}{\epsilon-2} - \frac{1}{6}\frac{1}{\epsilon+4} + \frac{15}{32}\frac{1}{\epsilon-6} + F_\psi(\epsilon), \tag{3.9}$$

where $F_\psi(\epsilon)$ is analytic in a very wide interval (wider than $-4 \leqslant \epsilon \leqslant 6$). It is easy to verify Eq. (3.9) from the exact expression of Eq. (2.5), which shows that $I_\psi(\epsilon)/\Gamma(D/2)$ is a meromorphic function with further ultraviolet poles at twice the negative even integers and further infrared poles at twice the positive odd integers.

From the information that we have about $I_\psi(\epsilon)/\Gamma(D/2)$ it is possible to determine the first two coefficients in the Taylor's series $F_\psi(\epsilon) = C_0 + C_1\epsilon + \cdots$. This information consists of corrections to the asymptotic form near $\epsilon = 0$ and 2. From Eq. (2.4) we see that

$$I_\psi(\epsilon)/\Gamma(D/2) = 2/\epsilon + \tfrac{3}{2} + \psi(1) + O(\epsilon), \tag{3.10}$$

for $\epsilon \simeq 0$, and from Eq. (2.13)

$$I_\psi(\epsilon)/\Gamma(D/2) = 1/(2-\epsilon) + 2\ \ln 2 - \tfrac{1}{2}\psi(1) + O(2-\epsilon), \tag{3.11}$$

for $\epsilon \simeq 2$. From Eqs. (3.9)–(3.11), we find

$$C_0 = \psi(1) + \tfrac{215}{192} \simeq 0.543 \tag{3.12}$$

and

$$C_1 = \ln 2 - \tfrac{3}{4}\psi(1) - \tfrac{2275}{2304} \simeq 0.138. \tag{3.13}$$

Thus

$$\frac{I_\psi(\epsilon)}{\Gamma(D/2)} \frac{2}{\epsilon} - \frac{1}{\epsilon-2} + \frac{1}{6}\frac{1}{\epsilon+4} + \frac{15}{32}\frac{1}{\epsilon-6} + 0.543 + 0.138. \tag{3.14}$$

At $\epsilon = 1$, Eq. (3.14) gives $I_\psi(1)/\Gamma(3.2) = 3.554$.

The two-pole approximation of Sec. II, when applied to $I_\psi(\epsilon)/\Gamma(D/2)$, gives 3.548 at $\epsilon = 1$. Thus the effect of introducing the two unphysical poles is to produce a mere 0.15% effect. This could have been anticipated on the basis of Eq. (3.1). The main effect of introducing the unphysical pole at $\epsilon = \epsilon_i$ is to produce the curvature $-2\,a_i/\epsilon_i^3$. If $|a_i| \approx 1$ and $|\epsilon_i| \gg 1$, the effect is small. In the present case the curvature from the two unphysical poles is

$$R_2 = -2\,\frac{15}{32}\frac{1}{6^3} - 2\,\frac{1}{6}\frac{1}{4^3} = -0.01. \tag{3.15}$$

In general the change in $F_\psi(\epsilon)$ resulting from curvature is

$$\Delta F_\psi(\epsilon) = -\tfrac{1}{2}R\epsilon(2-\epsilon), \tag{3.16}$$

vanishing at $\epsilon = 0$ and $\epsilon = 2$ as required by Eqs. (2.4) and (2.13). At $\epsilon = 1$ the numerical change in $I_\psi(1)$ is

$$\Delta E_\psi(1) = -R/2 = 0.005, \tag{3.17}$$

or a 0.15% increase.

Although numerically Eq. (3.17) yields a small effect, it clearly illustrates the role and influence of the unphysical poles. The more remote ones will have even less influence, provided their strengths remain $|a_i| \approx 1$. The present example is an exception, as seen from Table I and Fig. 1. The next unphysical infrared pole at $\epsilon = 10$ has the relatively large strength of $-10395/2048 = -5.08$, so that its contribution to the curvature is $-2(-5.08)\times 10^{-3} = 0.01$, exactly canceling Eq. (3.15). The remaining poles give similar but smaller contributions, converging to a net curvature $R_\infty = +0.006$ from the infinite set of unphysical poles. The small value of R_∞ is the reason that the two-pole approximation works so well. This will not be true in general. In particular, if the unphysical poles are close to the physical region and if their strengths are large, then inclusion of the physical poles alone will not be sufficient. The introduction of the first two unphysical poles will then produce a substantial improvement, as will be seen for $I_S(\epsilon)$, to which we now turn.

In studying $I_S(\epsilon)$ it is not necessary to divide out by $\Gamma(D/2)$. This is because the first unphysical infrared pole occurs before $D = 0$ is reached (where C_D has its first zero). To find the poles and principal parts of $I_S(\epsilon)$, we analytically continue Eq. (2.22) into the ultraviolet region $\epsilon < 0$ until we encounter the divergence of $J_v(\epsilon)$ at $\epsilon = -4$. In this vicinity the intgral behaves as

$$J_v(\epsilon) = \frac{2}{C_D}\int_{p\geqslant 1} \frac{d^D p}{p^{4+D/2}} \times\left(3\,\frac{4+3D/4}{D(D+2)} + \frac{1}{4} - 3\,\frac{1+D}{D}\right), \tag{3.18}$$

with principal part equal to $-2^{-1}(\epsilon+4)^{-1}$.

Going in the other direction, we extract the first unphysical infrared singularity from the regularized integral, Eq. (2.26). This integral does not have an infrared singularity

TABLE I. Principal parts of "physical and "unphysical" poles of $I_\psi/\Gamma(D/2)$ in the ultraviolet and infrared regions.

Singularity	"Physical"	"unphysical" $n \geqslant 1$
ultraviolet	$\frac{2}{\epsilon}$	$\frac{2(-1)^n\Gamma(1+2n)}{(\epsilon+4n)\Gamma(2+3n)\Gamma(1+n)}$
infrared	$-\frac{1}{\epsilon-2}$	$-\frac{\Gamma(n+\frac{1}{2})}{[\epsilon-2(2n+1)\ ]\Gamma(\frac{1}{2}-3n)\Gamma(1+2n)}$

until $D = 4/3$. It can easily be seen from the divergence at $p \approx 0$ and at $p' \approx 0$ that there is a simple pole of the integral at $\epsilon = 8/3$, with the principal part equal to $4/(3\epsilon - 8)$. We thus have the four-pole truncated Mittag–Leffler expansion

$$I_S(\epsilon) = \frac{1}{\epsilon} - \frac{2}{\epsilon - 2} - \frac{1}{2}\frac{1}{\epsilon + 4} + \frac{4}{3\epsilon - 8} + F_S(\epsilon), \tag{3.19}$$

where $F_S(\epsilon)$ is analytic over an extended interval (broader than $-4 \leqslant \epsilon \leqslant 8/3$). From Eqs. (2.24) and (2.31) we obtain the linear approximation

$$F_S(\epsilon) = 0.095 + 0.146\epsilon. \tag{3.20}$$

At $\epsilon = 1$, Eq. (3.20) substituted into Eq. (3.19) yields $I_{LS}(1) = 2.34$, overshooting the exact answer of 2.25 ± 0.02 and representing an increase of 0.45 over the two-pole approximation of Sec. II. This can be attributed almost entirely to the unphysical infrared pole, it being the stronger and closer of the two unphysical poles of Eq. (3.19). According to Eqs. (3.1) and (3.16) it should contribute $-2(4/3)(5/3)^{-3}$ to R_2, and therefore $-R_2/2 = 0.28$ to $I_S(1)$. (For this purpose we have evaluated the curvature at $\epsilon = 1$.)

Because of the nearness of the pole, the first terms of the Taylor's series for $\Delta F(\epsilon)$ are not sufficiently accurate and we need to use for the contribution of the ith pole

$$\Delta F_i(1) = a_i\left[\frac{1}{1 - \epsilon_i} - \frac{1}{2}\left(\frac{1}{-\epsilon_i} + \frac{1}{2 - \epsilon_i}\right)\right] = \frac{a_i}{\epsilon_i(1 - \epsilon_i)(2 - \epsilon_i)}, \tag{3.21}$$

wehich reduces to $a_i\epsilon_i^{-3}$ for $\epsilon_i \gg 1$. For the unphysical infrared pole, Eq. (3.21) gives $9/20 = 0.45$, in agreement with the above result for the total deviation. The contribution of the unphysical ultraviolet pole is less than one percent of this.

As anticipated at the beginning of this section, including the unphysical poles can produce a significant change in the value of the integral. Because of the sensitivity of the integral to the poles, on whose strength and location we necessarily have limited information, it is useful to consider an alternative scheme which minimizes the number of the poles. The $\bar{\epsilon}$-expansion of the next section is such a scheme.

IV. $\bar{\epsilon}$-EXPANSION

In this section a different expansion scheme is described for evaluating the integrals I_ψ and I_S at $D = 3$. We define the functions $\bar{I}_\psi(\bar{\epsilon})$ and $\bar{I}_S(\bar{\epsilon})$ as follows

$$\bar{I}_\psi(\bar{\epsilon}) = \frac{1}{C_3}\int \frac{d^3p}{p^2 p'^{1+\bar{\epsilon}}} \tag{4.1}$$

and

$$\bar{I}_S(\bar{\epsilon}) = \frac{1}{C_3}\int \frac{d^3p}{p^2 p'^2}\frac{(p^2 - p'^2)^2}{p^{1+\bar{\epsilon}} + p'^{1+\bar{\epsilon}}}. \tag{4.2}$$

Setting $\bar{\epsilon} = 1$ gives the $D = 3$ version of Eqs. (1.1) and (1.2). We note that $\bar{\epsilon} = 0$ is the source of ultraviolet divergence, just as $\epsilon = 0$ was in the case of the ϵ-expansion in Sec. II. We study the first corrections to this pole at $\bar{\epsilon} = 0$ and expect the two-term approximant to be better than the corresponding one in the ϵ-expansion. The reason is in the delay of the onset of the infrared divergence in Eq. (4.1) and its complete removal in Eq. (4.2). Equation (4.1) shows an infrared singularity at $\bar{\epsilon} = 2$, but the point of interest, $\bar{\epsilon} = 1$, is well removed from the divergence and consequently is relatively little affected by it. The integral $\bar{I}_S(\bar{\epsilon})$ has no infrared singularity at all and therefore should be an even better candidate for the two-term $\bar{\epsilon}$-expansion. The infrared singularity is an unnecessary complication in the evaluation of such integrals at the physical dimension $D = 3$. By removing it completely or by lessening its effect, we have succeeded in the task of reducing the number of poles.

Another advantage of this scheme should be noted. To obtain the correction to the asymptotic behavior at $\bar{\epsilon} = 0$, we need to study $\bar{I}_\psi(\bar{\epsilon})$ and $\bar{I}_S(\bar{\epsilon})$ at $\bar{\epsilon} = 0$. There is yet another value of $\bar{\epsilon}$ at which these integrals can be done analytically—namely $\bar{\epsilon} = 1$. This provides further information on the $\bar{\epsilon}$-expansion.

Treating $\bar{I}_\psi(\bar{\epsilon})$ first, we see that

$$\bar{I}_\psi(\bar{\epsilon}) = \frac{1}{C_3}\int_{p>1} \frac{d^3p}{p^2 p^{1+\bar{\epsilon}}} + \frac{1}{C_3} \times \int \frac{d^3p}{p^2}\left(\frac{1}{p'^{1+\bar{\epsilon}}} - \frac{\theta(p-1)}{p^{1+\bar{\epsilon}}}\right). \tag{4.3}$$

The first term gives the singularity at $\bar{\epsilon} = 0$, while the term in large parentheses, evaluated at $\bar{\epsilon} = 0$, provides the first correction. Thus

$$\bar{I}_\psi(\bar{\epsilon}) = 1/\bar{\epsilon} + 1 + O(\bar{\epsilon}). \tag{4.4}$$

Evaluated at $\bar{\epsilon} = 1$, Eq. (4.4) gives $\bar{I}_\psi(1) = 3$ as compared to 2.5 from the two-term ϵ-expansion and the exact answer of $\pi = 3.14$. The integral $\bar{I}_\psi(\bar{\epsilon})$ can be carried out exactly, giving

$$\bar{I}_\psi(\bar{\epsilon}) = \frac{\pi}{2(1 - \bar{\epsilon})}\cot\frac{\pi\bar{\epsilon}}{2}. \tag{4.5}$$

If we go further out in $\bar{\epsilon}$, we see that at $\bar{\epsilon} = 1$ the exact answer is $\pi^2/4 = 2.46$, while Eq. (4.4) yields the 20% lower value of 2. This is because of the closeness to the infrared singularity at $\bar{\epsilon} = 2$. The closer one gets to $\bar{\epsilon} = 2$, the worse the approximation of Eq. (4.4) becomes. If we now take the infrared pole into account and get the four-term approximant analogous to Eqs. (2.19) and (2.34), then

$$\bar{I}_\psi(\bar{\epsilon}) = \frac{1}{\epsilon} + \frac{1}{2 - \bar{\epsilon}} + \frac{1}{2} + O(\bar{\epsilon}^2). \tag{4.6}$$

At $\bar{\epsilon} = 1$, this gives $\bar{I}_\psi(1) = 2.50$, now only 2% too high. The improvement is obvious. Further improvement can be achieved if we proceed as we did in the previous section and include the nearest ultraviolet and infrared poles beyond $\bar{\epsilon} = 0$ and 2. These are the poles at $\bar{\epsilon} = -2$ and 4 and the four-pole approximant becomes

$$\bar{I}_\psi(\bar{\epsilon}) = \frac{1}{\bar{\epsilon}}\frac{1}{\bar{\epsilon} - 2} + \frac{1}{3}\cdot\frac{1}{\bar{\epsilon} + 2}\frac{1}{3}\cdot\frac{1}{\bar{\epsilon} - 4} + \frac{1}{4} + O(\bar{\epsilon}^2). \tag{4.7}$$

The $O(\bar{\epsilon})$ term of the analytic function vanishes. The values at $\bar{\epsilon} = 1$ and 1 are

$$\bar{I}_\psi(1) = 3.145 \tag{4.8a}$$

and

$$\bar{I}_\psi(1) = 2.472, \tag{4.8b}$$

with accuracies of one-tenth and one-third of one percent, respectively. We can now further improve the approximation of Eq. (4.7) at $\bar{\epsilon} = \frac{1}{2}$ by using the known value of $\bar{I}_\psi(\bar{\epsilon})$ at $\bar{\epsilon} = 1$. This is the added advantage of the $\bar{\epsilon}$-expansion that we had mentioned in the beginning of this section. Using the known value of $\bar{I}_\psi(1)$, we can make a parabolic approximation to the residual analytic part of $\bar{I}_\psi(\bar{\epsilon})$ (after the subtraction of the four poles). Thus

$$\bar{I}_\psi(\bar{\epsilon}) = \frac{1}{\bar{\epsilon}} - \frac{1}{\bar{\epsilon}-2} + \frac{1}{3}\cdot\frac{1}{\bar{\epsilon}+2} - \frac{1}{3}\cdot\frac{1}{\bar{\epsilon}-4} + \frac{1}{4} - 0.016\,\bar{\epsilon} + 0.008\,\bar{\epsilon}^2. \tag{4.9}$$

At $\bar{\epsilon} = \frac{1}{2}$,

$$\bar{I}_\psi(\tfrac{1}{2}) = 3.139, \tag{4.10}$$

now a slight underestimate by approximately one-tenth of one percent.

We regularize $\bar{I}_S$ similarly, as we did I_S in Sec. II.

$$\bar{I}_S(\bar{\epsilon}) = \frac{1}{C_3}\int_{p>1} \frac{d^3p}{p^{3+\bar{\epsilon}}} + \bar{J}_v(\bar{\epsilon}), \tag{4.11}$$

where

$$\bar{J}_v(\bar{\epsilon}) = \frac{1}{C_3}\int \frac{d^3p}{p^2p'^2}\left(\frac{(p^2-p'^2)^2}{p^{1+\bar{\epsilon}}+p'^{1+\bar{\epsilon}}} - \frac{p'^2\theta(p-1)}{p^{1+\bar{\epsilon}}}\right). \tag{4.12}$$

is the ultraviolet regularized integral. The asymptotic behavior near $\bar{\epsilon}\sim 0$ is given by

$$\bar{I}_S(\bar{\epsilon}) = \frac{2}{3\bar{\epsilon}} + \bar{J}_v(0) = \frac{2}{3\bar{\epsilon}} + \frac{8}{9}. \tag{4.13}$$

We note that this two-term expansion gives

$$\bar{I}_S(\tfrac{1}{2}) = 2.22, \tag{4.14}$$

which is closer to the numerical integration result of 2.25 ± 0.02 then the $\bar{\epsilon}$-expansion with all the various improvements of Secs. II and III. We emphasize that this accuracy of a two-term approximation is due to the absence of infrared poles and a consequent simplification of the analytic structure. A further check on the two-term expansion is provided by the values $\bar{I}_S(1) = 1.55$ and $\bar{I}_S(2) = 1.22$, in good agreement with the exact values of $\pi/2 = 1.57$ and $2\pi/(3\sqrt{3}) = 1.21$, respectively.

In the spirit of Sec. III, we study the effect of the nearest unphysical pole by continuing analytically the regularized integral $\bar{J}_v(\bar{\epsilon})$. The first pole is at $\bar{\epsilon} = -2$ with the principal part $\frac{3}{10}(\bar{\epsilon}+2)^{-1}$. The two-pole truncated Mittag–Leffler expansion of $\bar{I}_S(\bar{\epsilon})$ is accordingly

$$\bar{I}_S(\bar{\epsilon}) = \frac{2}{3\bar{\epsilon}} + \frac{3}{10}\frac{1}{\bar{\epsilon}+2} + \frac{133}{180}, \tag{4.15}$$

with

$$\bar{I}_S(\tfrac{1}{2}) \simeq 2.19. \tag{4.16}$$

A further improvement can be obtained by using the exact values quoted above for $\bar{\epsilon} = 1$ and $\bar{\epsilon} = 2$. We can now obtain the linear and quadratic terms in the residual analytic function. The final result is

$$\bar{I}_S(\bar{\epsilon}) = \frac{2}{3\bar{\epsilon}} + \frac{3}{10}\cdot\frac{1}{\bar{\epsilon}+2} + \frac{133}{180} + 0.0995\bar{\epsilon} - 0.0343\epsilon^2, \tag{4.17}$$

yielding

$$\bar{I}_S(\tfrac{1}{2}) \simeq 2.23. \tag{4.18}$$

This is in satisfactory agreement with the numerical answer of 2.25 ± 0.02.

We now evaluate $\bar{I}_S(\bar{\epsilon})$ for $\bar{\epsilon} \gg 1$ and show that it diverges when $\bar{\epsilon} \to \infty$. Introducing the variable $\mathbf{q}$, in terms of which

$$\vec{p} = \frac{\vec{1}}{2} + \frac{\vec{q}}{2}, \tag{4.19a}$$

and

$$\vec{p}' = \frac{\vec{1}}{2} - \frac{\vec{q}}{2}, \tag{4.19b}$$

we have for $\bar{\epsilon} \gg 1$,

$$\begin{aligned} p^{1+\bar{\epsilon}} + p'^{1+\bar{\epsilon}} &= 2^{-(1+\bar{\epsilon})}[(1+q^2+2\mu q)^{(1+\bar{\epsilon})/2} + (1+q^2-2\mu q)^{(1+\bar{\epsilon})/2}] \\ &\simeq 2^{-(1+\bar{\epsilon})}(1+q^2)^{(1+\bar{\epsilon})/2}\cosh\frac{2\mu q\bar{\epsilon}}{1+q^2}. \end{aligned} \tag{4.20}$$

The contribution to $\bar{I}_S(\bar{\epsilon})$ comes almost entirely from the region $q \simeq 0$, as otherwise the right-hand side of Eq. (4.20) becomes very large and consequently the integrand vanishes in Eq. (4.1). With this in mind we make approximations appropriate to $q \simeq 0$ to write

$$p^{1+\bar{\epsilon}} + p'^{1+\bar{\epsilon}} \simeq 2^{-(1+\bar{\epsilon})}\exp\left[\left(\frac{1+\bar{\epsilon}}{2}\right)^2\right]\cosh(2\mu q\bar{\epsilon}) \tag{4.21}$$

and

$$\bar{I}_S(\bar{\epsilon}) = 2^{\bar{\epsilon}}\int_0^\infty dq\cdot q^4\exp\left[-\left(\frac{1+\bar{\epsilon}}{2}\right)q^2\right] \times \int_{-1}^{1} d\mu\,\frac{\mu^2}{\cosh(2\mu q\bar{\epsilon})}. \tag{4.22}$$

For any given value of q, the hyperbolic function in the denominator of Eq. (4.22) increases rapidly with increasing μ and hence no appreciable error is made if the limits of integration are extended to $\pm\infty$. This yields

$$\bar{I}_S(\bar{\epsilon}) = \frac{\pi^3 2^{\bar{\epsilon}\to\infty}}{64\bar{\epsilon}^3}\int_0^\infty dq\cdot q\cdot\exp\left[-\left(\frac{1+\bar{\epsilon}}{2}\right)q^2\right] = \frac{\pi^3 2^{\bar{\epsilon}}}{32\bar{\epsilon}^3(1+\bar{\epsilon})}, \tag{4.23}$$

This clearly shows that, for $\bar{\epsilon} \gg 1$, $\bar{I}_S(\bar{\epsilon})$ rises, ultimately diverging as $\bar{\epsilon} \to \infty$. The function also diverges at $\bar{\epsilon} = 0$ and hence there must be a minimum for some value of $\bar{\epsilon}$.

FURTHER APPLICATIONS

In this section we look at some additional integrals as well as point out certain complexities in the analytic properties of integrals like $I_S(\epsilon)$. We consider first the integral $I_F(\epsilon)$ which occurs in the critical dynamics of the isotropic ferromagnet,

$$I_F(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2}\cdot\frac{(p^2-p'^2)^2}{p^{1+D/2}+p'^{1+D/2}}. \tag{5.1}$$

This is identical to $I_S(\epsilon)$, except for the differing exponent of p and p' in the denominator. The difference is largely a result of the existence of a conservation law for the order parameter (i.e., the magnetization.) As before we are ultimately interested in obtaining the answer at $D = 3$. Analytic solution is impossible because of the fractional powers of p and p' in the denominator at this value of D. The "physical" ultraviolet divergence occurs, not at $D = 4$ as in $I_S(\epsilon)$, but at $D = 6$. The expansion parameter therefore is $\epsilon = 6 - D$. Subtraction of the leading singularity leads to

$$I_F(\epsilon) = \frac{2}{DC_D}\int_{p>1}\frac{d^D p}{p^{3+D/2}} + I_v(\epsilon), \tag{5.2}$$

where $L_v(\epsilon)$ is the ultraviolet regularized integral

$$L_v(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2} \times\left(\frac{(p^2-p'^2)^2}{p^{1+D/2}+p'^{1+D/2}} - \frac{2p'^2\theta(p-1)}{Dp^{D/2}}\right). \tag{5.3}$$

The asymptotic behavior near the ultraviolet pole at $D = 6$ ($\epsilon = 0$) together with the leading correction is

$$\begin{aligned} I_F(\epsilon) &= 2/3\epsilon + \tfrac{1}{9} + L_v(0) \\ &= 2/3\epsilon + \tfrac{1}{9} + \tfrac{1}{12}(\pi + \tfrac{1}{3} - 4\ln 2) \\ &\simeq 2/3\epsilon + 0.053. \end{aligned} \tag{5.4}$$

The "physical" infrared pole is at $D = 2$ and regularization at this pole leads to

$$I_F(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2} + L_r(\epsilon), \tag{5.5}$$

where

$$L_r(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2}\left(\frac{(p^2-p'^2)^2}{p^{1+D/2}+p'^{1+D/2}} - 1\right). \tag{5.6}$$

The asymptotic part together with the leading correction is

$$\begin{aligned} I_F(\epsilon) &= 2/(D-2) + L_r(4) \\ &= 2/(4-\epsilon) - 2\ln 2. \end{aligned} \tag{5.7}$$

A "two-pole-plus-analytic-function" expansion in the spirit of Sec. II, now yields

$$I_F(\epsilon) = 2/3\epsilon + 2/(4-\epsilon) - 0.447 - 0.277\epsilon. \tag{5.8}$$

At $D = 3$, i.e., $\epsilon = 3$, this yields

$$I_F(3) = 0.94, \tag{5.9}$$

about 30% under the numerical integration of 1.27 ± 0.02, but much improved relative to the two-term expression of Eq. (5.4) which yields 0.27. This is a consequence of the closeness of $\epsilon = 3$ to the infrared pole.

To improve this, we now proceed as we did in Sec. III and introduce the first ultraviolet and infrared "unphysical" poles. These are the points where $L_v(\epsilon)$ and $L_r(\epsilon)$ diverge. We find that $L_v(\epsilon)$ diverges at $\epsilon = -4$ with the principal part $-(7/10)(\epsilon+4)^{-1}$, and $L_r(\epsilon)$ diverges at $\epsilon = 17/3$ with the principal part $-4/(3\epsilon - 17)$. The four-pole truncated Mittag–Leffler expansion is thus

$$I_F(\epsilon) = \frac{2}{3\epsilon} - \frac{2}{\epsilon-4} - \frac{7}{10}\frac{1}{\epsilon+4} + \frac{4}{3\epsilon-17} - 0.037 - 0.157\epsilon. \tag{5.10}$$

At $\epsilon = 3$

$$I_F(3) = 1.15, \tag{5.11}$$

about a 10% underestimate.

Another integral of interest results from introducing a parameter z into the denominator of a single-loop integral, corresponding to the modification of the response of the system at a finite frequency. Modified in this way, $I_S(\epsilon)$ becomes

$$I_S(\epsilon,z) = \frac{1}{C_D}\int \frac{d^D p}{p^2 p'^2}\frac{(p^2-p'^2)^2}{z+p^{D/2}+p'^{D/2}}. \tag{5.12}$$

At $D = 3$, this is the integral needed for a study of the fluctuation spectrum at the Lambda point of liquid helium. Clearly,

$$I_S(\epsilon) = I_S(\epsilon,0). \tag{5.13}$$

To demonstrate the effect of a nonzero z for $z = 1$, we simply quote the following results:

Two-term expansions

Ultraviolet

$$\begin{aligned} I_S(\epsilon,1) &= 1/\epsilon + \tfrac{1}{4} + (\tfrac{9}{4}\ln\tfrac{4}{3} - \tfrac{5}{8}) + O(\epsilon) \\ &\simeq 1/\epsilon + 0.27. \end{aligned} \tag{5.14}$$

Infrared

$$\begin{aligned} I_S(\epsilon,1) &= -1/(\epsilon-2) - 7 + 10\ln 2 \\ &\simeq -1/(\epsilon-2) - 0.07. \end{aligned} \tag{5.15}$$

Two-pole Mittag–Leffler expansion

$$I_S(\epsilon,1) = 1/\epsilon - 1/(\epsilon-2) - 0.23 - 0.08\epsilon\,. \tag{5.16}$$

Four-pole Mittag–Leffler expansion

$$I_S(\epsilon,1) \simeq \frac{1}{\epsilon} - \frac{1}{\epsilon-2} - \frac{1}{2}\cdot\frac{1}{\epsilon+4} + \frac{2}{\epsilon-8} + 0.155 + 0.19\epsilon\,. \tag{5.17}$$

From Eqs. (5.16) and (5.17) we obtain for $I_S(1,1)$, 1.69 and 1.85, respectively, as compared to the numerical integration value of 1.93 ± 0.01.

The corresponding modification of $\bar{I}_S(\bar{\epsilon})$ gives $\bar{I}_S(\frac{1}{2},1) = 1.92$ from a two-term expansion similar to Eq. (4.13). A two-pole Mittag-Leffler expansion similar to Eq. (4.17) (and fixed at $\bar{\epsilon} = 1$) yields 1.92.

We now return to $I_S(\epsilon)$ to point out certain analytic properties in the complex ϵ plane that we ignored in Sec. III. The infrared regularized integral $J_r(\epsilon)$ has a pole at $\epsilon = 8/3$. If we make an additional subtraction and analytically continue beyond $\epsilon = 8/3$, we find the next pole at $\epsilon = 3$. To systematize this, we note that these singularities are those associated with the integral

$$I_L(\epsilon) = \frac{1}{C_D}\int \frac{d^D p}{p^2(1+p^{D/2})}. \tag{5.18}$$

This integral can be analytically evaluated to yield

$$I_L(\epsilon) = \frac{2\pi}{D}\csc\left(\frac{\pi\epsilon}{D}\right). \tag{5.19}$$

The singularities for positive ϵ are at

$$\epsilon/((4-\epsilon) = \epsilon/D = n$$

or

$$\epsilon = 4 - 4/(n+1), \tag{5.20}$$

where n is a positive nonzero integer. $n = 1$ corresponds to the physical infrared singularity at $\epsilon = 2$. $n = 2$ yields the first unphysical singularity at $\epsilon = 8/3$ that we treated in Sec. III. $n = 3$ produces a singularity at $\epsilon = 3$ and so on. Each value of n gives rise to singularities that are bunched closer and closer together and for $n \to \infty$, we get the singularity at $\epsilon = 4$, which is a limit point for this set of singularities. Thus $\epsilon = 4$ is a point of essential singularity for the function. This property was ignored in Sec. III. Being some distance away from the physical region, it did not affect the situation very much. However, the essential singularity can be a source of trouble if it is closer to the physical region and should be taken into account if necessary.

VI. SUMMARY

We have considered integrals which are important in the study of critical dynamics of liquid helium, ferromagnets, etc. These integrals cannot in general be evaluated in a three-dimensional space. The usual technique is to consider the dimensionlity D as a variable and develop the integral as a power series in the variable $\epsilon = D_C - D$. D_C is a critical dimension at which the integral has an ultraviolet singularity. We have considered the analytic properties of such integrals in the complex ϵ plane and have shown how the ϵ-expansion can be improved upon by the inclusion of further singularities. For functions with a complicated singularity structure, it is found that inclusion of more poles leads to a better representation in the "physical region." This motivates the $\bar{\epsilon}$-expansion of Sec. IV. This is a technique where a significant simplification occurs in the pole structure, and consequently an expansion based on a single pole gives excellent results. We conclude on a cautionary note by pointing out the existence of an essential singularity in one of the integrals studied. However, so long as such a singularity is well removed from the physical region, one can expect to get accurate results.

ACKNOWLEDGMENT

The authors wish to acknowledge support from the National Science Foundation under grant DMR 79-01172.

[1]P.C. Hobenberg and B.I. Halperin, Rev. Mod. Phys. **49**, 435 (1977).
[2]B.I. Halperin, P.C. Hohenberg, and E.D. Siggia, Phys. Rev. B **13**, 1299 (1976).
[3]C. DeDominicis and L. Peliti, Phys. Rev. B **18**, 353 (1978).
[4]R.A. Ferrell and J.K. Bhattacharjee, J. Low Temp. Phys. **36**, 165 (1979).
[5]R. Freedman and G.F. Mazenko, Phys. Rev. B **13**, 4967 (1976).
[6]V. Dohm, Solid State Commun. **20**, 657 (1967).
[7]J.K. Bhattacharjee and R.A. Ferrell (to be published), U. of Maryland, Tech. Rep. #80-030.
[8]R.A. Ferrell, N. Menyhárd, H. Schmidt, F. Schwabl, and P. Szépfalusy, Phys. Rev. Lett **18**, 891 (1967); Ann. Phys. (N.Y.) **47**, 565 (1968).
[9]B.I. Halperin and P.C. Hohenberg, Phys. Rev. **177**, 952 (1969).
[10]K.G. Wilson and J. Kogut, Phys. Rep. C **12**, 75 (1974).
[11]G. t'Hooft and M. Veltman, Nucl. Phys. B **44**, 189 (1972).
[12]See for example, L.V. Ahlfors, *Complex Analysis* (McGraw-Hill, New York, 1966).

On exact solutions of the nonlinear Heisenberg–Klein–Gordon equation in a space–time of constant spacelike curvature

Eckehard W. Mielke

Institut für Reine und Angewandte Kernphysik der Christian–Albrechts–Universität Kiel, Olshausenstrasse 40—60, 2300 Kiel, Federal Republic of Germany

(Received 5 April 1979; accepted for publication 18 May 1979)

A nonlinear complex *scalar* field theory associated with a "squared" *Heisenberg–Pauli–Weyl nonlinear spinor equation* is considered. In a $d+1$ dimensional universe of constant spatial curvature *exact* localized solutions for the resulting $[|\varphi|^{2d/(d-2)} - \text{const}\ |\varphi|^{2(d-1)/(d-2)}]$ model are constructed. For "soliton-like" solutions with quantized (nontopological) charge the filed energy and the Heisenberg uncertainty principle are analyzed.

1. INTRODUCTION

In a proposed *geometrodynamical model*[1] *of extended particles* strong interactions are described by a GL($2n$,C) gauge theory[2] with an Einstein–Cartan–Dirac-type Lagrangian. At the center of such a model is the massive *Heisenberg–Pauli–Weyl nonlinear spinor equation*[3,4]

$$\left\{ i\gamma^{\mu}\nabla_{\mu} + \frac{3\epsilon}{16}\ell^2[\gamma_{\mu}(\bar{\psi}\gamma^{\mu}\psi) + \gamma_{\mu}\gamma_5(\bar{\psi}\gamma^{\mu}\gamma^5\psi)] - (\mu c/\hbar)\right\}\psi = 0, \qquad \epsilon = \pm 1. \tag{1}$$

generalized to curved space–time. Similarly, as in Heisenberg's unified field theory[3] the nonlinear terms induced into the Dirac equation by *Cartan's torsion*[5] are expected to produce a binding force between the *n fundamental* Dirac spinors $\psi^{(q)}$ (distinguished by n different flavors or colors q). Moreover, according to some ideas anticipated already by Wheeler, these quark fields may become confined by the curvature of the underlying space–time in geon-type objects.[24] This curved geometry is self-consistently created by the stress–energy of the nonlinear spinor fields via Einstein-type field equations. The dimension in which this should happen is characterized by the *modified* Planck length

$$L^* \equiv (\hbar G_S/c^3)^{1/2} = \hbar/cM^* \sim 10^{-14}\ \text{cm} \tag{2}$$

of *strong gravity*.[6]

This is an ambitious program. In a related, less pretentious investigation,[7] the nonlinear spinor equation (1) has been "squared" in order to avoid algebraic complications but to maintain essentially the same dynamical features. Localized solutions of the resulting Heisenberg–Klein–Gordon equation for the corresponding *scalar* fields then have been obtained numerically in a flat and exterior Schwarzschild background. In this paper, *exact* quasistable[8] solutions of a similar (with n internal degrees of freedom) scalar equation are studied in a space–time having *constant* Gaussian *spacelike* curvature. The derivation generalizes a remarkable construction of Okolowski[8] by means of conformal techniques[9] already employed[10] in the Rosen model.[11]

2. THE MODEL

The n complex (charged) scalar fields $\varphi \equiv \{\varphi^{(q)}(x) | q = 1,...,n\}$ will be considered in a curved pseudo-Riemannian space–time owning $d > 2$ spatial dimensions. The model is defined by the Lagrangian density

$$\mathcal{L}_{\text{HKG}} = \frac{\hbar^2}{2\mu}\sqrt{-f}\,[\partial^{\nu}\varphi^*\partial_{\nu}\varphi - U(\varphi)], \tag{3}$$

where

$$U(\varphi) \equiv d(d-2)\left(\frac{\ell}{4}\right)^{4/(d-2)}(\varphi^*\varphi)^{d/(d-2)} - d(d-2) \times \frac{\epsilon\mu c}{\hbar}\left(\frac{\ell}{4}\right)^{2/(d-2)}(\varphi^*\varphi)^{(d-1)/(d-2)} + (\mu c/\hbar)^2\varphi^*\varphi \tag{4}$$

denotes the field potential.

Variation of $\mathcal{L}_{\text{HKG}}$ with respect to φ^* yields the nonlinear field equations

$$\left[\frac{\partial^2}{c^2dt^2} - L^{*-2}\Delta - \epsilon d(d-1)(\mu c/\hbar)\left(\frac{\ell}{4}\right)^{2/(d-2)} \times (\varphi^*\varphi)^{1/(d-2)} + d^2\left(\frac{\ell}{4}\right)^{4/(d-2)}(\varphi^*\varphi)^{2/(d-2)} + (\mu c/\hbar)^2\right]\varphi = 0, \tag{5}$$

which, for $d = 3$, coincides with the *Heisenberg-Klein-Gordon equation* of Ref. 7. Then it can be related to the "squared" form of Eq. (1). It admits the constant solution

$$\varphi_0^{(n)} = \frac{4}{\ell}\left[\frac{\mu c}{2d\hbar}(\epsilon(d-1) \mp \sqrt{(d-1)^2-4})\right]^{(d-2)/2}, \qquad \varphi_0^{(q)} = 0 \ \text{else}, \tag{6}$$

that for $\epsilon = +1$ might signal the occurrence of a *spontaneous breaking* of the internal U(n) *symmetry* in the vacuum sector.[12]

3. LOCALIZED SOLUTIONS IN A CURVED BACKGROUND

Passing on to a rest frame, stationary solutions of (5) will be studied by means of the familiar separation ansatz

$$\varphi^{(q)}(x) = \frac{4}{\ell}\left(\frac{M^*c}{n\hbar}\right)^{(d-2)/2}F(\mathbf{z},\omega) \times \exp[-i(t-t_0)\omega\mu c^2/\hbar]. \tag{7}$$

In terms of the modified Planck length (2), the spatial coordinates $\mathbf{x} \equiv (x^{\alpha} | \alpha = 1,...,d)$ may be replaced by the dimen-

 0022-2488/80/030543-04$1.00

sionless variables

$$\mathbf{z} \equiv L^{*-1}\mathbf{x}. \tag{8}$$

For definiteness, it will be assumed that the Planck mass M^* associated with the curved background is related via

$$M^* = \beta\mu, \tag{9}$$

to the *bare* mass μ of the nonlinear model (4) under consideration.

Then, the dimensionless field $F(\mathbf{z},\omega)$ (in the following abbreviated with F) satisfies the equation

$$\frac{4(d-1)}{d-2}\frac{d}{4}\Delta F - RF + \bar{R}F^{(d+2)/(d-2)} + \frac{d(d-1)}{\beta(d-2)}\frac{R}{K}F^{d/(d-2)} = 0, \tag{10}$$

where

$$\Delta \equiv L^{*2}\frac{1}{\sqrt{|{}^{(d)}f|}}\partial_\alpha\left({}^{(d)}f^{ab}\sqrt{|{}^{(d)}f|}\,\partial_b\right) \tag{11}$$

denotes the dimensionless spacelike, Laplace–Beltrami operator. In (10), the abbreviations

$$R = \frac{d(d-1)}{\beta^2(d-2)}(1-\omega^2), \tag{12}$$

$$K \equiv \frac{R}{\epsilon d(d-1)} = \frac{(1-\omega^2)}{\epsilon(d-2)\beta^2}, \tag{13}$$

$$\bar{R} = -\frac{(d-1)d^3}{d-2}, \tag{14}$$

are used to bring out a close connection with differential geometry. In fact, Eq. (10) may then be viewed as a *deformed* version of Lichnerowicz's equation.[13] In the geometric context the latter governs the realization of arbitrary scalar curvatures $(4/d)\bar{R}$ and $(4/d)R$ by means of a conformal change[9]

$$dS^2 \to d\mathcal{S}^2 = F^{4/(d-2)}\,dS^2, \tag{15}$$

of a hypothetical Riemannian metric. As in a similar exact soluble model[10] on a d-dimensional manifold, the alteration

$$F = \phi^{(2-d)/2}, \tag{16}$$

of the "scale dimensions" of the field transforms (10) into the simpler looking equation

$$\bar{R} = R\left(\phi^2 - \frac{d(d-1)}{K(d-2)\beta}\phi\right) + \frac{d}{4}(d-1)\left[2\phi\Delta\phi - d\partial^c\phi\partial_c\phi\right], \tag{17}$$

which is reminiscent of (A5) of Ref. 9.

Following Okolowski,[8] we confine ourselves to an underlying space–time owning a $(d-1)$-dimensional Euclidean hypersurface of *constant* (Gaussian) *sectional curvature* K [defined by (13)]. Then, the corresponding line element can be written (for both signs of K!) as[14]

$$ds^2 \equiv f_{\mu\nu}\,dx^\mu\,dx^\nu = c^2\,dt^2 - L^{*2}\left(d\chi^2 + \frac{\sin^2(\chi K^{1/2})}{K}{}^{(d-1)}\delta_{kl} \times dz^k\,dz^l\right). \tag{18}$$

In this strong gravitational[6] background it is feasible to consider only those solutions of (17) which, in addition, satisfy the inhomogeneous Laplace equation

$$\Delta\phi = -\frac{8}{d-2}\left[\Omega d + (d+1)\left(\frac{K(d-2)\beta}{d}\phi - 1\right)\right]. \tag{19}$$

A spherically symmetric solution of (19) is

$$\phi = \frac{d}{\beta K(d-2)}\{1 - \Omega\sin^2(\chi K^{1/2})\}. \tag{20}$$

A bit of algebra reveals that the ansatz (20) also satisfies (17) provided that

$$\Omega(\omega) = 1 + \frac{d-2}{d}\beta^2 K = 1 + \epsilon\frac{1-\omega^2}{d}. \tag{21}$$

The physical meaning of the obtained solutions becomes more transparent by changing the coordinate χ to

$$\rho^2 \equiv \frac{1}{K}\sin^2(\chi\sqrt{K}). \tag{22}$$

With respect to this *Schwarzschild radial coordinate*[14] ρ the line element (18) reads as follows:

$$ds^2 \equiv c^2\,dt^2 - L^{*2}\left[\frac{d\rho^2}{1-K\rho^2} + \rho^{2(d-1)}\delta_{kl}\,dz^k\,dz^l\right]. \tag{23}$$

It is obvious from (23) that $2\pi\rho$ measures the "proper circumference," as desired. Then, the radially *localized* solutions of (10) finally read

$$F(\rho,\omega) = \left[\frac{d}{\beta K(d-2)}(1-\Omega K\rho^2)\right]^{(2-d)/2} = \left[\epsilon\frac{\beta d}{1-\omega^2} - \frac{d}{\beta(d-2)} \times\left(1+\epsilon\frac{1-\omega^2}{d}\right)\rho^2\right]^{(2-d)/2}. \tag{24}$$

It is interesting to note the strong interrelation between the solutions and the background space–time. In a *closed* ($K>0$, i.e., $0<\omega<1$ for $\epsilon=+1$ and $\omega>1$ for $\epsilon=-1$) minicosmos, the solutions have a pole at

$$\rho_0 = \left(K + \beta^2\frac{d-2}{d}K^2\right)^{-1/2} < K^{-1/2}, \tag{25}$$

which is located *inside* the universe of hadronic dimensions. This is in contradistinction to the *open* case ($K<0$). There, the solutions (24) are *regular everywhere*, provided that

$$\epsilon < \omega^2 < d + \epsilon. \tag{26}$$

From the viewpoint of differential geometry[15] it is a remarkable but not completely understood fact, that the obtained solutions (24), those of a similar model,[10] as well as the instanton solutions[16] of Yang–Mills theory, are all closely related with *spaces of constant curvature*.

4. FIELD ENERGY OF "SOLITONS" WITH QUANTIZED CHARGE

Lacking a more fundamental formalism, a very crude way of *quantizing* a nonlinear field theory already noted by

Finkelstein *et al.*[17] is as follows: Postulate the total charge

$$Q \equiv \frac{e\hbar}{2\mu c} i \int d^d x \sqrt{|\,^{(d)}f\,|}\, f^{00}[\varphi^* \partial_0 \varphi - (\partial_0 \varphi^*)\varphi\,], \tag{27}$$

of a solution (24) to be an integral multiple k of the quantum of charge e, i.e.,

$$Q = ke. \tag{28}$$

For stationary solutions this is equivalent to imposing the *Bohr–Sommerfeld quantization conditions.*[12] The insertion of the ansatz (7) yields

$$k = 16\omega \frac{L^{*2}}{\ell^2} \int_{S^{d-1}} d\mu_g \int_0^{\rho} d\rho \frac{\rho^{d-1}}{(1-K\rho^2)^{1/2}} F^2; \quad \rho = 1/\sqrt{K} \text{ for } K>0, \quad \rho = \infty \text{ for } k<0. \tag{29}$$

The remaining integral can be evaluated in closed form by noting the integral representation

$${}_2F_1(a,b;c;\Omega) \equiv \frac{2\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 \rho^{2b-1}(1-\rho^2)^{c-b-1}(1-\Omega\rho^2)^{-a} d\rho, \quad \mathrm{Re}\, c > \mathrm{Re}\, b > 0, \quad |\arg(1-\Omega)| < \pi, \tag{30}$$

or

$${}_2F_1(a,b;c;1-\Omega) \equiv \frac{2\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^{\infty} \rho^{2b-1}(1+\rho^2)^{a-c}(1+\Omega\rho^2)^{-a} d\rho, \quad \mathrm{Re}\, c > \mathrm{Re}\, b > 0, \quad |\arg\Omega| < \pi, \tag{30'}$$

of the hypergeometric function.[18]

The result (written below for $K>0$ only)

$$k = 8\beta^{d-2}\omega \frac{L^{*2}}{\ell^2} |K|^{-d/2} \left(\frac{d}{|K|(d-2)}\right)^{2-d} \left(\int_{S^{d-1}} d\mu_g\right) \frac{\pi^{1/2}\Gamma(d/2)}{\Gamma[(d+1)/2]} {}_2F_1(d-2,d/2;(d+1)/2;\Omega), \tag{31}$$

may be viewed as a *normalization* of the reduced Compton wave length βL^* of the soliton with respect to the *coupling constant* (fundamental length) ℓ of the model and to other occuring parameters.

An important piece of information is provided by the *classical field energy* of localized solutions. The reason being, that it constitutes a first approximation to the energy eigenvalues of the corresponding quantum mechanical states.[12] In a background with $f_{0b} = 0$, the Halmiltonian may be derived from (3) as

$$E = \frac{\hbar^2}{2\mu} \int d^d x \sqrt{|f|}\, \{ f^{00}(\partial_0 \varphi)^*(\partial_0 \varphi) - f^{ab}(\partial_a \varphi)^*(\partial_b \varphi) + U(\varphi)\}. \tag{32}$$

The calculation will be facilitated by subtracting from (32) the boundary term

$$E_0 \equiv -\frac{\hbar^2}{2\mu} \int d^d x\, \partial_a (\varphi^* \sqrt{|f|}\, f^{ab} \partial_b \varphi). \tag{33}$$

Because of the field equations (5), this procedure yields

$$E - E_0 = \frac{\hbar^2}{\mu} \int d^d x \sqrt{|f|} \left\{ f^{00}(\omega\mu c/\hbar)^2 \varphi^* \varphi + \epsilon \frac{d\mu c}{2\hbar} \left(\frac{\ell}{4}\right)^{2/(d-2)} (\varphi^* \varphi)^{(d-1)/(d-2)} - d\left(\frac{\ell}{4}\right)^{4/(d-2)} (\varphi^* \varphi)^{d/(d-2)} \right\}. \tag{34}$$

After inserting the ansatz (7), the expression

$$E - E_0 = 16\mu c^2 \frac{L^{*2}}{\ell^2} \int_{S^{d-1}} d\mu_g \int_0^{1/\sqrt{K}} d\rho \frac{\rho^{d-1}}{(1-K\rho^2)^{1/2}} \left\{ \omega^2 F^2 + \epsilon\beta \frac{d}{2} F^{2(d-1)/(d-2)} - \beta^2 dF^{2d/(d-2)} \right\}, \tag{35}$$

may be obtained.

After performing the integrations with the aid of (30) and applying the quantization condition (29) the *mass formula* for Heisenberg solitons in a singlet state finally reads:

$$\frac{E - E_0}{c^2} = k\mu \left\{ \omega - \frac{1-\omega^2}{2\omega} \frac{{}_2F_1(d-1,d/2;(d+1)/2;\Omega) + (2/d)(1-\omega^2)\,{}_2F_1(d,d/2;(d+1)/2;\Omega)}{{}_2F_1(d-2,d/2;(d+1)/2;\Omega)} \right\}. \tag{36}$$

According to it, the "bare" mass μ appears to be (classically) "*renormalized*" by self-energy contributions. Considered as a function of ω the energy (36) may be endowed with local minima as in Ref. 7. This may be the result of an "*Archimedes effect*"[2] in the hadronic environment that originates from the nonlinear self-interactions of all quark-type fields $\varphi^{(q)}$.

5. THE HEISENBERG UNCERTAINTY RELATION

The normalization (31) is also reflected in the *mean size* of radially localized solutions. Their mean radius may be

measured by

$$\langle|\mathbf{x}|^2\rangle_\varphi^{1/2} \equiv L^*\left(\frac{\int \rho^2\varphi^*\varphi(-{}^{(d)}f)^{1/2}\,d^dx}{\int \varphi^*\varphi(-{}^{(d)}f)^{1/2}\,d^dx}\right)^{1/2}. \tag{37}$$

(Compare with Ref. 19 for a similar definition.)

In the case of "quantized solitons,"

$$\langle|\mathbf{x}|^2\rangle_\varphi = \frac{\ell^2}{4}\frac{k}{\omega}\beta^{2-d}|K|^{d/2}\left(\frac{d}{|K|(d-2)}\right)^{d-2}\left(\int_{S^{d-1}} d\mu_g\right)^{-1}\frac{1}{\pi} \times \frac{\Gamma((d+1)/2)\,d}{(d+1)\Gamma(d/2)}\frac{{}_2F_1(d-2,(d+2)/2;(d+3)/2;\Omega)}{{}_2F_1(d-2,d/2;(d+1)/2;\Omega)^2}, \tag{38}$$

may be obtained.

According to a formal proof given by Weyl,[20] the relation

$$\langle|\mathbf{x}|^2\rangle_\varphi\langle|\mathbf{p}|^2\rangle_\varphi \geqslant \frac{d^2}{4}\hbar^2, \tag{39}$$

holds for the lower bound in *Heisenberg's uncertainty principle*. Up to the surface term $2\mu E_0/c$ the *mean* squared *momentum operator* may be defined by

$$\langle|\mathbf{p}|^2\rangle_\varphi \equiv \hbar^2\int d^dx\{f^{ab}\partial_a\varphi^*\partial_b\varphi\}\sqrt{|{}^{(d)}f|}\Big/\left(\int d^dx\varphi^*\varphi\sqrt{|{}^{(d)}f|}\right). \tag{40}$$

The above form is also employed by Lieb[21] in an investigation of the implications of (39) in nonlinear quantum mechanics.

A calculation reveals that

$$\hbar^{-2}\langle|\mathbf{p}|^2\rangle_\varphi = \frac{8}{\ell^2}\frac{\omega}{k}\Omega^2(2-d)^2|K|^{-d/2}\left(\frac{d}{|K|(d-2)}\right)^{2-d}\beta^{d-2}\left(\int_{S^{d-1}} d\mu_g\right)\frac{\Gamma((d+2)/2)\pi^{1/2}}{\Gamma((d+5)/2)} \times {}_2F_1(d,d/2+1;(d+5)/2;\Omega). \tag{41}$$

Therefore, the uncertainty relation (39) yields the restriction

$$\frac{8}{\pi^{1/2}}\Omega^2\frac{1}{(d+1)^2(d+3)}\frac{{}_2F_1(d-2,(d+2)/2;(d+3)/2;\Omega)\,{}_2F_1(d,(d+2)/2;(d+5)/2;\Omega)}{{}_2F_1(d-2,d/2;(d+1)/2;\Omega)^2} \geqslant 1, \tag{42}$$

on ω. A violation of this unequality might signal a serious breakdown with respect to a *probabilistic* interpretation of any nonlinear semiclassical theory. (This curious phenomenon has been termed "ultraquantum" in a related study.[22]) In view of this, it might be fruitful to implement the ideas of de Broglie[23] into such models.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Professor J.A. Wheeler and Professor F.W. Hehl for their generous support and encouragement. Furthermore, I am happy to thank Kathryn Carradus for her inspirational influence. This work was supported by the Deutsche Forschungsgemeinschaft, Bonn.

[1]E.W. Mielke, Phys. Rev. Lett. **39**, 530 (1977); **39**, 851 (E) (1977).
[2]A. Salam, in *Five Decades of Weak Interactions*, Marshak Festschrift, (Am. Academy of Arts and Sciences, New York, 1977) Vol. 294, p. 12.
[3]W. Heisenberg, *Introduction to the Unified Field Theory of Elementary Particles* (Wiley, London, 1966).
[4]H. Weyl. Phys. Rev. **77**, 699 (1950).
[5]F.W. Hehl, P. von der Heyde, G.D. Kerlick, and J.M. Nester, Rev. Mod. Phys. **48**, 393 (1976).
[6]C.J. Isham, A. Salam, and J. Strathdee, Phys. Rev. D **3**, 867 (1971).
[7]W. Deppert and E.W. Mielke, "Localized Solutions of the nonlinear Heisenberg–Klein–Gordon Equation. In flat and exterior Schwarzschild space-time," Phys. Rev. D (to be published).
[8]J.A. Okolowski, J. Math. Phys. **19**, 2253 (1978).
[9]E.W. Mielke, Gen. Rel. Grav. **8**, 321 (1977), Appendix, Eqs. (A5) and (A6).
[10]E.W. Mielke, Phys. Rev. D **18**, 4525 (1978).
[11]G. Rosen, J. Math. Phys. **6**, 1269 (1965).
[12]R. Jackiw, Rev. Mod. Phys. **49**, 681 (1977).
[13]A. Lichnerowicz, J. Math. Pures Appl. **23**, 37 (1944).
[14]C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), pp. 721–723.
[15]E.W. Mielke, "On Pseudoparticle Solutions in Riemannian Geometry" (submitted to Gen. Rel. Grav.).
[16]A.A. Belavin, A.M. Polyakov, A.S. Schwartz, and Yu.S. Tyupkin, Phys. Lett. B **59**, 85 (1975).
[17]R. Finkelstein, R. LeLevier, and R. Ruderman, Phys. Rev. **83**, 326 (1951).
[18]W. Magnus, F. Oberhettinger, and R.P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, Berlin, 1966), p. 54.
[19]L. Vazquez, Lett. Nuovo Cimento **19**, 37; 561 (1977).
[20]L.D. Landau and E.M. Lifschitz, *Quantum Mechanics—Nonrelativistic Theory* (Pergamon, London, 1967), p. 45.
[21]E.H. Lieb, Rev. Mod. Phys. **48**, 553 (1976).
[22]A.F. Rañada and L. Vazquez, Phys. Rev. D **19**, 493 (1979).
[23]L. de Broglie, *Non-Linear Wave Mechanics* (Elsevier, Amsterdam, 1960).
[24]J.A. Wheeler, *Geometrodynamics* (Academic, New York, 1962).

Finite field equation of Yang–Mills theory

Richard A. Brandt
Department of Physics, New York University, 4 Washington Place, New York, New York, 10003

Ng Wing-Chiu[a)]
Physikalisches Institut der Universität Bonn, 53 Bonn 1, Nussallee 12, F.R. Germany

Wai-Bong Yeung[b)]
Department of Physics, Queen Mary College, University of London, Mile End Road, London, E1 4NS, England

(Received 1 December 1978; accepted for publication 20 April 1979)

We consider the finite local field equation
$-\{[1+1/\alpha(1+f_4)]g^{\mu\nu}\Box-\partial^\mu\partial^\nu\}A^{\nu a}=-(1+f_3)g^2N[A^{c\nu}A^{a\mu}A^c_\nu]+\cdots+(1-s)^2M^2A^{a\mu}$,
introduced by Lowenstein to rigorously describe SU(2) Yang–Mills theory, which is written in terms of normal products. We also consider the operator product expansion
$A^{c\nu}(x+\xi)A^{a\mu}(x)A^{b\lambda}(x-\xi)\sim\Sigma M^{cab\nu\mu\lambda}_{c'a'b'\nu'\mu'\lambda'}(\xi)N[A^{\nu'c'}A^{\mu'a'}A^{\lambda'b}](x)$,
and using asymptotic freedom, we compute the leading behavior of the Wilson coefficients $M^{\cdots}_{\cdots}(\xi)$ with the help of a computer, and express the normal products in the field equation in terms of products of the c-number Wilson coefficients and of operator products like $A^{c\nu}(x+\xi)\,A^{a\mu}(x)\,A^{b\lambda}(x-\xi)$ at separated points. Our result is
$-\{[1+(1/\alpha)(1+f_4)]g^{\mu\nu}\Box-\partial^\mu\partial^\nu\}A^{\nu a}=-(1+f_3)g^2\lim_{\xi\to0}\{(\ln\xi)^{-0.28/2b}$
$\times[A^{c\nu}(x+\xi)A^{a\mu}(x)A^c_\nu(x-\xi)+\epsilon^{abc}A^{\mu c}(x+\xi)\partial^\nu A^b_\nu(x)+\cdots]+\cdots\}+(1-s)^2M^2A^{a\mu}$,
where $\beta(g)=-bg^3$, and so $(\ln\xi)^{-0.28/2b}$ is the leading behavior of the c-number coefficient multiplying the operator products in the field equation.

I. INTRODUCTION

We continue the program we inaugurated[1] by extending our inquiry to Yang–Mills theory. We shall arrive at a finite field equation in point-separation form for the Yang–Mills equation, in which the source current of the interacting theory is a finite composite field operator, defined in terms of the short-distance limit of the products of singular functions and of elementary operators at separated points. We compute explicitly the c-number Wilson coefficients in these operator-product expansions, making use of the asymptotic freedom of the Yang–Mills theory. Thus, the structure of the *composite* interaction source current is now completely known in terms of these c-number coefficients and point-separated products of elementary quantum fields.

The short-distance behavior of the Yang–Mills theory is of course intimately tied to the renormalization. While it has been shown to be renormalizable for some time now, a detailed renormalization study was only more recently available, which enables the form of the finite field equation in the form of normal products to be established[2]:

$$\begin{aligned}-\left\{\left[1+\frac{1}{\alpha}(1+f_4)\right]\delta^{\mu\nu}\Box-\partial^\mu\partial^\nu\right\}A^{\nu a}&=(1+f_1)gN[\epsilon^{abc}A^{\mu c}\partial^\nu A^b_\nu]\\&\quad-2(1+f_1)gN[\epsilon^{abc}A^{c\nu}\partial_\nu A^{b\mu}]\\&\quad+(1+f_2)gN[\epsilon^{abc}A^{c\nu}\partial^\mu A^b_\nu]\\&\quad-(1+f_3)g^2N[A^{c\mu}A^{a\nu}A^c_\nu]\\&\quad-(1+a)\frac{g}{2}N[\epsilon^{abc}(\partial^\mu\bar c^b)c^c]\\&\quad-(1+f_3)g^2N[A^{c\nu}A^{a\mu}A^c_\nu]\\&\quad+(1-s)^2M^2A^{a\mu}.\end{aligned}\tag{1.1}$$

For illustrative purposes we consider just the term $N[A^\nu_c A^{a\mu}A^c_\nu]$. This would occur on the right-hand side of the operator-product expansion $A_c(x)A^{a\mu}(0)A^c(-x)$, for example, but it is not the only contribution to that expansion. All the normal products in Eq. (1.1), being of the same dimension and quantum number, can also occur in the OPE. Thus, we need to consider other OPE's, like $\epsilon^{abc}A^{\nu c}(x)\times\partial^\mu A^b_\nu(0)$, which also has the whole menagerie of normal products on the RHS. So we have to consider a large set of different operator-product expansions of identical quantum numbers in order to be able to determine each of these normal products in terms of c-number coefficients (to be calculated), and point-separated products of elementary fields. In other words, the normal products are determined by inverting our set of operator-product expansions.

In the case of Yang–Mills theory, this is a horrendous task indeed. Eventually, we have as many as 25 operators that mix among themselves in this array of operator-product expansions. This is only so for the case of the SU(2) gauge group, which is the only case we treat in this paper; for higher groups there will be even more participating operators. (We have to leave it for future generations, who hopefully would have evolved better calculational tools.)

Why are we doing this? The main motivation is to be able to check renormalization-induced invariances in field

a)Present address: Department of Physics, McGill University, Montreal, P.Q., Canada.
b)Present address: Department of Physics, National Central University, Chung-Li, Taiwan 320, Republic of China.

 0022-2488/80/030547-14$1.00

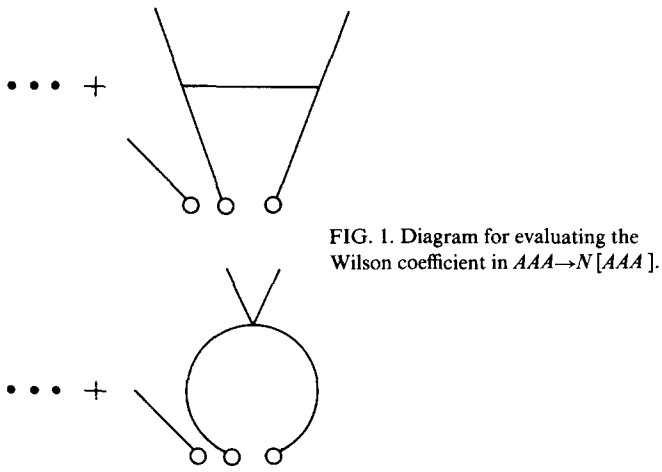

FIG. 1. Diagram for evaluating the Wilson coefficient in $AAA \to N[AAA]$.

theory with greater confidence. For example, under the R transformation

$$\mathbf{A}_\mu(x) \to \mathbf{A}_\mu(x) + \mathbf{r}_\mu, \tag{1.2}$$

the finite field equation in point-separated form we shall obtain is uniquely suited to investigating whether this invariance is present in the theory. We have discussed at length the procedure for doing this elsewhere[3] and will not repeat the analysis here. For the interested reader, we note only that it is extremely convenient for this purpose to perform a Lorentz invariant average over direction $x_\mu \to 0$ in defining local field products, before the R transformation (1.1) is implemented.

II. ENUMERATION OF OPERATORS

The finite field equation for Yang–Mills theory in terms of normal products was first obtained by Lowenstein[2] and it takes the form

$$\begin{aligned}
&-\left\{\left[1+\frac{1}{\alpha}(1+f_4)g^{\mu\nu}\right]\Box-\partial^\mu\partial^\nu\right\}A^{\nu a}\\
&\quad=(1+f_1)gN[\epsilon^{abc}A^{\mu c}\partial^\nu A^b_\nu]\\
&\qquad-2(1+f_1)gN[\epsilon^{abc}A^{\nu c}\partial_\nu A^{b\mu}]\\
&\qquad+(1+f_2)gN[\epsilon^{abc}A^{\nu c}\partial^\mu A^b_\nu]\\
&\qquad-(1+f_3)g^2N[A^{c\mu}A^{a\nu}A^c_\nu]\\
&\qquad-(1+a)(g/2)N[\epsilon^{abc}(\partial^\mu\bar{c}^b)c^c]\\
&\qquad-(1+f_3)g^2N[A^{c\nu}A^{a\mu}A^c_\nu]\\
&\qquad+(1-s)^2M^2A^{a\mu}.
\end{aligned} \tag{2.1}$$

In the above, f_1, f_2, f_3, f_4, a, and s are finite quantities expressible as power series of the coupling constant g. M is a mass parameter such that the Green's functions of the theory are defined partially at the origin and partially at the mass M.

We shall use Wilson's short-distance operator product expansions to relate each of these normal products or composite operators in terms of short-distance limits of products of c-number Wilson coefficients and point-separated products of elementary operators. However, to obtain the six composite operators in Eq. (2.1), it will not be enough to consider only the six point-separated operator products $\epsilon^{abc}A^{\mu c}(\xi)\partial^\nu A^b_\nu(0)$, etc. The expansion of one of these products gives further composite operators beyond the original six, since any operator of the right quantum numbers, including scale dimension, may mix in the Wilson expansion. In the language of renormalization, the renormalization of bare composite operators requires the introduction of extra operators as counterterms whose renormalization requires still other operators. This goes on until we find a (finite) set $\mathscr{C} = \{\mathcal{O}_i\}$ of operators with a corresponding renormalized set $\{N[\mathcal{O}i]\}$ such that the renormalization of any operator O_i in the set generates counterterms which are all members of the same set. We say that this set $\mathscr{C}$ is then closed under renormalization. The renormalization matrix $\mathbf{Z}$ satisfying

$$N[\mathcal{O}_j]Z_{ji} = \mathcal{O}_i, \tag{2.2}$$

is then the object of our calculation.

Thus, we shall have to enumerate all operators of the same canonical dimension which can contribute to the renormalization of the six composite operators. For SU(2), thanks to $\epsilon^{abc}\epsilon^{cde} = \delta^{ad}\delta^{be} - \delta^{ae}\delta^{bd}$, the internal symmetry indices are manageable, and we find the following set $\mathscr{C}$ of 25 operators closed under renormalization which includes the original six as a subset:

$$\begin{aligned}
O_1 &= N[A^{c\nu}A^{a\mu}A^c_\nu],\\
O_2 &= N[A^{a\nu}A^{c\mu}A^c_\nu],\\
O_3 &= N[n\cdot A^c n\cdot A^c A^{a\mu}],\\
O_4 &= N[n\cdot A^c n\cdot A^a A^{c\mu}],\\
O_5 &= N[n^\mu(n\cdot A^c)(A^c\cdot A^a)],\\
O_6 &= N[n^\mu n\cdot A^a A^c\cdot A^c],\\
O_7 &= N[n^\mu n\cdot A^c n\cdot A^a n\cdot A^c],\\
O_8 &= N[\epsilon^{abc}A^{c\nu}\partial^\mu A^b_\nu],\\
O_9 &= N[\epsilon^{abc}A^{c\mu}\partial^\nu A^b_\nu],\\
O_{10} &= N[\epsilon^{abc}A^{c\nu}\partial_\nu A^{b\mu}],\\
O_{11} &= N[\epsilon^{abc}n\cdot A^c\partial^\mu(n\cdot A^b)],\\
O_{12} &= N[\epsilon^{abc}n\cdot A^c n\cdot\partial A^{b\mu}],\\
O_{13} &= N[\epsilon^{abc}A^{c\mu}n\cdot\partial n\cdot A^b],\\
O_{14} &= N[\epsilon^{abc}n^\mu n\cdot A^c\partial\cdot A^b],\\
O_{15} &= N[\epsilon^{abc}n^\mu A^{c\nu}n\cdot\partial A^b_\nu],\\
O_{16} &= N[\epsilon^{abc}n^\mu A^{c\nu}\partial_\nu n\cdot A^b],\\
O_{17} &= N[\epsilon^{abc}(\partial^\mu\bar{c}^b)c^c],\\
O_{18} &= N[\epsilon^{abc}n^\mu(n\cdot\partial\bar{c}^b)c^c],\\
O_{19} &= \Box A^{\mu a},\\
O_{20} &= \partial^\mu(\partial\cdot A^a),
\end{aligned}$$

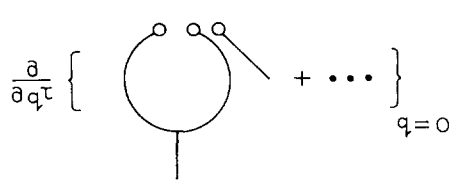

FIG. 2. Diagram for evaluating the Wilson coefficient in $AAA \to N[A\partial A]$.

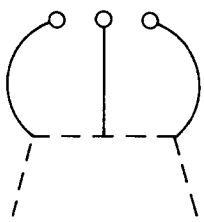

FIG. 3. Diagram for evaluating the Wilson coefficient in $AAA \to N[(\partial\bar{C})C]$.

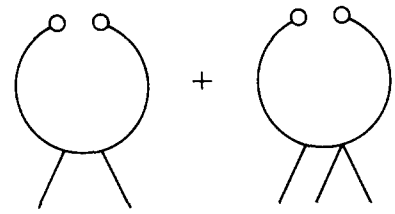

FIG. 4. Diagram for evaluating the Wilson coefficient in $A\partial A \to N[AAA]$.

$$O_{21} = (n\cdot\partial)(n\cdot\partial)A^{a\mu},$$
$$O_{22} = \partial^{\mu}(n\cdot\partial n\cdot A^{a}),$$
$$O_{23} = n_{\mu}\Box(n\cdot A^{a}),$$
$$O_{24} = n^{\mu}n\cdot\partial\partial\cdot A^{a},$$
$$O_{25} = n^{\mu}n\cdot\partial n\cdot\partial n\cdot A^{a}. \tag{2.3}$$

The operators O_1, O_2, O_8, O_9, O_{10}, O_{17}, O_{19}, and O_{20} occur in the field equation (2.1). The rest are generated by the need for renormalization to give a closed set; they all involve extra *direction dependence* in the form of the unit vector

$$n^{\mu} = x^{\mu}/\sqrt{x^2}. \tag{2.4}$$

It is a general feature of regularization by point separation that such extra direction-dependent terms need to be introduced.

III. LOW-ORDER CALCULATIONS

The calculation of the Wilson coefficients to lowest order of the expansion of the set of 25 operator products in terms of the same 25 composite operators presents no problem of principle, and involves merely application of well-known methods. In practice, the computation turns out to be extremely tedious. We shall classify the calculation into 13 types, each requiring the evaluation of Feynman diagrams of a definite genre.

A. *AAA→N[AAA]:*

We need to evaluate Feynman diagrams of the genre of Figs. 1(a) and 1(b). If we let

$$A^{c\nu}(x+\xi)A^{a\mu}(x)A^{b\lambda}(x-\xi) \sim \sum M^{cab\nu\mu\lambda}_{c'a'b'\nu'\mu'\lambda'} N[A^{c'\nu'}A^{\mu'a'}A^{\lambda'b'}], \tag{3.1}$$

then, to second order in g

$$\begin{aligned} M^{cab\nu\mu\lambda}_{c'a'b'\nu'\mu'\lambda'} = \frac{g^2}{16\pi^2}(-\ln\xi\mu)\Big\{\Big[&-(\tfrac{1}{2}+\tfrac{7}{4}\beta)(2\delta^{cb}\delta^{c'b'}-\delta^{ca'}\delta^{c'b}-\delta^{c'c}\delta^{ba'})g_{\lambda\nu}g_{\nu'\mu'}+(-\tfrac{1}{2}+\tfrac{1}{4}\beta)(2\delta^{cb}\delta^{c'a'}-\delta^{ca'}\delta^{c'b} \\ &-\delta^{c'c}\delta^{ba'})(g_{\nu\nu'}g_{\lambda\mu'}+g_{\lambda\nu'}g_{\nu\mu'})+\tfrac{3}{2}(\delta^{cc'}\delta^{ba'}-\delta^{ca'}\delta^{bc'})(g_{\nu\mu'}g_{\lambda\mu'}-g_{\lambda\nu'}g_{\nu\mu'})\Big]\delta^{ab'}g^{\mu\lambda'} \\ &+\Big[-(\tfrac{1}{2}+\tfrac{7}{4}\beta)(2\delta^{ab}\delta^{c'a'}-\delta^{aa'}\delta^{c'b}-\delta^{c'a}\delta^{ba'})g_{\lambda\mu}g_{\nu'\mu'}+(-\tfrac{1}{2}+\tfrac{1}{4}\beta)(2\delta^{ab}\delta^{c'a'}-\delta^{aa'}\delta^{c'b}-\delta^{c'a}\delta^{ba'})(g_{\mu\nu}g_{\lambda\mu'}+g_{\lambda\nu}g_{\mu\mu'}) \\ &+\tfrac{3}{2}(\delta^{aa'}\delta^{ba'}-\delta^{aa'}\delta^{bc'})(g_{\mu\nu}g_{\lambda\mu'}-g_{\lambda\nu}g_{\mu\mu'})\Big]\delta^{cb'}g^{\nu\lambda'}+\Big[-(\tfrac{1}{2}+\tfrac{7}{4}\beta)(2\delta^{ca}\delta^{c'a'}-\delta^{a'c}\delta^{c'a}-\delta^{c'c}\delta^{aa'})g_{\mu\nu}g_{\nu'\mu'} \\ &+(-\tfrac{1}{2}+\tfrac{1}{4}\beta)(2\delta^{ca}\delta^{c'a'}-\delta^{ca'}\delta^{c'b'}-\delta^{c'c}\delta^{a'c})(g_{\nu\nu}g_{\mu\mu'}+g_{\mu\nu}g_{\nu\mu'}) \\ &+\tfrac{3}{2}(\delta^{cc'}\delta^{aa'}-\delta^{ca'}\delta^{ac'})(g_{\nu\nu}g_{\mu\mu'}-g_{\mu\nu'}g_{\nu\mu'})\Big]\delta^{ab'}g^{\lambda\lambda'}\Big\}, \end{aligned} \tag{3.2}$$

where $\beta = 1-\alpha$, and α is the gauge parameter.

From this general expression (3.2), we can then obtain the expansion for the other operator products involving A's by making appropriate contractions.

B. *AAA→N[A∂A]*

This is obtained by differentiation at zero momentum of the Feynman graphs of the type illustrated in Fig. 2. The result is

$$\begin{aligned} A^{c\nu}(x+\xi)A^{\mu a}(x)A^{b\lambda}(x-\xi) \sim \frac{-g}{192\pi^2}\ln(\xi\mu)\{&(-6\beta+36)N\ [\epsilon^{cbb'}A^{\mu a}\partial^{\lambda}A^{\nu b'}]+(6\beta-36)N\ [\epsilon^{cbb'}A^{\mu a}\partial^{\nu}A^{\lambda b'}]\} \\ &+(-6\beta+36)N\ [\epsilon^{abb'}A^{\nu c}\partial^{\lambda}A^{\mu b'}]+(6\beta-36)N\ [\epsilon^{abb'}A^{\nu c}\partial^{\mu}A^{\lambda b'}] \\ &+(-6\beta+36)N\ [\epsilon^{cab'}A^{\lambda b}\partial^{\mu}A^{\nu b'}]+(6\beta-36)N\ [\epsilon^{cab'}A^{\lambda b}\partial^{\nu}A^{\mu b'}]. \end{aligned} \tag{3.3}$$

Once again, the expansion of the other operator products with greater indicial symmetry is obtained by contraction.

Here the lowest order calculation is of $O(g^1)$. One might wonder if a further calculation to $O(g^2)$ is necessary to give the proper input to the renormalization group calculation. We shall show in Sec. IV that this lowest order calculation is sufficient.

C. *AAA→N[(∂C)C]*

The expansion coefficient is given by Fig. 3, which gives an $O(g^3)$ contribution. We need not evaluate this contribution, however, for our purposes, by the argument that will be given in Sec. IV.

D. *A∂A→N[AAA]*

This is given by graphs like Figs. 4(a) and 4(b). Again, by consideration in Sec. IV, even though Figs. 4(a) and 4(b) are of $O(g^3)$, their contribution cannot be ignored now for the renormalization group calculation. The calculation here is particularly arduous. The result for $P^{cb\beta\lambda\alpha}_{c'b'\nu'\mu'\lambda'}$, where

$$A^{c\beta}(x+\xi)\partial^{\lambda}A^{ab}(x)\sim\sum P^{cb\beta\lambda\alpha}_{defa\sigma\tau}N[A^{d\alpha}A^{e\sigma}A^{f\tau}], \tag{3.4}$$

is

$$\begin{aligned}
&\frac{g^2}{192\pi^2}(-\ln\xi\mu)\{g_{\mu\tau}g_{\alpha\nu}g_{\sigma\lambda}[\delta^{ae}\delta^{df}(-8\beta^2+96\beta-288)+\delta^{af}\delta^{ed}(-8\beta^2-8\beta-162)+\delta^{ad}\delta^{ef}(4\beta^2-72\beta+120)]\\
&+g_{\mu\tau}g_{\alpha\lambda}g_{\sigma\nu}[\delta^{ae}\delta^{df}(24\beta^2-32\beta)+\delta^{af}\delta^{ed}(6\beta^2-56\beta+144)+\delta^{ad}\delta^{ef}(-4\beta^2-24\beta+144)]\\
&+g_{\mu\alpha}g_{\tau\sigma}g_{\nu\lambda}[\delta^{ae}\delta^{df}(16\beta^2-24\beta+48)+\delta^{af}\delta^{ed}(8\beta^2-72\beta+72)+\delta^{ad}\delta^{ef}(-4\beta^2-48\beta+24)]\\
&+g_{\mu\alpha}g_{\tau\nu}g_{\sigma\lambda}[\delta^{ae}\delta^{df}(16\beta^2-24\beta+48)+\delta^{af}\delta^{ed}(8\beta^2-104\beta+72)+\delta^{ad}\delta^{ef}(-4\beta^2-48\beta+24)]\\
&+g_{\mu\sigma}g_{\tau\alpha}g_{\nu\lambda}[\delta^{ae}\delta^{df}(8\beta^2-16\beta)+\delta^{af}\delta^{ed}(32\beta^3-88\beta+160)+\delta^{ad}\delta^{ef}(28\beta^2-72\beta+160)]\\
&+g_{\mu\sigma}g_{\tau\alpha}g_{\nu\lambda}[\delta^{ae}\delta^{df}(16\beta^2-72\beta+96)+\delta^{af}\delta^{ed}(16\beta^2+24\beta-32)+\delta^{ad}\delta^{ef}(4\beta^2+96\beta-128)]\\
&+g_{\mu\sigma}g_{\tau\lambda}g_{\sigma\nu}[\delta^{ae}\delta^{df}(16\beta^2-72\beta+96)+\delta^{af}\delta^{ed}(16\beta^2+24\beta-32)+\delta^{ad}\delta^{ef}(4\beta^2+96\beta-128)]\\
&+g_{\mu\nu}g_{\lambda\sigma}g_{\tau\alpha}[\delta^{ae}\delta^{df}(8\beta^2-16\beta+48)+\delta^{af}\delta^{ed}(32\beta^2-72\beta-16)+\delta^{ad}\delta^{ef}(28\beta^2-56\beta)]\\
&+g_{\mu\nu}g_{\tau\sigma}g_{\alpha\lambda}[\delta^{ae}\delta^{df}(24\beta^2-84\beta+96)+\delta^{af}\delta^{ed}(24\beta^2+56\beta-24)+\delta^{ad}\delta^{ef}(-44\beta^2+136\beta-120)]\\
&+g_{\nu\mu}g_{\tau\lambda}g_{\alpha\sigma}[\delta^{ae}\delta^{df}(-16\beta^2+88\beta-16)+\delta^{af}\delta^{ed}(-88\beta^2+296\beta-264)+\delta^{ad}\delta^{ef}(-68\beta^2+208\beta-168)]\\
&+g_{\mu\tau}g_{\alpha\sigma}g_{\nu\lambda}[\delta^{ae}\delta^{df}(8\beta^2-32\beta+192)+\delta^{af}\delta^{ed}(32\beta^2-104\beta+528)+\delta^{ad}\delta^{ef}(28\beta^2-72\beta+528)]\\
&+g_{\mu\tau}g_{\alpha\nu}g_{\sigma\lambda}[\delta^{ae}\delta^{df}(-8\beta^2+96\beta-288)+\delta^{af}\delta^{ed}(32\beta^2+24\beta-192)+\delta^{ad}\delta^{ef}(4\beta^2-72\beta+120)]\\
&+g_{\mu\lambda}g_{\alpha\tau}g_{\nu\sigma}[\delta^{ae}\delta^{df}(8\beta^2-16\beta+48)+\delta^{af}\delta^{ed}(32\beta^2-104\beta+48)+\delta^{ad}\delta^{ef}(28\beta^2-56\beta)]\\
&+g_{\mu\lambda}g_{\tau\sigma}g_{\alpha\nu}[\delta^{ae}\delta^{df}(24\beta^2-80\beta+96)+\delta^{af}\delta^{ed}(-24\beta^2-104\beta-72)+\delta^{ad}\delta^{ef}(-44\beta^2+136\beta-120)]\\
&+g_{\mu\lambda}g_{\tau\nu}g_{\alpha\sigma}[\delta^{ae}\delta^{df}(-16\beta^2+88\beta-96)+\delta^{af}\delta^{ed}(-88\beta^2+296\beta-264)+\delta^{ad}\delta^{ef}(-68\beta^2+208\beta-168)]\}.
\end{aligned} \tag{3.5}$$

Again, the other operator products with greater indicial symmetry are deducible from Eq. (3.5).

E. *A∂A→N[A∂A]*

This of $O(g^2)$ and follows from Figs. 5(a) and 5(b). Let

$$\epsilon^{acb}A^{c\beta}(x+\xi)\partial^{\lambda}A^{b\alpha}(x)\sim\sum\epsilon^{acb}R^{bc\beta\lambda\alpha}_{a'b'\alpha'\nu'\lambda'}N[A^{a'\alpha}\partial^{\nu'}A^{\lambda'b'}], \tag{3.6}$$

and the calculation gives to second order

$$\begin{aligned}
\epsilon^{abc}R^{bc\beta\lambda\alpha(2)}_{b'c'\mu''\tau\nu}=\;&\frac{g^2}{1536\pi^2}(-\ln\xi\mu)\epsilon^{ac'b'}[(16\beta^2-92\beta+68)g_{\mu''\tau}g_{\nu\beta}g_{\alpha\lambda}+(16\beta^2+20\beta-124)g_{\mu''\tau}g_{\nu\lambda}g_{\beta\alpha}\\
&+(16\beta^2-92\beta+68)g_{\mu''\tau}g_{\nu\alpha}g_{\beta\lambda}+(56\beta^2-4\beta-4)g_{\mu''\nu}g_{\tau\beta}g_{\alpha\lambda}+(96\beta^2-4\beta-4)g_{\mu''\nu}g_{\tau\alpha}g_{\beta\lambda}\\
&+(56\beta^2+20\beta-596)g_{\mu''\nu}g_{\tau\lambda}g_{\beta\alpha}+(-32\beta^2+84\beta-44)g_{\mu''\beta}g_{\tau\nu}g_{\alpha\lambda}\\
&+(-16\beta^2+76\beta-36)g_{\mu''\beta}g_{\tau\alpha}g_{\nu\lambda}+(-8\beta^2-492\beta+1004)g_{\mu''\beta}g_{\tau\lambda}g_{\alpha\nu}\\
&+(-32\beta^2+84\beta-44)g_{\mu''\alpha}g_{\tau\nu}g_{\beta\lambda}+(-8\beta^2+76\beta-36)g_{\mu''\alpha}g_{\tau\beta}g_{\nu\lambda}\\
&+(-8\beta^2+372\beta-948)g_{\mu''\alpha}g_{\tau\lambda}g_{\nu\beta}+(-32\beta^2+4\beta+52)g_{\mu''\lambda}g_{\tau\nu}g_{\beta\alpha}\\
&+(-8\beta^2-20\beta-36)g_{\mu''\lambda}g_{\tau\beta}g_{\nu\alpha}+(-16\beta^2-20\beta-36)g_{\mu''\lambda}g_{\tau\alpha}g_{\nu\beta}].
\end{aligned} \tag{3.7}$$

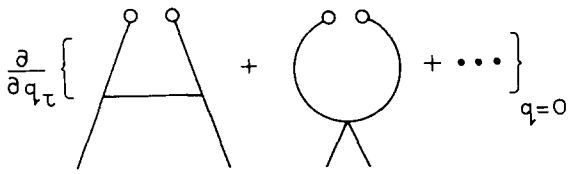

FIG. 5. Diagram for evaluating the Wilson coefficient in $A\partial A \to N[A\partial A]$.

F. $A\partial A \to N[(\partial\bar{C})C]$

The contribution comes from Fig. 6 and gives

$$\epsilon^{abc}(\partial^\mu A^{b\sigma})A^{c\beta} \sim \frac{g^2}{192\pi^2}(-\ln\xi\mu)[(\partial^b\bar{c})\,c^c g_{\mu\beta} \times(-8+6\beta+2\ \beta^2)+(\partial^\beta\bar{c}^b)c^c g_{\sigma\mu}(-4+3\beta-5\ \beta^2) + (\partial^\mu\bar{c}^b)c^c g_{\sigma\beta}(-2+6\beta-4\ \beta^2)]. \tag{3.8}$$

G. $AAA \to N[(\partial\bar{C})C]$

The contribution from Fig. 7 is of $O(g^3)$, and can be ignored in this case.

H. $(\partial\bar{C})C \to N[A\partial A]$

The $O(g^2)$ contribution from the sum of Figs. 8(a) and 8(b) vanishes. Higher order contributions turn out to be ignorable.

I. $(\partial\bar{C})C \to N[(\partial\bar{C})C]$

The result from Fig. 9 is

$$\epsilon^{abc}\partial^\mu\bar{c}^b(x)c^b(x+\xi)\sim\frac{-g^2}{32\pi^2}(-\ln\xi\mu) \times(-5+5\beta)N[\epsilon^{ab'c'}(\partial^\mu\bar{c}^{b'})c^{c'}]. \tag{3.9}$$

J. $(\partial\bar{C})C \to N[AAA]$

The $O(g^3)$ contribution from Fig. 10 vanishes.

K. $AAA \to A$; $A\partial A \to A$; $(\partial\bar{C})C \to A$

We shall find in Sec. IV that the values of these contributions from Figs. 11, 12, and 13, respectively, need not be known.

It is then simple arithmetic to calculate the γ matrix for the set $\mathscr{C}$. We display the first 18×18 elements as matrix I in Table I.

The 19–25 operators involve only single A's. The only γ matrix elements of relevance to us from 19 to 25 are the diagonal ones, which involve Z_3's and so need no further elaboration.

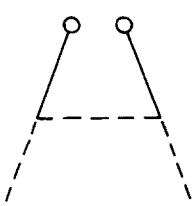

FIG. 6. Diagram for evaluating the Wilson coefficient in $A\partial A \to N[(\partial\bar{C})C]$.

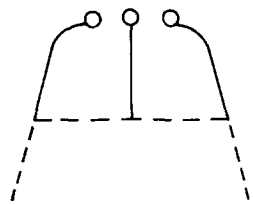

FIG. 7. Diagram for evaluating the Wilson coefficient in $AAA \to N[(\partial\bar{C})C]$.

IV. HOMOGENIZATION

Now that we have the lowest order values for $\mathbf{Z}$ and γ, we shall have to use them as input to the renormalization group equation

$$\mu\frac{d}{d\mu}Z_{ik} = -\sum_j \gamma_{ij}Z_{jk}. \tag{4.1}$$

This equation is formally solved by

$$Z_{ik}(g,\alpha,t) = T\left\{\exp\int_0^t dt'\gamma[\bar{g}(t'),\bar{\alpha}(t')]\right\}_{ij} Z_{jk}(\bar{g},\bar{\alpha},0). \tag{4.2}$$

In practice, this solution is hardly feasible in our present case. The lowest order calculations for γ are not homogeneous in g, i.e., some elements begin with $O(g^1)$, others with $O(g^2), O(g^3)$, etc. Thus, the $\bar{g}$ dependence cannot be factored out of the T-ordered exponential, and a computation of that object becomes an impossible proposition for the 25×25 γ matrix we have.

Instead, we have to define other matrices $\mathbf{Z}^h$ and γ^h such that their lowest order contributions are homogeneous in g, and then proceed to solve the equation by diagonalization. We shall describe below this process of "*homogenization*."[4]

When the low order expressions for γ are substituted into Eq. (4.1), we have

$$\mu\frac{d}{d\mu}Z_{ik} = -\sum_j g^{m_{ij}}\gamma_{ij}^{(0)}(\alpha)Z_{jk}. \tag{4.3}$$

Z_{ik} is essentially the coefficient of the ith operator on the right-hand side in the expansion of the k th operator. The fact that we have found the finite set of 25 operators closed under renormalization means Z is a square (25×25) matrix. In the Yang–Mills theory we are considering, the diagonal elements of γ always start with $O(g^2)$. Thus, we shall homogenize $\mathbf{Z}$ and γ to give a new γ^h such that all elements of γ^h start in the second order.

When the γ matrix contains terms *both* of $O(g^1)$ and $O(g^2)$, then both give singular contributions to

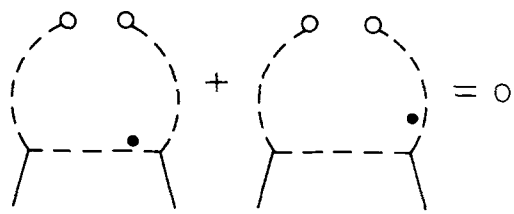

FIG. 8. Diagram for evaluating the Wilson coefficient in $(\partial\bar{C})C \to N[A\partial A]$.

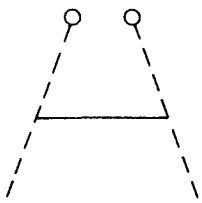

FIG. 9. Diagram for evaluating the Wilson coefficient in $(\partial\bar{C})C\to N[(\partial\bar{C})C]$.

$$\int_0^t dt'\gamma[\bar{g}(t'),\bar{\alpha}(t')], \tag{4.4}$$

at the upper limit of the integration as $t\to\infty$, and we simply do not know how to evaluate explicitly the T-ordered exponential

$$T\left\{\exp\int_0^t dt'\gamma[\bar{g}(t'),\bar{\alpha}(t')]\right\}, \tag{4.5}$$

needed for the solution to the renormalization group equation. If all the terms are of $O(g^2)$ and higher, then we can factor out the g^2:

$$\gamma(g,\alpha)=g^2\gamma^{(0)}(g,\alpha), \tag{4.6}$$

and the only singular part of the integral (4.4) comes from the g^2 factor. The T-ordered exponential (4.5) can then be evaluated by calculating eigenvalues of $\gamma^{(0)}$ in the usual way.

To explain the homogenization process, let us simplify the discussion by pretending that there are not 25 operators but only five, defined by the following:

(1) AAA is the generic operator with three A, i.e., O_1 through O_7.

(2) $[AA]_1$ is the generic operator with two A, whose operators contribute to the expansion of AAA, i.e., $O(8,10,11,12,15,16)$.

(3) $[AA]_2$ is the generic operator with two A, whose operators do not contribute to the expansion of AAA, i.e., $O(9,13,14)$.

(4) CC is the generic operator with two C fields, i.e., O_{17} and O_{18}.

(5) A is the generic operator with only one A, i.e., O_{19} through O_{25}.

The γ matrix thus has the simplified look:

$$\gamma=\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ & AAA & [AA]_1 & [AA]_2 & CC & A \\ 1,AAA & \Gamma_{11}g^2 & \Gamma_{12}g^3 & \Gamma_{13}g^3 & 0 & 0 \\ 2,[AA]_1 & \Gamma_{21}g & \Gamma_{22}g^2 & \Gamma_{23}g^3 & 0 & 0 \\ 3,[AA]_2 & 0 & \Gamma_{32}g^2 & \Gamma_{33}g^2 & 0 & 0 \\ 4,CC & \Gamma_{41}g^3 & \Gamma_{42}g^2 & \Gamma_{43}g^2 & \Gamma_{44}g^2 & 0 \\ 5,A & \Gamma_{51}g^2 & \Gamma_{52}g & \Gamma_{53}g & \Gamma_{54}g & \Gamma_{55}g^2 \end{array}, \tag{4.7}$$

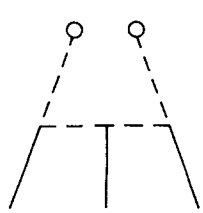

FIG. 10. Diagram for evaluating the Wilson coefficient in $(\partial\bar{C})C\to N[AAA]$.

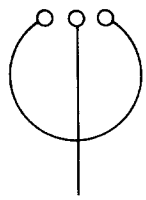

FIG. 11. Diagram for evaluating the Wilson coefficient in $AAA\to\partial\partial A$.

where the Γ_{ij} are numerical constants.

We first consider the "dangerous" matrix element $\Gamma_{21}g$. The component renormalization group equation in which it occurs is, for any k,

$$\mu\frac{d}{d\mu}Z_{2k}=-(\Gamma_{21}gZ_{1k}+\Gamma_{22}g^2Z_{2k}+\Gamma_{23}g^2Z_{3k}). \tag{4.8}$$

We multiply Eq. (4.8) by g and use the product rule to give

$$g\mu\frac{d}{d\mu}(Z_{2k})=\mu\frac{d}{d\mu}(gZ_{2k})-\beta(g)Z_{2k}, \tag{4.9}$$

and so, with $\beta(g)=-bg^3$,

$$\mu\frac{d}{d\mu}(gZ_{2k})=-[\Gamma_{21}g^2Z_{1k}+(\Gamma_{22}g^2+bg^2)(gZ_{2k})+\Gamma_{23}g^2gZ_{3k}]. \tag{4.10}$$

Thus, the homogenized $\mathbf{Z}$ and γ element from Eq. (4.10) is

$$\begin{aligned} Z^h_{2k}&=gZ_{2k},\\ \gamma^h_{21}&=\Gamma_{21}g^2,\\ \gamma^h_{22}&=\Gamma_{22}g^2+bg^2. \end{aligned} \tag{4.11}$$

Now consider the equation for Z_{3k}:

$$\begin{aligned} \mu\frac{d}{d\mu}Z_{3k}&=-(\Gamma_{32}g^2Z_{2k}+\Gamma_{33}g^2Z_{3k})\\ &=-(\Gamma_{32}gZ^h_{2k}+\Gamma_{33}g^2Z_{3k}). \end{aligned} \tag{4.12}$$

By the need to homogenize $Z_{2k}\to Z^h_{2k}$, a dangerous coefficient $\Gamma_{32}g$ is induced in Eq. (4.12). We must then eliminate it by the same method to give

$$\mu\frac{d}{d\mu}(gZ_{3k})=-[\Gamma_{32}g^2Z^h_{2k}+(\Gamma_{33}g^2+bg^2)(gZ_{3k})], \tag{4.13}$$

and hence

$$Z^h_{3k}=gZ_{3k},\quad \gamma^h_{33}=(\Gamma_{33}+b)g^2. \tag{4.14}$$

Exactly the same "inducement" effect forces us to change the equation for Z_{4k} to

$$\mu\frac{d}{d\mu}(gZ_{4k})=-[\Gamma_{41}g^3gZ_{1k}+\Gamma_{42}g^2(gZ_{2k})+\Gamma_{43}g^2(gZ_{3k})+(\Gamma_{43}+b)g^2(gZ_{4k})], \tag{4.15}$$

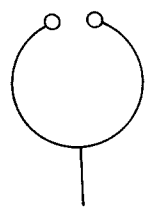

FIG. 12. Diagram for evaluating the Wilson coefficient in $A\partial A\to\partial\partial A$.

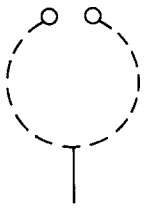

FIG. 13. Diagram for evaluating the Wilson coefficient in $(\partial\bar{C})C \to \partial\partial A$.

and so

$$Z^h_{4k} = gZ_{4k}, \quad \gamma^h_{41} = \Gamma_{41}g^4, \quad \gamma^h_{43} = (\Gamma_{43} + b)g^2. \tag{4.16}$$

For Z_{1k}, the effect is

$$\mu\frac{d}{d\mu}Z_{1k} = -[\Gamma_{11}g^2Z_{1k} + \Gamma_{12}g^2(gZ_{2k}) + \Gamma_{13}g^2(gZ_{3k})]. \tag{4.17}$$

The need to homogenize Z_{2k} and Z_{3k} absorbs a factor of g from γ_{12} and γ_{13}, and so one *must* know the value of Γ_{12} and Γ_{13}, even though they are coefficients of g^3, which one might think can be discarded as being of higher order. The homogenized γ elements are

$$\gamma^h_{12} = \Gamma_{12}g^2, \quad \gamma_{13} = \Gamma_{13}g^2. \tag{4.18}$$

Finally, we consider the equation for Z_{5k}:

$$\begin{aligned}\mu\frac{d}{d\mu}Z_{5k} &= -(\Gamma_{51}g^2Z_{1k} + \Gamma_{52}gZ_{2k} + \Gamma_{53}gZ_{3k} + \Gamma_{54}gZ_{4k} + \Gamma_{55}g^2Z_{55}) \\ &= -(\Gamma_{51}g^2Z_{1k} + \Gamma_{52}Z^h_{2k} + \Gamma_{53}Z^h_{3k} + \Gamma_{54}Z^h_{4k} + \Gamma_{55}g^2Z_{55}),\end{aligned} \tag{4.19}$$

where the last line arises from the need to homogenize Z_{2k}, Z_{3k}, and Z_{4k}. We therefore have to multiply Eq. (4.19) by g^2 to give

$$\mu\frac{d}{d\mu}(g^2Z_{5k}) = -[\Gamma_{51}g^4Z_{1k} + \Gamma_{52}g^2Z^h_{2k} + \Gamma_{53}g^2Z^h_{3k} + \Gamma_{54}g^2Z^h_{4k} + (\Gamma_{55}g^2 + 2bg^2)(g^2Z_{55})], \tag{4.20}$$

and so we identify

$$\begin{aligned}Z^h_{5k} &= g^2Z_{5k}, \\ \gamma_{51} &= \Gamma_{51}g^4, \\ \gamma_{52} &= \Gamma_{52}g^2, \\ \gamma_{53} &= \Gamma_{53}g^2, \\ \gamma_{54} &= \Gamma_{54}g^2, \\ \gamma_{55} &= (\Gamma_{55} + 2b)g^2.\end{aligned} \tag{4.21}$$

Our homogenization process is now complete, and we have

TABLE I. Matrix I: an 18×18 matrix.[a]

Row	Column 1	2	3	4	5	6
1	$U(-6-34\beta)$	$U(-11-24\beta)$	$U(-2-7\beta)$	$U(-1-\frac{1}{2}\beta)$	0	0
2	$U(-6+8\beta)$	$U(5+19\beta)$	0	$U(1+\frac{7}{2}\beta)$	0	0
3	0	0	$U(4+\beta)$	$U(-4+\frac{1}{2}\beta)$	0	0
4	0	0	$U(-8+\beta)$	$U(\frac{3}{2}\beta)$	0	0
5	0	0	$U(2+7\beta)$	$U(1+\frac{7}{2}\beta)$	$U(+5+19\beta)$	$U(-6+8\beta)$
6	0	0	$U(-2-7\beta)$	$U(-2-7\beta)$	$U(-11-24\beta)$	$U(-6-34\beta)$
7	0	0	0	0	0	0
8	$Y(-12\beta+72)$	$Y(6\beta-36)$	0	0	0	0
9	0	0	0	0	0	0
10	$Y(12\beta-72)$	$Y(-6\beta+36)$	0	0	0	0
11	0	0	$Y(-12\beta+72)$	$Y(6\beta-36)$	0	0
12	0	0	$Y(12\beta-72)$	$Y(-6\beta+36)$	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	$Y(6\beta-36)$	$Y(-12\beta+72)$
16	0	0	0	0	$Y(-6\beta+36)$	$Y(12\beta-72)$
17	0	0	0	0	0	0
18	0	0	0	0	0	0

TABLE I. (continued)

Row	Column 7	8	9	10	11	12
1	0	$Y(-12\beta^2+12\beta+264)$	$Y(-16\beta^2+104\beta-50)$	$Y(-24\beta^2+236\beta-152)$	$Z(36\beta^2-144\beta+672)$	$Z(-124\beta\beta^2+416\beta-354)$
2	0	$Y(68\beta^2-128\beta+456)$	$Y(36\beta^2+28\beta-8)$	$Y(16\beta+100)$	$Z(136\beta^2-352\beta+1184)$	$Z(-24\beta^2+52\beta+92)$
3	0	0	0	0	$Z(8\beta^2-224\beta-114)$	$Z(40\beta^2+48\beta-208)$
4	0	0	0	0	$Z(80\beta^2+44\beta-416)$	$Z(48\beta^2-208\beta-322)$
5	$U(2+7\beta)$	0	0	0	$Z(-56\beta^2+124\beta-320)$	$Z(112\beta^2-212\beta+1036)$
6	$U(-4-14\beta)$	0	0	0	$Z(-248\beta^1+8325-678)$	$Z(-88\beta^1+336)$
7	$U(-4+2\beta)$	0	0	0	0	0
8	0	$Z(-360\beta^2+48\beta+2336)$	$Z(-336\beta^2-56\beta+688)$	$Z(-464\beta^2-56\beta+688)$	$Z(-56\beta^2-20\beta+536)$	$Z(-96\beta^2+4\beta+4)$
9	0	$Z(232\beta^2-68\beta-12)$	$Z(224\beta^2-8\beta-800)$	$Z(208\beta^2-972\beta+1152)$	$Z(32\beta^2-4\beta-52)$	$Z(32\beta^2-84\beta+44)$
10	0	$Z(-88\beta^2+28\beta+396)$	$Z(-72\beta^2+88\beta+768)$	$Z(-80\beta^2+652\beta-1184)$	$Z(-16\beta^2-20\beta+124)$	$Z(-18\beta^2+92\beta-68)$
11	0	0	0	0	$Z(16\beta^2+120\beta-56)$	$Z(32\beta^2-56\beta+72)$
12	0	0	0	0	$Z(24\beta^2-152\beta+72)$	$Z(16\beta^2-572\beta-968)$
13	0	0	0	0	$Z(24\beta^2+40\beta+72)$	$Z(16\beta^2-448\beta+984)$
14	0	0	0	0	$Z(64\beta^2+168\beta+88)$	$Z(64\beta^2-88\beta-18)$
15	0	0	0	0	$Z(-152\beta^2+8\beta+8)$	$Z(-112\beta^2-16\beta+600)$
16	0	0	0	0	$Z(-32\beta^2+184\beta-132)$	$Z(-32\beta^1+72\beta+56)$
17	0	$Y(-20+33\beta-19\beta^2)$	$Y(-26+24\beta-22\beta^2)$	$Y(-36+33\beta-\beta^2)$	$Y(-2+6\beta-4\beta^2)$	$Y(-8+6\beta+2\beta^2)$
18	0	0	0	0	$Y(-12+8\beta-3\beta^2)$	$Y(-6+9\beta-9\beta^2)$

TABLE I. (continued)

Row	Col 13	14	15	16	17	18
1	$Z(-124\beta^2+416\beta-336)$	0	0	0	0	0
2	$Z(-24\beta^2+162\beta-128)$	0	0	0	0	0
3	$Z(40\beta^2+16\beta-208)$	0	0	0	0	0
4	$Z(48\beta^2-176\beta-514)$	0	0	0	0	0
5	$Z(96\beta^2-356\beta+744)$	$Y(36\beta^2+28\beta-8)$	$Y(68\beta^2-128\beta+456)$	$Y(16\beta+100)$	0	0
6	$Z(-88\beta^2+172\beta+336)$	$Y(-16\beta^2+10\beta-50)$	$Y(-12\beta^2+12\beta+264)$	$Y(-24\beta^2+236\beta-152)$	0	0
7	0	0	0	0	0	0
8	$Z(-56\beta^2+4\beta+4)$	0	0	0	0	0
9	$Z(32\beta^2-848\beta+44)$	0	0	0	0	0
10	$Z(-16\beta^2+92\beta-68)$	0	0	0	0	0
11	$Z(16\beta^2-56\beta+72)$	0	0	0	0	0
12	$Z(24\beta^2-352\beta+984)$	0	0	0	0	0
13	$Z(24\beta^2+416\beta-968)$	0	0	0	0	0
14	$Z(64\beta^2-128\beta-8)$	$Z(224\beta^2-8\beta-800)$	$Z(232\beta^2-68\beta-12)$	$Z(208\beta^2-872\beta+1152)$	0	0
15	$Z(-152\beta^2-16\beta+600)$	$Z(-360\beta^2-56\beta+688)$	$Z(-360\beta^2+48\beta+2336)$	$Z(-464\beta^2-56\beta+688)$	0	0
16	$Z(-32\beta^2+72\beta+56)$	$Z(-72\beta^2+88\beta+768)$	$Z(-88\beta^2+28\beta+396)$	$Z(-80\beta^2+852\beta-1184)$	0	0
17	$Y(-4+3\beta-5\beta^2)$	0	0	0	$U(5/2-5/2\beta)$	0
18	$Y(-10+12\beta-2\beta^2)$	$Y(-26+24\beta-22\beta^2)$	$Y(-20+33\beta-2\beta^2)$	$Y(-38+33\beta-\beta^2)$	0	$U(5/2-5/2\beta)$

[a] $U=1/16\pi^2$, $Y=1/192\pi^2$, $Z=1/1536\pi^2$, and $\beta=1-\alpha$, with α as the gauge parameter.

TABLE II. Matrix II: an 18×18 matrix.[a]

Row	Column 1	2	3	4	5	6
1	$U(-6-34\beta)$	$U(-11-24\beta)$	$U(-2-7\beta)$	$U(-1-\frac{1}{2}\beta)$	0	0
2	$U(-6+8\beta)$	$U(5+19\beta)$	0	$U(1+\frac{7}{2}\beta)$	0	0
3	0	0	$U(4+\beta)$	$U(-4+\frac{1}{2}\beta)$	0	0
4	0	0	$U(-8+\beta)$	$U(\frac{1}{2}\beta)$	0	0
5	0	0	$U(2+7\beta)$	$U(1+\frac{7}{2}\beta)$	$U(+5+19\beta)$	$U(-6+8\beta)$
6	0	0	$U(-2-7\beta)$	$U(-2-7\beta)$	$U(-11-24\beta)$	$U(-6-34\beta)$
7	0	0	0	0	0	0
8	$Y(-12\beta+72)$	$Y(6\beta-36)$	0	0	0	0
9	0	0	0	0	0	0
10	$Y(12\beta-72)$	$Y(-6\beta+36)$	0	0	0	0
11	0	0	$Y(-12\beta+72)$	$Y(6\beta-36)$	0	0
12	0	0	$Y(12\beta-72)$	$Y(-6\beta+36)$	0	0
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	$Y(6\beta-36)$	$Y(-12\beta+72)$
16	0	0	0	0	$Y(-6\beta+36)$	$Y(12\beta-72)$
17	0	0	0	0	0	0
18	0	0	0	0	0	0

TABLE II. (continued)

Row	Column 7	8	9	10	11	12
1	0	$Y(-12\beta^2+12\beta+264)$	$Y(-16\beta^2+104\beta-50)$	$Y(-267+236\beta-152)$	$Z(36\beta^2-144\beta+672)$	$Z(-124\beta^2+4163-354)$
2	0	$Y(68\beta^2-128\beta+456)$	$Y(36\beta^2+28\beta-8)$	$Y(16\beta+100)$	$Z(136\beta^2-352\beta+1184)$	$Z(-24\beta v^2+52\beta+92)$
3	0	0	0	0	$Z(8\beta^2-224\beta-114)$	$Z(40\beta^2+48\beta-208)$
4	0	0	0	0	$Z(80\beta^2+44\beta-416)$	$Z(48\beta^2-208\beta-322)$
5	$U(2+7\beta)$	0	0	0	$Z(-56\beta^2+124\beta-320)$	$Z(112\beta^2-212\beta+1036)$
6	$U(-4-14\beta)$	0	0	0	$Z(-248\beta^2+832\beta-678)$	$Z(-88\beta^2+272\beta+336)$
7	$U(-4+2\beta)$	0	0	0	0	0
8	0	$W+Z(-360\beta^2+488+2336)$	$Z(-336\beta^2-56\beta+688)$	$Z(-464\beta^2-56\beta+688)$	$Z(-56\beta^2-20\beta+596)$	$Z(-96\beta^2+4\beta+4)$
9	0	$Z(232\beta^2-68\beta-12)$	$Z(224\beta^2-8\beta-800)$	$Z(208\beta^2-872\beta-1152)$	$Z(32\beta^2-4\beta-52)$	$Z(32\beta^2-848\beta+44)$
10	0	$Z(-88\beta^2+28\beta+396)$	$Z(-72\beta^2+885+768)$	$Z(-80\beta^2+852\beta-1184)+W$	$Z(-16\beta^2-20\beta+124)$	$Z(-16\beta^2+92\beta-68)$
11	0	0	0	0	$Z(16\beta^2+120\beta-56)+W$	$Z(32\beta^2-56\beta+72)$
12	0	0	0	0	$Z(24\beta^2-152\beta+72)$	$Z(16\beta^2+12\beta-968)+W)$
13	0	0	0	0	$Z(24\beta^2+40\beta+72)$	$Z(16\beta^2-448\beta+984)$
14	0	0	0	0	$Z(64\beta^2+168\beta+88)$	$Z(64\beta^2-88\beta-18)$
15	0	0	0	0	$Z(-152\beta^2+8\beta+8)$	$Z(-112\beta^2-16\beta+600)$
16	0	0	0	0	$Z(-32\beta^2+184\beta-132)$	$Z(-32\beta^2+72\beta+56)$
17	0	$Y(-20+33\beta-19\beta^2)$	$Y(-26+24\beta-22\beta^2)$	$Y(-36+33\beta-\beta^2)$	$Y(-2+6\beta-4\beta^2)$	$Y(-8+6\beta+2\beta^2)$
18	0	0	0	0	$Y(-12+8\beta-3\beta^2)$	$Y(-6+9\beta-9\beta^2)$

TABLE II. (continued)

Row	Column 13	14	15	16	17	18
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	$Y(36\beta^2+28\beta-8)$	$Y(68\beta^2+28\beta-8)$	$Y(16\beta+100)$	0	0
6	0	$Y(-16\beta^2+10\beta-50$	$Y(-12\beta^2+125\beta+264)$	$Y(-24\beta^2+236\beta-152)$	0	0
7	0	0	0	0	0	0
8	$Z(-56\beta^2+4\beta+4)$	0	0	0	0	0
9	$Z(32\beta^2-848\beta+44)$	0	0	0	0	0
10	$Z(-16\beta^2+92\beta-68)$	0	0	0	0	0
11	$Z(16\beta^2-56\beta+72)$	0	0	0	0	0
12	$Z(24\beta^2-352\beta+984)$	0	0	0	0	0
13	$Z(24\beta^2+416\beta-968)$	0	0	0	0	0
14	$Z(64\beta^2-128\beta-8)$	$Z(224\beta^2-8\beta-800)$	$Z(232\beta^2-68\beta-12)$	$Z(208\beta^2-872\beta+1152)$	0	0
15	$Z(-152\beta^2-16\beta+600)$	$Z(-360\beta^2-56\beta+688)$	$Z(-360\beta^2+485+2336)+W$	$Z(-464\beta^2-56+688)$	0	0
16	$Z(-32\beta^2+72\beta+56)$	$Z(-72\beta^2+88+768)$	$Z(-88\beta^2+28\beta+396)$	$Z(-80\beta^2+852\beta-1184)+W$	0	0
17	$Y(-4+3\beta-5\beta^2)$	0	0	0	$U(5/2-5/2\beta)$	0
18	$Y(-10+12\beta-2\beta^2)$	$Y(-26+24\beta-22\beta^2)$	$Y(-20+33\beta-2\beta^2)$	$Y(-38+33\beta-\beta^2)$	0	$U(5/2-5/2\beta)$

[a]The new quantity W here means $+b$.

the differential equation

$$\mu \frac{d}{d\mu}\mathbf{Z}^h = -\gamma^h \mathbf{Z}^h, \tag{4.22}$$

where

$$Z^h_{1k} = Z_{1k}, \quad Z^h_2 = gZ_{2k}, \quad Z^h_3 = gZ_{3k}, \quad Z^h_4 = gZ_{4k}, \quad Z^h_5 = g^2 Z_{5k}, \tag{4.23}$$

and the γ^h matrix is

	1	2	3	4	5
	AAA	$[AA]_1$	$[AA]_2$	CC	A
1,AAA	$\Gamma_{11}g^2$	$\Gamma_{22}g^2$	$\Gamma_{13}g^2$	0	0
2,$[AA]_1$	$\Gamma_{21}g^2$	$(\Gamma_{22}+b)g^2$	$\Gamma_{23}g^2$	0	0
3,$[AA]_2$	0	$\Gamma_{32}g^2$	$(\Gamma_{33}+b)g^2$	0	0
4,CC	$\Gamma_{41}g^4$	$\Gamma_{42}g^2$	$\Gamma_{43}g^2$	$(\Gamma_{44}+b)g^2$	0
5,A	$\Gamma_{51}g^4$	$\Gamma_{52}g^2$	$\Gamma_{53}g^2$	$\Gamma_{54}g^2$	$(\Gamma_{55}+2b)g^2$

(4.24)

By comparing Eq. (4.24) with (4.7), we see that all matrix elements in γ^h are at least of $O(g^2)$. All $O(g^1)$ terms have been elevated to being $O(g^2)$. The entries γ_{12} and γ_{13} were of $O(g^3)$, but are $O(g^2)$ in γ^h and so have to be known. The entries γ_{41} and γ_{51} were $O(g^3)$ and $O(g^2)$, respectively, in γ, but have been downgraded both to $O(g^4)$ in γ^h, and so the values of Γ_{41} and Γ_{51} need not be known for solving the equation (4.22).

In actuality, we perform this procedure on the full 25×25 γ matrix, and we obtain the γ^h matrix, the first 18×18 part of which is displayed as matrix II in Table II.

V. FINITE FIELD EQUATION

The equation (4.22) for $\mathbf{Z}^h$ can be solved in the usual manner by calculating the eigenvalues of the matrix $\gamma^h_{(0)}$, where

$$\gamma^h = \gamma^h_{(0)} g^2 \tag{5.1}$$

to leading order. Because of the zeroes above the diagonal in columns CC and A [see Eq. (4.24)], the eigenvalues in that sector decouple and are just the diagonal entries. The rest of the eigenvalues are obtained by computer on the first 16×16 part of the $\gamma^h_{(0)}$ matrix:

$$\begin{aligned}
&\lambda_1 = 0.69, && \text{IMAG} = 0.00\\
&\lambda_2 = -0.02, && \text{IMAG} = 0.21,\\
&\lambda_3 = -0.02, && \text{IMAG} = -0.21,\\
&\lambda_4 = -0.28, && \text{IMAG} = 0.00\\
&\lambda_5 = -0.11, && \text{IMAG} = 0.00\\
&\lambda_6 = 0.69, && \text{IMAG} = 0.00,\\
&\lambda_7 = -0.02, && \text{IMAG} = 0.21,\\
&\lambda_8 = -0.02, && \text{IMAG} = -0.21,\\
&\lambda_9 = -0.28, && \text{IMAG} = 0.00,\\
&\lambda_{10} = -0.07, && \text{IMAG} = 0.00,\\
&\lambda_{11} = -0.11, && \text{IMAG} = 0.00,\\
&\lambda_{12} = -0.23, && \text{IMAG} = 0.00,\\
&\lambda_{13} = 0.14, && \text{IMAG} = 0.00,\\
&\lambda_{14} = -0.01, && \text{IMAG} = 0.00,\\
&\lambda_{15} = 0.09, && \text{IMAG} = 0.00,\\
&\lambda_{16} = -0.07, && \text{IMAG} = 0.00.
\end{aligned} \tag{5.2}$$

The zeroes in the last two columns of Eq. (4.24) mean that two more (multiple) eigenvalues are given by simply the 17–25 diagonal entries:

$$\begin{aligned}
&\lambda_{17} = \lambda_{18} = 0.16, \quad \text{IMAG} = 0,\\
&\lambda_{19} = \lambda_{20} = \lambda_{21} = \lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0.18,\\
&\qquad\qquad \text{IMAG} = 0.
\end{aligned} \tag{5.3}$$

The smallest of these eigenvalues is

$$\lambda_{\min} = -0.28, \quad \text{IMAG} = 0.00, \tag{5.4}$$

and so we conclude that the behavior of the $\mathbf{Z}$ matrix, or the Wilson coefficients E, is

$$E(\xi) \underset{\xi\to0}{\sim} (\ln\xi)^{+0.28/2b}. \tag{5.5}$$

We invert the operator product expansion and obtain the composite normal-product current in point-separated form to be

$$N[A^{c\nu}A^{a\mu}A^{\ c}_{\nu}] = \lim_{\xi\to0}\{(\ln\xi)^{-0.28/2b} \times [A^{c\nu}(x+\xi)A^{a\mu}(x)A^{\ c}_{\nu}(x-\xi)+\ldots]\}. \tag{5.6}$$

Combining the results in Eqs. (2.7) and (5.6), we obtain what we have promised in the Introduction, namely, an equation for the Yang–Mills field in which the source current is written in terms of elementary operators at separated points multiplied by a Wilson coefficient whose asypmtotic behavior is known:

$$-\left\{\left[1+\frac{1}{\alpha}(1+f_4)\right]g^{\mu\nu}\Box-\partial^\mu\partial^\nu\right\}A^{\nu a}$$

$$=-(1+f_3)g^2\lim_{\xi\to 0}\{(\ln\xi)^{-0.28/2b}[A^{c\nu}(x+\xi)A^{a\mu}(x)$$

$$\times A^{c}_{\nu}(x-\xi)+\epsilon^{abc}A^{\mu c}(x+\xi)\partial^\nu A^{b}_{\nu}(x)+\ldots]$$

$$+\ldots\}+(1-s)^2M^2A^{a\mu}. \tag{5.7}$$

[1]R.A. Brandt, W.C. Ng, and W.B. Yeung, Phys. Rev. D **19**, 503 (1979).
[2]J. Lowenstein, Nucl. Phys. B **96**, 189 (1975).
[3]R.A. Brandt, W.C. Ng, and K. Young, Phys. Rev. D **15**, 2885 (1977).
[4]This was done in the context of the ϕ^4 theory in Ref. 1.

Superconformal group and curved fermionic twistor space

J. Lukierski[a)]
International Centre for Theoretical Physics, Trieste, Italy

(Received 27 November 1978)

It is shown that the superconformal transformations describe the isometry group of curved fermionic twistor space, with suitably generalized Fubini–Study Hermitian metric. Further, such a geometry is applied to derive a two-dimensional supersymmetric quark–twistor string model, described by the fermionic nonlinear SU(2,1;1)-invariant σ-model.

I. INTRODUCTION

Geometrization of supersymmetry leads to the introduction of superspace—a linear vector space with independent commuting (bosonic) and anticommuting Grassmann (fermionic) coordinates. However, if we look into the structure of superalgebra, we can conclude that fermionic degrees of freedom are more fundamental than the bosonic ones. Indeed, in the case of graded de Sitter [1,2] and graded conformal [1,3–5] groups, performing twice the supersymmetry transformations generated by supercharges (supertranslations), one gets *all* the transformations of the bosonic sector. It seems possible, therefore, that one can build super-de Sitter or superconformal geometry only with *primary fermionic degrees of freedom*—and introduce all other coordinates as composite geometric objects. It is interesting to mention here that a similar situation occurs in hadrodynamics: Fundamental building blocks are *fermionic quarks*, and all observable degrees of freedom are their composites.

In the first part of this paper we show that one can introduce curved Grassmann twistor space $\widetilde{T}_4(c)$ with the group of isometries described by the Wess–Zumino graded conformal group. [5] We have chosen superconformal geometry because:

(i) The graded conformal algebras SU(2,2;N) [3,4] seem to be group-theoretically the best candidates for the unification of geometric and internal degrees of freedom.

(ii) Fundamental representations of the conformal group are known as twistors (see, e.g., Ref. 6). In the case of commuting twistor components (bosonic twistor geometry) the idea of constructing composite space–time coordinates from spinorial twistor variables forms a basis of the Penrose twistor theory. [7] In this paper we follow the idea that the quark model is related to the parallel construction for *twistors with anticommuting coordinates*—provided that we replace fermionic nilpotent twistor components by one-dimensional (for fixed time) quark–twistor Grassmann field variables.

(iii) The only other candidates for the unification of symmetries—graded de Sitter groups—can be obtained as the subgroups of graded conformal groups (see, e.g., Ref. 9).

In the discussion on possible application of our fermionic superconformal geometry—in the second part of the paper—we shall follow the arguments which determined the Nambu-Goto action from the knowledge of the line element in pseudo-Euclidean space. [10] In such a way we introduce the two-dimensional nonlinear quark-twistor model, which can be described by a fermionic two-dimensional SU(2,2;1)-invariant field theory. Passing then for "pseudoclassical" anticommuting coordinates to quantized Grassmann variables—in accordance with the quantization rules of the Grassmann variant of QM (see, e.g., Refs. 11, 13, and 14)—we obtain fundamental quantized quark-twistor variables, which should play a similar role in hadrodynamics to that of the Heisenberg pairs of canonical variables in quantum mechanics of point particles.

It should be mentioned here that the dual models with ad hoc introduced fundamental two-dimensional spinorial quark fields have been discussed in the past; the most interesting example is the Bardakci–Halpern quark dual model. [15] Our aim here is to provide the geometric basis for such a type of quark dual theory. A more detailed analysis of the model proposed here, and in particular,

—introduction of internal (flavor and color) symmetries and corresponding two-dimensional gauge fields (gluons),

—application to the description of hadronic states, will be discussed in the near future.

II. TWISTORS AND SUPERTWISTORS

Flat twistor space $T_4(c)$ is a complex four-dimensional vector space with indefinite Hermitian metric

$$\eta = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix},$$

i.e., with the following U(2,2)-invariant norm:

$$T_4(c): |t|^2 = t^*_\alpha \eta^{\alpha\beta} t_\beta = \bar{t}t = |t_1|^2 + |t_2|^2 - |t_3|^2 - |t_4|^2 = \text{inv}, \tag{1}$$

where $\bar{t} = t^*\eta$. The isomorphism between SU(2,2) and conformal algebra is obtained by the following two inequivalent choices of 4×4 Hermitian matrices [16,17]:

$$p_\mu = \mp\tfrac{1}{2}\gamma_\mu(1 \pm i\gamma_5), \quad k_\mu = \pm\tfrac{1}{2}\gamma_\mu(1 \mp i\gamma_5),$$
$$m_{\mu\nu} = \sigma_{\mu\nu} = \tfrac{1}{4}[\gamma_\mu,\gamma_\nu], \quad d = \mp\tfrac{1}{2}\gamma_5, \tag{2}$$

where $\eta = \gamma_0$ and $\bar{t}$ describes the adjoint Dirac bispinor; the norm (1) corresponds to the choice $\gamma_0 = \sigma_3 \otimes 1$ of the representation of γ_μ matrices. The conformal generators on twis-

[a)]Permanent address: Institute of theoretical Physics, University of Wrocław, Poland.

 0022-2488/80/030561-07$1.00

tor space are as follows:

$$\hat{p}_\mu = t p_\mu^T \frac{\partial}{\partial t} + t^* p_\mu \frac{\partial}{\partial t^*}, \quad \text{etc.,} \tag{3}$$

where we also included a term describing the action on complex conjugate twistor space $T_4^*(c)$.

The twistor coordinates t_α can be treated as:

(a) Projective coordinates on three-dimensional complex space CP(3). In such a way one identifies twistors with the coordinates t_α and ut_α, where $|u| = 1$, i.e., one restricts U(2,2) to SU(2,2).

(b) One can add to the pseudo-unitary rotations the twistor translations, and consider inhomogeneous pseudo-unitary Lie group IU(2,2) as the group of motions in twistor space. In such a case the metric $\eta_{\alpha\beta}$ describes the scalar product in the tangent bundle over the twistor space.

If we wish to represent SU(2,2;1) as the group of linear transformations on supertwistor space, there are two possible ways of generalizing flat bosonic twistor space:

(i) One adds to four commuting twistor variables (t_1,t_2,t_3,t_4) a fifth complex Grassmann coordinate u ($u^2 = 0$). In such a way one obtains Ferber's *supertwistors*[18] (see also Ref. 19). The norm (1) is generalized as follows:

$$|t|^2 \to |t|^2 + |u|^2 = |t_1|^2 + |t_2|^2 - |t_3|^2 - |t_4|^2 + |u|^2 = R^2. \tag{4}$$

(ii) One can assume that the complex twistor coordinates *anticommute*, i.e., $\{t_\alpha,t_\beta\} = 0$, and the fifth complex coordinate u is a bosonic one. In such a case again the group of graded linear transformations, leaving the form (4) invariant, defines the superconformal group.[20]

We shall choose the second possibility because (a) such a choice is consistent with the spin statistics theorem: Twistors are fermions because they transform under SL(2,C) as Dirac bispinors, and fermionic degrees of freedom are described by Grassmann's variant of classical mechanics,[22] (b) only in case (ii) one can introduce in a convenetional way the projective Grassmann coordinates

$$z_\alpha = (t_\alpha/u)R, \tag{5}$$

which we shall use for generating the metric structure on complex projective Grassmann space $\widetilde{CP}(4)$.

III. LIE ROOTS AND JORDAN ROOTS OF LIE GROUP GENERATORS

Before considering the superconformal group, let us observe the general rule that:

—generalized translations in curved bosonic spaces are generated by *Lie roots* of O(p,q) (real pseudo-Euclidean metric) or U(p,q) (complex pseudo-Hermitian metric) generators

—generalized supertranslations in curved Grassmann spaces are generated by the *Jordan roots* ("super-roots") of Sp($2n$) (real symplectic metric) or U(p,q) (complex pseudo-Hermitian metric) generators.

Let us consider the following familiar examples.

A. Lorentz group O(3,1)

It is well known that the Lie roots $\mathscr{P}_\mu$ of Lorentz generators $M_{\mu\nu}$,

$$[\mathscr{P}_\mu,\mathscr{P}_\nu] = (1/R^2)M_{\mu\nu}, \tag{6}$$

generate translations in curved Minkowski space with constant curvature (see, e.g., Refs. 24 and 25). Curving Minkowski space means the following change of the group of motions:

$$\underset{\text{(Poincaré group)}}{\text{IO(3,1)}} \xrightarrow[R\neq 0]{} \underset{\text{(de Sitter groups)}}{\text{O(4,1) or O(3,2)}} \tag{7}$$

B. Spinorial de Sitter group Sp(4)

In order to obtain a nonvanishing norm of vectors, one should define the real symplectic group Sp(4) as leaving the symplectic norm invariant in "flat" four-dimensional real Grassmann space $G_4(R)$ (see, e.g., Refs., 26–30). One can further show[29,30] thus the super-roots Q_α of the Sp(4) generators $\mathscr{M}_{\alpha\beta}$,

$$\{Q_\alpha,Q_\beta\} = \mathscr{M}_{\alpha\beta}, \tag{8}$$

generate supertranslations in curved real Grassmann space with constant curvature, where curving $G_4(R)$ means the following modification of the group of isometries [31]:

$$\underset{\substack{\text{Inhomogeneous}\\ \text{Sp(4) with anticommuting}\\ \text{spinor}\\ \text{translations}}}{\text{ISp(4)}} \xrightarrow[R\neq 0]{} \underset{\substack{\text{Graded orthosymplectic}\\ \text{group}\\ \text{with Sp(4) bosonic}\\ \text{sector}}}{\text{OSp(4,1)}} \tag{9}$$

C. Conformal $+\gamma_5$ group U(2,2)

Due to the fact that one can consider the twistor norm (1) as describing the metric in bosonic as well as fermionic (Grassmann) twistor space, we shall have the possibility of defining Lie roots as well as super-roots of the generators of U(2,2). We obtain:

(a) Lie roots: If the group of isometries in bosonic twistor space is changed as follows:

$$\text{IU(2,2)} \xrightarrow[R\neq 0]{} \text{SU(2,3)}; \tag{10a}$$

(b) Super-roots: If the group of isometries in fermionic twistor space is modified in the following way:

$$\left(\begin{array}{c}\text{Inhomogeneous}\\ \text{U(2,2) with anticommuting}\\ \text{twistor}\\ \text{translations}\end{array}\right) \xrightarrow[R\neq 0]{} \left(\begin{array}{c}\text{Wess-Zumino}\\ \text{Superconformal}\\ \text{groups}\end{array}\right) \tag{10b}$$

We shall discuss relation (10b) in more detail below.

IV. NONLINEAR REALIZATIONS OF SUPERCONFORMAL GROUP ON CURVED FERMIONIC TWISTOR SPACE

Let us write the U(2,2) algebra in a natural unitary basis, consisting of 10 symmetric ($S_{\alpha\beta} = S_{\beta\alpha}$) and six antisymmetric ($A_{\alpha\beta} = -A_{\beta\alpha}$) anti-Hermitian generators:

$$\begin{aligned}
[A_{\alpha\beta},A_{\gamma\delta}] &= \eta_{\beta\gamma}A_{\alpha\delta} - \eta_{\alpha\gamma}A_{\beta\delta} - \eta_{\alpha\gamma}A_{\alpha\gamma} + \eta_{\alpha\delta}A_{\beta\gamma},\\
[A_{\alpha\beta},S_{\gamma\delta}] &= \eta_{\alpha\gamma}S_{\beta\delta} - \eta_{\beta\gamma}S_{\alpha\delta} - \eta_{\beta\delta}S_{\beta\gamma} + \eta_{\alpha\delta}S_{\beta\gamma},\\
[S_{\alpha\beta},S_{\gamma\delta}] &= \eta_{\beta\gamma}A_{\alpha\delta} + \eta_{\alpha\gamma}A_{\beta\delta} + \eta_{\beta\delta}A_{\beta\gamma} + \eta_{\alpha\delta}A_{\beta\gamma}.
\end{aligned} \tag{11}$$

The differential realizations of the generators $S_{\alpha\beta}, A_{\alpha\beta}$ on fermionic twistor space are as follows [33]:

$$S_{\alpha\beta} = i\left(t_\alpha \frac{\partial}{\partial t^\beta} + t_\beta \frac{\partial}{\partial t^\alpha}\right) - i\left(t^*_\alpha \frac{\partial}{\partial t^{\beta *}} + t^*_\beta \frac{\partial}{\partial t^{\alpha *}}\right),$$
$$A_{\alpha\beta} = \left(t_\alpha \frac{\partial}{\partial t^\beta} - t_\beta \frac{\partial}{\partial t^\alpha}\right) + \left(t^*_\alpha \frac{\partial}{\partial t^{\beta *}} - t^*_\beta \frac{\partial}{\partial t^{\alpha *}}\right). \tag{12}$$

It is easy to check that the infinitesimal generators preserve the norms (1) and (4). In order to introduce all 25 generators leaving the norm (4) invariant, we define:

(a) Generators of Grassmann rotations in four complex (t_α, u) planes, generated by [34]

$$S_\alpha = \frac{i}{R}\left(u \frac{\partial}{t^\alpha} + t_\alpha \frac{\partial}{\partial u}\right) - \frac{i}{R}\left(u^* \frac{\partial}{\partial t^{\alpha *}} - t^*_\alpha \frac{\partial}{\partial u^*}\right),$$
$$A_\alpha = \frac{1}{R}\left(u \frac{\partial}{\partial t^\alpha} - t_\alpha \frac{\partial}{\partial u}\right) + \frac{1}{R}\left(u^* \frac{\partial}{\partial t^{\alpha *}} + t^*_\alpha \frac{\partial}{\partial u^*}\right). \tag{13}$$

(b) Phase generator for u variable: $S = u\partial/\partial u$. Taking into consideration the relations

$$\left\{t_\alpha, \frac{\partial}{\partial t_\beta}\right\} = \delta_{\alpha\beta}, \quad \left[u, \frac{\partial}{\partial u}\right] = 1, \quad \text{etc.}, \tag{14}$$

one obtains the SU(2,2;1) algebra written in the natural twistor basis [35]:

$$\{A_\alpha, A_\beta\} = \frac{i}{R^2}(\widetilde{S}_{\alpha\beta} + 2A\,\eta_{\alpha\beta}),$$
$$\{S_\alpha, S_\beta\} = \frac{-1}{R^2}(\widetilde{S}_{\alpha\beta} + 2A\,\eta_{\alpha\beta}), \tag{15}$$
$$\{A_\alpha, S_\beta\} = \frac{i}{R^2} A_{\alpha\beta},$$

where

$$\widetilde{S}_{\alpha\beta} = S_{\alpha\beta} - \frac{i}{2}\eta_{\alpha\beta}\left(t^\gamma \frac{\partial}{\partial t_\gamma} - t^{*\gamma} \frac{\partial}{\partial t^*_\gamma}\right), \tag{16a}$$

$$A = i\left(\frac{1}{4} t^\gamma \frac{\partial}{\partial t_\gamma} + u \frac{\partial}{\partial u} - \frac{1}{4} t_{*\gamma} \frac{\partial}{\partial t^*_\gamma} - u^* \frac{\partial}{\partial u^*}\right). \tag{16b}$$

We see, therefore, that from the generators of *two* U(1) groups,

$$t'_\alpha = e^{i\alpha} t_\alpha, \quad u' = e^{i\beta} u, \tag{17}$$

commuting with the conformal SU(2,2) group [with the generators (16a)] and leaving the supertwistor norm (4) invariant, only *one* U(1) generator (16b) occurs in the superalgebra relations (15). It is easy to recognize [1] that the choice $\beta = 4\alpha$ describes the axial γ_5 transformations, which transform the variable z_α as follows:

$$z'_\alpha = e^{-3i\beta/4} z_\alpha. \tag{18}$$

Let us now parametrize the pseudosphere (4) in supertwistor space by homogeneous coordinates (5), transforming under SU(2,2) as twistor components, and a real angle variable φ, where $0 \leqslant \varphi \leqslant 2\pi$. The coordinates (t_α, u) on the manifold (4) are the following functions of the new variables (z_α, φ):

$$t_\alpha = \frac{z_\alpha e^{i\varphi}}{(1 + |z|^2/R^2)^{1/2}}, \quad u = \frac{e^{i\varphi} R}{(1 + |z|^2/R^2)^{1/2}}. \tag{19}$$

Formulas (17) imply the following metric form on the projective Grassmann space $\widetilde{CP}(4)$ with twistor coordinates z_α, z^*_α, denoted below the Z_A $(A = 1,\ldots,8)$:

$$ds_2 = \left(\frac{\partial t^*_\alpha}{\partial Z_A}\frac{\partial t^\alpha}{\partial Z_B} + \frac{\partial u^*}{\partial Z_A}\frac{\partial u}{\partial Z_B}\right) dZ_A dZ_B$$
$$= \frac{1}{2}\frac{dz^*_\alpha dz^\alpha}{(1 + |z|^2/R^2)} - \frac{1}{R^2}\frac{(z^*_\alpha dz_\alpha)(z_\beta dz^*_\beta)}{(1 + |z|^2/R^2)^2}, \tag{20}$$

which is a standard Fubini–Study metric on a projective complex plane (see, e.g., Ref. 36) generalized to the case of complex fermionic manifold with indefinite Hermitian metric. [37]

In order to write the differential realizations of the superconformal group on $\widetilde{CP}(4)$ one can employ two methods:

(a) Use the infinitesimal form of the superconformal transformations on supertwistors (t_α, u), determined by the 5×5 matrix representation of superconformal algebra (see, e.g., Refs. 21 and 40):

$$\delta\begin{pmatrix} t_\alpha \\ u \end{pmatrix} = i\begin{pmatrix} \delta M_{\alpha\beta} & \delta\epsilon_\alpha \\ -\delta\bar{\epsilon}_\alpha & \delta a \end{pmatrix}\begin{pmatrix} t_\alpha \\ u \end{pmatrix}, \tag{21}$$

where

$\delta M^{+}_{\alpha\beta} = \eta \delta M \eta$,

$\delta\omega_\alpha$ is the arbitrary complex Grassmann twistor,

$\delta M_{\alpha\alpha} = \delta a$ (vanishing supertrace condition).

Substituting in formula (5), one obtains the following formulas for the 24 generators of the superconformal group.

(i) $\widetilde{S}_{\alpha\beta}$ and $A_{\alpha\beta}$ are expressed by formulas (12) and (16a) with the replacement $t_\alpha \to z_\alpha$.

(ii) The supertranslation generators S_α and A_α take the form

$$A_\alpha = \frac{1}{i}\left(\frac{\partial}{\partial z^\alpha} - \frac{z_\alpha z_\beta}{R^2}\frac{\partial}{\partial z^\beta}\right) + \frac{1}{i}\left(\frac{\partial}{\partial z^{\alpha *}} + \frac{z^*_\alpha z^*_\beta}{R^2}\frac{\partial}{\partial z^{\beta *}}\right),$$
$$S_\alpha = \left(\frac{\partial}{\partial z^\alpha} + \frac{z_\alpha z_\beta}{R^2}\frac{\partial}{\partial z^\beta}\right) - \left(\frac{\partial}{\partial z^{\alpha *}} - \frac{z^*_\alpha z^*_\beta}{R^2}\frac{\partial}{\partial z^{\beta *}}\right). \tag{22}$$

(iii)

$$A = -i\frac{3}{4}\left(z^\alpha \frac{\partial}{\partial z_\alpha} - z^{\alpha *}\frac{\partial}{\partial z^*_\alpha}\right) \tag{23}$$

if we assume the standard form for γ_5 generators, with the entries on diagonal $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 1)$.

(b) A mapping $(t_\mu, u) \to z_\alpha$ is a projection of the circles $(0 \leqslant \varphi < 2\pi)$ into the points of $\widetilde{CP}(4)$ generalizing Hopf fibering $S^9 \to P_4(c)$ to the case of indefinite fermionic space. In order to obtain local *isometric* mapping, one should perform the change of variables $(t_\mu, u) \to (z_\alpha, \varphi, R)$, where $e^{i2\varphi} = u_5/u^*_5$ and z_α, R as defined by formulas (4) and (5). The supertwistor norm (5) induces the following formula on the product $\widetilde{CP}(4) \otimes U(1)$, where U(1) describes the struc-

tural group of a bundle attached to every point z_α of our curved twistor space

$$dS^2 = \frac{dz^*_\alpha dz^\alpha}{1+|z|^2/R^2} + R^2(d\varphi - \psi^*_\alpha(z)dz^\alpha)(d\varphi - \psi_\alpha(z)dz^{\alpha^*}), \quad (24)$$

where

$$\psi_\alpha(z) = \frac{z_\alpha}{R}\frac{1}{1+|z|^2/R^2}. \quad (25)$$

It is interesting to observe that the spinor field (25) recalls analogous formulas obtained in quanternionic space in the discussion on nontrivial topologically Yang–Mills instanton fields (see, e.g., Ref. 41); the metric (20) is obtained if we put $d\varphi = 0$.

Using the mapping $(t_\mu, u) \to (z_\alpha, \varphi, R)$, we also obtain the generalization of formulas (22) and (23),

$$\widetilde{A}_\alpha = A_\alpha + \frac{1}{2iR^2}(z_\alpha - z^*_\alpha)\frac{\partial}{\partial\varphi}, \quad (26a)$$

$$\widetilde{S}_\alpha = S_\alpha - \frac{1}{2R^2}(z_\alpha + z^*_\alpha)\frac{\partial}{\partial\varphi}, \quad (26b)$$

and

$$\widetilde{A} = A + \frac{\partial}{\partial\varphi}. \quad (26c)$$

Formulas (26) provide an extension of the realization of the superconformal group from the curved twistor space $\widetilde{CP}(4)$ to the supermanifold $\widetilde{CP}(4) \otimes U(1)$. It is easy to see that the realizations (26) are reducible.

It is important to mention that the parametrization (19) of the pseudosphere (4) can be replaced by any other U(2,2)-invariant parametrization. One can use, for example, the complex stereographic projections, [42] which give more complicated formulas for the supertranslation generators but imply simpler (conformally flat) formulas for the induced metric on curved fermionic twistor space $\widetilde{CP}(4)$.

V. FROM FERMIONIC COORDINATES TO FERMIONIC STRINGS: SUPERSYMETRIC SU(2,2;1)-INVARIANT σ-MODEL

One can generalize the Nambu–Goto derivation of reparametrization-invariant string action in pseudo-Euclidean space to the case of curved space with the metric $g_{AB}(x)$, with the line element given by

$$ds^2 = g^{AB}(x)dx_A dx_B. \quad (27)$$

The action for the point particle $(x_A \to x_A(\xi))$ can be written in two ways:

$$S = \int ds = \begin{cases} \int d\xi \left(g^{AB}(x)\frac{\partial x_A}{\partial\xi}\frac{\partial x_B}{\partial\xi}\right)^{1/2}, & (28a) \\ \int ds g^{AB}(x)\frac{\partial x_A}{\partial s}\frac{\partial x_B}{\partial s}, & (28b) \end{cases}$$

where formula (28a) is reparametrization-invariant and formula (28b) is obtained in the proper time parametrization. If we pass to the two-dimensional fields, describing strings

$$x_A \to x_A(\xi_1, \xi_2), \quad (29)$$

the most convenient way of writing down a reparametrization-invariant string Lagrangian is to couple it to the two-dimensional gravitational field. [44] We get $(i,j = 1,2)$

$$\mathcal{L} = g^{AB}(x)\frac{\partial x_A}{\partial\xi_i}\frac{\partial x_B}{\partial\xi_j}h_{ij}\sqrt{-h}, \quad (30)$$

where $h_{ij}(\xi_1,\xi_2)$ is an arbitrary two-dimensional metric tensor in the "internal" string space $\xi = (\xi_1, \xi_2)$. The metric tensor components h_{ij} play the role of Lagrange multipliers for the Virasoro conditions: Indeed, one obtains, from (30),

$$T_{ij}(\xi) = \frac{\delta \int d^2\xi \mathcal{L}}{\delta h_{ij}(\xi)} = 0 \quad (31)$$

and after substituting (31) into (30) one obtains

$$\mathcal{L} = (-\det G)^{1/2}, \quad (32)$$

where the 2×2 matrix

$$G_{ij} = g^{AB}(x)\frac{\partial x_A}{\partial\xi_i}\frac{\partial x_B}{\partial\xi_j} \quad (33)$$

describes the metric structure on a two-dimensional surface $X_A = X_A(\xi_1,\xi_2)$.

Let us now apply such a scheme to our curved fermionic geometry, with the metric structure (20). The Lagrangian (30) takes the following form:

$$\mathcal{L} = \frac{1}{(1+|q|^2/R^2)}\left[\frac{1}{2}\frac{\partial q^*_\alpha}{\partial\xi_i}\frac{\partial q^\alpha}{\partial\xi_j} - \frac{1}{1+|q|^2/R^2} \times \frac{1}{R^2}\left(q^*_\alpha\frac{\partial q^\alpha}{\partial\xi_i}\right)\left(q_\beta\frac{\partial q^{\beta^*}}{\partial\xi_j}\right)\right]\sqrt{-h}\,h_{ij}, \quad (34)$$

where we introduced quark–twistor variables by the substitution [compare with (29)],

$$z_\alpha \to q_\alpha(\xi_1,\xi_2). \quad (35)$$

In such a formulation, the choice of two-dimensional metric h_{ij} characterizes the choice of the parametrization. We shall consider, further, the case of flat two-dimensional metric; by allowing complex reparametrization one can use flat Minkowski or Euclidean metric. In such a way we obtain the Lagrangian for *two-dimensional fermionic nonlinear* SU(2,2;1)*-ivariant σ, model,* with two-dimensional fields q_α transforming as twistors under the conformal transfromations. In order to single out the states which do not depend on reparametrization, we should look for the solutions of the Euler–Lagrange equation satisfying also the condition [45]

$$T_{ij} = \frac{1}{1+|q|^2/R^2}\left(\frac{\partial q^*_\alpha}{\partial\xi_i}\frac{\partial q^\alpha}{\partial\xi_j} - \frac{1}{2}\delta_{ij}\frac{\partial q^*_\alpha}{\partial\xi_k}\frac{\partial q^\alpha}{\partial\xi_k}\right) - \frac{1}{(1+|q|^2/R^2)^2}\left[\left(q^*_\alpha\frac{\partial q^\alpha}{\partial\xi_i}\right)\left(q_\beta\frac{\partial q^{\beta^*}}{\partial\xi_j}\right) + (i\leftrightarrow j) - \delta_{ij}\left(q^*_\alpha\frac{\partial q^\alpha}{\partial\xi_k}\right)\left(q_\beta\frac{\partial^{\beta^*}}{\partial\xi_k}\right)\right] = 0. \quad (36)$$

It is easy to see that the trace condition $T_{ii} = 0$ is automatically satisfied.

The Lagrangian (31) with flat h_{ij} can be quantized in accordance with the well-established rules in the Grassmann variant of QM. [13,14] If we calculate the canonically conjugat-

ed momenta [47]

$$\pi_{\alpha;i} = \frac{\partial \mathcal{L}}{\partial q_{\alpha;i}} = \frac{1}{(1+|q|^2/R^2)} \times \left(\frac{1}{2}\eta_\alpha{}^\beta - \frac{q^*_\alpha q^\beta}{1+|q|^2/R^2}\right) q^*_{\beta,i}, \tag{37a}$$

where we denoted $q_{\nu;i} \equiv \partial q_\beta / \partial \xi_i$, and similarly

$$\pi^*_{\alpha;i} = \frac{\partial \mathcal{L}}{\partial q^*_{\alpha;i}} = \frac{1}{1+|q|^2/R^2} \times \left(\frac{1}{2}\eta_\alpha{}^\beta - \frac{q_\alpha q^{\beta *}}{1+|q|^2/R^2}\right) q_{\beta;i}, \tag{37b}$$

they satisfy the canonical *symmetric* Poisson bracket relations, which, after quantization, give the canonical anticommutators

$$\{\pi_{\alpha;0}(\xi,\tau), q_\beta(\xi',\tau)\} = \delta(\xi - \xi'),$$
$$\{\pi^*_{\alpha;0}(\xi,\tau), q^*_\beta(\xi',\tau)\} = \delta(\xi - \xi'). \tag{38}$$

Having the canonical momenta, one can write the field-theoretic 2-currents and charges, generating SU(2,2;1) superconformal transformations of the quark–twistor fields $q_\alpha(\xi,\tau)$. One obtains, for example, the following fifteen canonical conformal currents in analogy with formula (3):

$$\mathcal{P}_{\mu;i} = q_\alpha (p^T_\mu)^{\alpha\beta} \pi_{\beta;i} + q^*_\alpha (p_\mu)^{\alpha\beta} \pi^*_{\beta;i},$$
$$\mathcal{M}_{\mu\nu;i} = q_\alpha (\sigma^T_{\mu\nu})^{\alpha\beta} \pi_{\beta;i} + q^*_\alpha (\sigma^T_{\mu\nu})^{\alpha\beta} \pi^*_{\nu;i}, \tag{39}$$

etc. In the framework of the "naive" canonical formalism, it can be checked that [49]:

(a) All 24 superconformal currents are conserved.

(b) All 24 time-independent generators

$$P_\mu = \int d\xi \, \mathcal{P}_{\mu;0}(\xi,\tau), \quad \text{etc.} \tag{40}$$

satisfy the algebra of SU(2,2;1).

We shall stop our considerations here. In relation to the model (34), one can raise several interesting questions:

(a) The existence of an equivalent current theory of the Sugawara type, with two-dimensional current coordinates.

(b) The role of the U(1) variable φ in the two-dimensional model (34). In particular, the *fields*, obtained by the substitution $\varphi \to \varphi(\xi_1,\xi_2)$, can have a nontrivial topology, described by the homotopy classes $\pi_1(S_1) = Z$, and represented by the mapping of a "big" circle in the (ξ_1,ξ_2) plane into the asymptotic values of the variable $\varphi(\xi_1,\xi_2)$.

(c) It is worth recalling that, in the ordinary two-dimensional σ model, the condition (31) characterizes topologically stable multi-instanton solutions. Because the physical excitations of quark–twistor string should satisfy the reparametrization-invariant gauge condition (31), they should be described by analogous instanton-like solutions of our fermionic system.

(d) One of the unconventional features of the model (34) is the second-order equation for the fermionic field. However, because topologically nontrivial solutions satisfy the equations of lower order, we conjecture that the fields describing the excitations of the quark-twistor string satisfy a *first-order equation, related to the* U(2,2)*-invariant Thirring model*; we recall that only the Thirring models are reparametrization-invariant (invariant under two-dimensional conformal transformations [51]) in two dimensions and describe non-trivial interaction.

VI. FINAL REMARKS

In conventional (Veneziano, Neveu–Schwarz) two-dimensional string models the external index denotes space–time vector components; the external symmetry is described by the Poincaré group. Since supersymmetries appeared as a chance for unifying geometric and internal symmetries (in the superalgebraic scheme *both* are the bilinear forms of the same set of supercharges), there were also proposals of introducing the supersymmetric string variables transforming under supersymmetry transformation as an irreducible set of bosonic and fermionic string coordinates. [52,53] The most attractive proposal was the spinning string theory in ten dimensions; by making six of these dimensions compact, a four-dimensional theory with SO(6)≃SU(4) internal symmetry group was obtained.

Our approach here is based also on the introduction of supersymmetric string coordinates, with two modifications, however, in comparison with Refs. 52 and 53:

(a) We prefer to use as "external" supersymmetry groups for a string model the graded conformal groups SU(2,2;N) instead of the orthosymplectic series OSp(4;N) or graded Poincaré groups. It is natural to assume that massless quarks are conformal spinors; the main aim of this paper was to show (for $N = 1$) that one can interpret geometrically the extension of U(2,2) to SU(2,2;1) if one introduces curved translations in fermionic twistor space.

(b) Since, on the fundamental "quark" level of hadrodynamics, it seems that the symmetry between fermionic and bosonic degrees of freedom is not a sound idea, [54] we exploit here the nonlinear realization of the superconformal group and use only four complex **fermionic** degrees of freedom as independent variables.

Finally, let us mention that Volkov and Akulov, in this pioneering work on supersymmetry, [55] considered the neutrino field as described by the spinor coordinate with spinor translations described by the Poincaré supergroup. It appeared that the idea did not work, e.g., due to the experimentally verifiable discrepancy of predictions for low-energy neutrino theorems. In this paper we propose to use instead the conformal supergroup and then to apply the analogous idea to the fundamental quark–twistor degrees of freedom in the two-dimensional string model. At this stage it is only a hypothesis, but we hope one which deserves further investigation. [56,57]

ACKNOWLEDGMENTS

The author is indebted to Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics. He also wishes to thank Dr. R. Casalbuoni, Dr. C. Rebbi, and Dr. J. Strathdee for useful comments.

[1]P.G.O. Freund and I. Kaplanski, J. Math. Phys. **17**, 228 (1976); W. Nahm, V. Rittenberg, and M. Scheunert, Phys. Lett. B **61**, 383 (1976); V. Kac, Commun. Math. Phys. **53**, 31 (1977).
[2]P.K. Townsend and P. van Nieuwenhuizen, Phys. Lett. B **67**, 439 (1977).
[3]J. Wess and B. Zumino, Nucl. Phys. B **70**, 39 (1974).
[4]R. Haag, J.T. Łopuszanski, and M. Sohnius, Nucl. Phys. B **88**, 257 (1975).
[5]S. Ferrara, M. Kaku, P. van Nieuwenhuizen, and P.K. Townsend, Nucl. Phys. B **129**, 125 (1977).
[6]R. Penrose, J. Math. Phys. **8**, 345 (1967); R. Penrose and M.A.H. MacCallum, Phys. Rep. C **6**, 242 (1972).
[7]See also Rzewuski *et al.*,[8] who introduced spinor space as the fundamental geometric manifold before Penrose. We quote here the first and the last paper of the Rzewuski group.
[8]J. Rzewuski, Nuovo Cimento **5**, 942 (1958); R. Abłamowicz, J. Mozrzymas, Z. Osiewicz, and J. Rzewuski, Rep. Math. Phys. **14**, 89 (1978).
[9]S. Ferrara, Phys. Lett. B **69**, 481 (1977).
[10]Similarly, one can derive the reparametrization-invariant action for the Neveu–Schwarz string from the line element in the Berexin–Marinov superspace for the spinning particle [11] described by the Minkowski coordinate and five (4-vector + pseudoscalar) Grassmann coordinates (see, e.g., Ref. 12).
[11]F.A. Berezin and M.S. Marinov, Ann. Phys. (N.Y.) **104**, 336 (1977).
[12]M.S. Marinov, Usp. Fiz. Nauk **121**, 377 (1977) (in Russian).
[13]J.L. Martin, Proc. Roy. Soc. (Lond.) A **251**, 536 (1959).
[14]R. Casalbuoni, Nuovo Cimento A **33**, 115 (1976).
[15]K. Bardakci and M.B. Halpern, Phys. Rev. D **3**, 2493 (1971); M.B. Halpern, Phys. Rev. D **4**, 3082 (1971).
[16]H.A. Kastrup, Ann. Physik (Leipz.) **9**, 388 (1962).
[17]G. Mack and A. Salam, Ann. Phys. (N.Y.) **53**, 174 (1968).
[18]A. Ferber, Nucl. Phys. B **132**, 55 (1978).
[19]R. Witten, Harvard Univ. Preprint, 1978.
[20]This second possibility is mentioned by Ferber [18] but not exploited. The twistors with anticommuting coordinates are used by Gürsey and Marchildon [21] in their discussion on nonlinear realizations of superconformal group on coset spaces. Our discussion on the fermionic superconformal geometry is complementary to that presented in Ref. 21.
[21]F. Gürsey and L. Marchildon, J. Math. Phys. **19**, 942 (1978); Phys. Rev. D **17**, 2038 (1978).
[22]We do not attach a direct physical meaning to Grassmann variables in "prequantized" fermion theory; simply, the description by means of anticommuting variables allows us to introduce without magic the quantization of the fermionic system—however, magic is still present in the interpretation of Grassmann coordinates (see, however, Ref. 23).
[23]A. Barducci, R. Casalbuoni, and L. Lusanna, Lett. Nuovo Cimento **19**, 581 (1977).
[24]H. Snyder, Phys. Rev. **71**, 38 (1947).
[25]F. Gürsey and T.D. Lee, Proc. Natl. Acad. Sci. USA **49**, 179 (1963).
[26]R. Arnowitt and P. Nath, Phys. Lett. B **56**, 177 (1975).
[27]P.P. Srivastava, Lett. Nuovo Cimento **13**, 657 (1975).
[28]R. Casalbuoni, Nuovo Cimento A **33**, 389 (1976).
[29]C. Fronsdal, Lett. Math. Phys. **1**, 165 (1976).
[30]J. Lukierski, Stony Brook Preprint ITP-SB-78-8, Lett. Nuovo Cimento **24**, 309 (1979).
[31]For the generalization of the notions of Riemannian geometry to Grassmann spaces see Refs. 26 and 29. We should like to mention here that one obtains a non-Riemannian geometry, with nonvanishing torsion, if one introduces superspace with independent bosonic and fermionic coordinates (see, e.g., Ref. 32).
[32]J. Wess and B. Zumino, Phys. Lett. B **56**, 361 (1977).
[33]It is interesting to observe that formulas (12) are also valid if t_α are bosonic Penrose twistors. We also mention that our derivatives over Grassmann variables are always *left* derivatives, i.e.,

$$\frac{\partial}{\partial t_\alpha}(t_\beta t_\gamma)=\left(\frac{\partial}{\partial t_\alpha}t_\beta\right)t_\gamma-t_\beta\left(\frac{\partial}{\partial t_\alpha}t_\gamma\right).$$

[34]The changes of sign in (13) in comparison with formulas (12) follow from the property that our derivatives are left derivatives, i.e.,

$$\frac{\partial}{\partial z^*_\alpha}z^{*\alpha}z_\alpha=z^\alpha \quad \text{but} \quad \frac{\partial}{\partial z_\alpha}z^{*\alpha}z_\alpha=-z^{*\alpha}.$$

[35]The supercharges S_α and A_α are introduced in a way which is consistent with twistor metric $\eta_{\alpha\beta}$; they are realized in the fundamental five-dimensional representation by 5×5 matrices as follows:

$$S_\alpha=\begin{pmatrix} & & & & \delta_{\alpha1}\\ & 0 & & & \delta_{\alpha2}\\ & & & & -\delta_{\alpha3}\\ & & & & -\delta_{\alpha4}\\ \delta_{\alpha1} & \delta_{\alpha2} & \delta_{\alpha3} & \delta_{\alpha4} & 0\end{pmatrix},$$

$$A_\alpha=i\begin{pmatrix} & & & & -\delta_{\alpha1}\\ & 0 & & & -\delta_{\alpha2}\\ & & & & \delta_{\alpha3}\\ & & & & \delta_{\alpha4}\\ \delta_{\alpha1} & \delta_{\alpha2} & \delta_{\alpha3} & \delta_{\alpha4} & 0\end{pmatrix},$$

etc., comparing with the entries $(C^+_{10},\ldots,C^+_{4\alpha},0)$ in the fifth column and $(\delta_{\dot\alpha1},\ldots,\delta_{\dot\alpha4})$ in the fifth row [where $C^{\pm}=\frac{1}{2}(1\pm\gamma_5)C$ and $\delta^{\pm}=\frac{1}{2}(1\pm\gamma_5)$; see Ref. 3], which correspond to the conventional choice of Marjorana supercharges and to the choice of antisymmetric Hermitian metric in twistor space. One can give, for example, simple relations expressing S_α and A_α in terms of supercharges written as Weyl spinors. We have, e.g.,

$$S_1=(1/\sqrt{2})(-iQ_1{}^+-S_2{}^+),$$

$$S_2=(1/\sqrt{2})(-iQ_2{}^++S_1{}^+), \quad \text{etc.},$$

where Q_α is a complex Weyl super-root of P_μ, and S_α is a complex Weyl super-root of K_μ (see, e.g., Ref. 4).
[36]R.O. Wells, Jr., *Differential Analysis on Complex Manifolds* (Prentice-Hall, Englewood Cliffs, N.J., 1973).
[37]The metric (18) describes pseudo-Kählerian structure on complex Grassmann space. The pseudo-Kähler indefinite metrics have been considered for complex bosonic manifolds (see, e.g., Refs. 38 and 39). The generalization of Kähler structures to the complex spaces with fermionic coordinates can be obtained provided that we take care of the order of Grassmann derivatives in the corresponding formulas (to be published).
[38]E.J. Flaherty, *Hermitian and Kählerian Geometry in Relativity* (Springer, New York, 1976).
[39]M. Ko, E.T. Newman, and R. Penrose, J. Math. Phys. **18**, 58 (1977).
[40]V. Rittenberg, in *Processings of the* 1977 *Tubingen Conference on Group-Theoretic Methods in Physics* (Springer, New York, 1978).
[41]A. Trautman, Int. Theoret. Phys. **16**, 561 (1977).
[42]Such a coordinate system was used in Ref. 43 for the case of bosonic Hermitian complex sphere, attached to every point of the Minkowski space.
[43]R. Casalbuoni, G. Domokos, and S. Kovesi-Domokos, Phys. Rev. D **17**, 2048 (1978).
[44]S. Deser and B. Zumino, Phys. Lett. B **65**, 369 (1976).
[45]The condition $T_{ij}=0$ follows if we observe that the integral $\delta G(\tau)=\int d\xi\, T_{0i}(\xi,\tau)\delta\xi_i(\xi,\tau)$ generates an arbitrary change of the coordinates $\xi_i\to\xi_i+\delta\xi_i(\xi,\tau)$ (where we put $\xi_1=\xi$ and $\xi_2=\tau$). It should be mentioned here that in the case of anticommuting variables, when the action integral is not a genuine c-number but belongs to the even sector of Grassmann algebra, the meaning of the variational principle is not clear; one also encounters problems in the derivation of the Euler–Lagrange equation and Noether theorem. We shall use here the variational principle, in accordance with Schwinger's prescription for fields with anticommuting variations (see Ref. 46) without entering into the mathematical justification.
[46]J. Schwinger, Phil. Mag., Ser. 7 **44**, 1171 (1953); G.R. Allcock, in *Cooperative Phenomena*, edited by H. Haken and M. Wagner (Springer-Verlag, Berlin, 1973).
[47]There are well-known ordering problems and nonequivalence of the Lagrangian and Hamiltonian formalisms for the nonlinear bosonic Lagrangians (see, e.g., Ref. 48). We expect similar difficulties in fermionic nonlinear models. We write here our "naive" formulae modulo ordering problems.
[48]J.M. Charap, J. Phys. A **6**, 393 (1973); M. Omote and H. Sato, Progr. Theoret. Phys. (Kyoto) **47**, 1367 (1972).
[49]Similarly to the situation with nonlinear O(N) σ-model (see, e.g., Ref. 50) we expect, in the superconformal currents, anomalies which are due to the renormalization procedure.
[50]A.M. Polyakov, Phys. Lett. B **72**, 224 (1977).
[51]S. Ferrara, R. Gatto, and A. Grillo, Nuovo Cimento A **12**, 959 (1972).

[52]F. Gliozzi, J. Scherk and D. Olive, Phys. Lett. B **65**, 282 (1976); Nucl. Phys. B **122**, 253 (1977).
[53]J.H. Schwarz, Lecture presented at Orbis Scientiae, January 1978; Preprint Calt-68-637 (1978).
[54]Gluons should appear as gauge fields, which are not required by global supersymmetry transformations.
[55]D.V. Volkov and V.P. Akulov, Phys. Lett. B **46**, 109 (1973).
[56]Some nonoverlapping material was presented in a lecture by the author at Symposium on Mathematical Physics, Liblice, 18–23 June 1978; published in Czech. J. Phys. B **29**, 44 (1979).
[57]*Note added in proof*: See also lectures given by the author at IVth Workshop on Theoretical Physics in Erice, October 1978; to be published in Proceedings by Plenum Press (1979).

Magnetic resonances between massive and massless spin-$\frac{1}{2}$ particles with magnetic moments

A. O. Barut[a)]
International Centre for Theoretical Physics, Trieste, Italy

(Received 2 October 1979; accepted for publication 9 November 1979)

The general effective radial potentials for a spin-$\frac{1}{2}$ particle interacting with scalar, electric, and magnetic potentials are given. In the $m = 0$ limit, it is shown that the magnetic potential provides a well deep enough to confine the massless particle. In particular, there are exact zero-energy solutions in which two of the four components of the massless particle are confined; only two can leak out into the asymptotic region. The scattering amplitude is analytic in the entire j plane, hence consists only of Regge poles.

1. INTRODUCTION

This is a continuation of a previous work[1] where we have solved the Dirac equation with an anomalous magnetic moment in the field of a fixed charge and showed that the magnetic terms, when treated nonperturbatively, can give rise to narrow very high mass resonances. In this paper we show that the same conclusion also holds for a massless relativistic particle, e.g., neutrino.

That a massless particle can form a bound state, or a narrow resonance state, is not known so far in any model. It is thus remarkable that the anomalous magnetic moment of the neutrino can produce such localized states of a particle moving with velocity c. These resonances have been used recently in the construction of hadron multiplets, where the basic truly elementary constituents of matter are the stable particles: proton, electron, and neutrino.[2] It is therefore of interest to study the mathematical behavior of narrow resonances of a massless particle in soluble models.

We first generalize the radial coupled Dirac equations given in Ref. 1 with the inclusion of three types of potentials, an electric V_e, a magnetic V_m, and a scalar potential V_s:

$$\frac{df}{dr} = \frac{\kappa - 1}{r} f + (m + V_s - E)g + V_e g + V_m f,$$
$$\frac{dg}{dr} = -\frac{\kappa + 1}{r} g + (m + V_s + E)f - V_e f - V_m g. \tag{1}$$

The substitutions

$$u_1 = rg, \quad u_2 = rf, \quad u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{2}$$

give the linear matrix eigenvalue problem

$$\frac{du}{dr} = Mu, \quad M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{3}$$

where

$$A = -(\kappa/r + V_m), \quad B = (m + V_s + E - V_e),$$
$$C = (m + V_s - E + V_e), \quad D = -A.$$

We now give a matrix method to transform the system (3) into a diagonal Sturm–Liouville eigenvalue system.

From (3) we have

$$u'' = (M^2 + M')u, \tag{4}$$

where, from now on, primes will indicate derivatives with respect to r. In order to diagonalize the matrix $(M^2 + M')$, let

$$u = X\Psi, \quad \det X \neq 0. \tag{5}$$

Hence

$$u' = X\Psi' + X'\Psi, \quad u'' = X\Psi'' + 2X'\Psi' + X''\Psi.$$

Solving these equations for Ψ'', using (4) and (3), and eliminating Ψ', we have

$$\Psi'' = L\Psi, \tag{6}$$

where

$$L = X^{-1}\{[(M^2 + M') - 2X'X^{-1}M]X + 2(X^{-1}X')^2 - X''\}.$$

We now determine X in such a way that L is diagonal. By direct calculation we have

Theorem: The system (3) with $B \neq 0$ and $C \neq 0$ is equivalent to the Sturm–Liouville system (5)–(6) with a diagonal L when

$$X = \begin{pmatrix} \sqrt{B} & 0 \\ 0 & \sqrt{C} \end{pmatrix}.$$

Then L is given by

$$L = \begin{pmatrix} A^2 + BC + A' - A\frac{B'}{B} - \frac{1}{2}\frac{B''}{B} + \frac{3}{4}\frac{B'^2}{B^2} & 0 \\ 0 & A^2 + BC - A' - A\frac{C'}{C} - \frac{1}{2}\frac{C''}{C} + \frac{3}{4}\frac{C'^2}{C^2} \end{pmatrix}.$$

For the physical problem (3), if we write (6) as

$$\Psi'' + (E^2 - m^2 - V_{\text{eff}})\Psi = 0, \tag{7}$$

[a)]Permanent address: Department of Physics, University of Colorado, Boulder, Colorado 80309.

 0022-2488/80/030568-03$1.00

we obtain the effective potentials for Ψ_1:

$$V^1_{\rm eff} = \frac{\kappa(\kappa+1)}{r^2} + 2E\ V_e - V_e^2 + V_s^2 + 2m\ V_s + V_m^2 + 2\ \frac{\kappa V_m}{r} - V_m' + \frac{1}{2}\frac{V_e'' - V_s'' + 2(V_s' - V_e')(\kappa/r + V_m)}{(m + E + V_s - V_e)} + \frac{3}{4}\frac{(V_s' - V_e')^2}{(m + E + V_s - V_e)^2}, \tag{8}$$

and for Ψ_2 we have the *symmetry*

$$V^{(2)}_{\rm eff}(\kappa, E, V_e, V_m) = V^{(1)}_{\rm eff}(-\kappa, -E, -V_e, -V_m). \tag{9}$$

Equations (7)–(9) are useful not only for our problem, but also for the relativistic treatment of, for example, the "charmonium" potential, which may be put either in V_s or in V_e.

Conversely, a system of the form (6) can be transformed into (1), which in special cases is easier to solve; for example, the $E = 0$ solution discussed below.

Note that the potential (8) is both energy (E) and angular momentum (κ) dependent. It must be evaluated at each E and κ, and then one must look for an eigenvalue for a bound state or resonance at this energy and κ.

The potential in general has three well separated wells. The maximum number of real zeros of $V_{\rm eff}$ is five. Beginning with large r values, the first well corresponds to the Coulomb level at distances $r \sim 1/\alpha m$, the second well ($r \sim \alpha/m$) corresponds to ordinary hadronic energies, and the third well at $r \sim \alpha^2/m$ corresponds to GeV energy range. The physical interpretation and application to hadron multiplets and high-energy particle physics processes has been given elsewhere.[2]

Equations (7) and (8) could form the basis of many investigations depending on the choice of the three potentials V_s, V_e, and V_m.

In our specific physical problem we have $V_s = 0$ and ($c = \hbar = 1$)

$$V_e = \frac{e_1 e_2}{r} F_e(r), \quad V_m = a\frac{e_1 e_2}{2m}\frac{1}{r^2} F_m(r), \tag{10}$$

where we have also introduced electric and magnetic form factors F_e and F_m, respectively. Here e_1 and e_2 are the charges of the particle and the force center and a, the anomalous magnetic moment of the relativistic particle.

2. NEUTRINO LIMIT

We now let $e_1 \to 0$ and $m \to 0$ such that $a_2(e_1/m) = \mu =$ anomalous magnetic moment of the neutrino. Hence $V_e = 0$.

Equations (7), (8), and (9) become

$$\Psi'' + (E^2 - V_{\rm eff})\Psi = 0 \tag{11}$$

$$V^{(1)}_{\rm eff} = \frac{\kappa(\kappa+1)}{r^2} + V_s^2 + V_m^2 + 2\ \frac{\kappa V_m}{r} - V_m' + \frac{1}{2}\frac{-V_s'' + 2(V_s')((\kappa/r) + V_m)}{E + V_s} + \frac{3}{4}\frac{V_s'^2}{(E + V_s)^2}, \tag{12}$$

$$V^{(2)}_{\rm eff}(\kappa, E, V_m) = V^{(1)}_{\rm eff}(-\kappa, -E, -V_m),$$

or, in the absence of a scalar potential V_s and with $V_m = (e\mu/r^2)F_m(r)$,

$$V^{(1)}_{\rm eff} = \frac{\kappa(\kappa+1)}{r^2} + 2e\mu\ \frac{\kappa+1}{r^3} F_m + \frac{e^2\mu^2}{r^4} F_m^2. \tag{13}$$

In dimensionless coordinates $r = |e\mu| y = r_0 y$, $\epsilon = \sin(e\mu)$

$$\left(\frac{d^2}{dy^2} - V_{\rm eff}(y) + \lambda^2\right)\Psi = 0 \quad \lambda^2 = |e\mu|^2 E^2, \tag{14}$$

$$V^{(1)}_{\rm eff}(y) = \frac{\kappa(\kappa+1)}{y^2} + \epsilon\ \frac{2(\kappa+1)}{y^3} F_m + \frac{1}{y^4} F_m^2,$$

$$V^{(2)}_{\rm eff}(y) = \frac{\kappa(\kappa-1)}{y^2} + \epsilon\ \frac{2(\kappa-1)}{y^3} F_m + \frac{1}{y^4} F_m^2. \tag{15}$$

It is interesting to note that the same equation (14), with the same potential (15), holds also for (a) a Schrödinger or Klein–Gordon particle in the field of a fixed quantum magnetic moment, and (b) a relativistic *massive* particle with magnetic moment in a fixed Coulomb field. Only the relation of the eigenvalue λ^2 to energy is different in these different cases.

3. ZERO-ENERGY SOLUTIONS

For $\epsilon = -1$, in particular $V_{\rm eff}$ exhibits a deep well, and bound state and resonances are possible. In general, for $F_m = 1$, but any ϵ and κ, Eq. (15) is exactly soluble analytically.[3] For a zero-energy bound state (or resonance) the total mass of the system is equal to the mass of the central charged particle. The $E = 0$ solution can in fact be obtained immediately from Eq. (1) for any V_m:

$$f = Cr^{\kappa-1} \exp\left(\int V_m(r)\, dr\right),$$

$$g = C' r^{-(\kappa+1)} \exp\left(-\int V_m\, dr\right). \tag{16}$$

In particular, for V_m given by (10) we obtain

$$f = Cr^{\kappa-1} \exp(-\epsilon r_0/r), \quad g = C' r^{-(\kappa+1)} \exp(+\epsilon r_0/r).$$

It is remarkable that for $\kappa = 1$, $\epsilon = -1$, for example, the f components of the wavefunction leaks out and approaches a constant as $r \to \infty$, whereas the g component is localized around $r = r_0/2$. For $\kappa = -1$, $\epsilon = +1$, the opposite holds.

Only at the energy $E = 0$ the two spinors f and g become uncoupled. The confined two-component solution is normalizable, whereas the leaking two-component solution is, as in the case of a free particle, not normalizable. The latter must be normalized in a box or with an energy δ function $\delta(E' - E)$.

More precisely, the acceptable localized, normalizable solutions are, for $\epsilon = -1$, of the form

$$\begin{pmatrix} ig_{\kappa=+1} \\ 0 \end{pmatrix},$$

and for $\epsilon = +1$

$$\begin{pmatrix} 0 \\ -f_{(\kappa=-1)} \end{pmatrix}.$$

Solutions normalizable in a box are of the form

$$\begin{pmatrix} ig_{(\kappa=-1)} \\ 0 \end{pmatrix} \quad \text{for } \epsilon = -1$$

and

$$\begin{pmatrix} 0 \\ -f_{(\kappa=+1)} \end{pmatrix} \quad \text{for } \epsilon = +1.$$

The scalar product is the usual one: $c\int(f^2+g^2)r^2\,dr$.

The physical interest for the above solution stems from the fact that neutrinos in asymptotic states seem to occur in nature as two-component spinors and that the other two components are missing, which is at the heart of parity violation.

Remarks

(1) Even if the magnetic moment μ of the massless particle is zero, but the form factor $F_m(r)$ is different from zero [note $\mu = F_m(\infty)$], the qualitative feature of the potential (15) does not change. This is because the centrifugal barrier $[\kappa(\kappa+1)]/y^2$ is independent of the magnetic moment, and the other two terms have dominant roles at short distances.

(2) For potentials with a singular repulsive core the scattering amplitude is an analytic function in the entire l plane.[3] The magnetic interactions expressed in potentials (8)–(10) show that such potentials are physical and must occur in nature. The amplitude is then a pure sum of Regge poles, no cuts and background integral occurring. This provides a new theoretical basis for the phenomenological analysis of the scattering amplitude in the present model.[2]

(3) The effect of the dipole-dipole interactions will be treated separately.

(4) Numerical studies by Skorupski and Senatorski[4] show that Eq. (14) and $\kappa = 1$, for $\lambda^2 \neq 0$, has no bound state solutions except the partly bound $\lambda^2 = 0$ solution, agreeing with the analytic result (16).

(5) Note that the standard results about the type of localizability or nonlocalizability refer to *free* massless particles with magnetic moment bound to a massive charge.[5]

4. "ZERO ENERGY" SOLUTIONS FOR $m \neq 0$

Even for a massive particle, coupled via a magnetic potential to a fixed charge, the "zero-energy" solutions (i.e., $E = m$) can be integrated exactly. From Eq. (1), with $V_e = 0$, $V_s = 0$, and $E = m$, we obtain for f the same solution as in (16), i.e.,

$$f = C_1 r^{\kappa-1} \exp\left(\int V_m(r)\,dr\right). \tag{17}$$

But now g satisfies the inhomogeneous equation

$$g' = \left(-\frac{\kappa+1}{r} - V_m\right)g + 2mf. \tag{18}$$

with (17), Eq. (18) can be integrated. If we put $g = hd$, then $g' = h'd + hd'$. Letting $h'd = 2mf$, we have $d' = (-(\kappa+1)/(r) - V_m)d$. Hence $d = C_2\exp[\int(-(\kappa+1)/(r) - V_m)dr]$. Consequently, $h' = (2mf/d) = 2mr^{2\kappa}\exp[2\int V_m\,dr]$. Thus the most general solution of (18) is

$$g = r^{-(\kappa+1)}\exp\left[-\int V_m\,dr\right]\left\{C + 2m\int dr r^{2\kappa} \times \exp\left[2\int V_m(r')\,dr'\right]\right\}, \tag{19}$$

which for $m = 0$, reduces back to Eq. (16).

For the magnetic potential given in (10), i.e., $V_m = r_0/r^2$, we then obtain explicitly

$$f = Cr^{\kappa-1}\exp(-r_0/r),$$

$$g = r^{-(\kappa+1)}\exp(r_0/r)\left[C + 2m\,e^{-2r_0/r} \times \sum_{\nu=0}^{-2\kappa-2}(-1)^\nu \frac{\Gamma(-2\kappa-1)}{\Gamma(-2\kappa-1-\nu)} r^{-2\kappa-2-\nu}\right], \quad \kappa = -1, -2, -3, \cdots$$

$$= r^{-(\kappa+1)}\exp(r_0/r)\left[C - 2m\,e^{-2r_0/r} \times \sum_{\nu=1}^{2\kappa+1}(2r_0)^{\nu-1}\frac{\Gamma(2\kappa+2-\nu)}{\Gamma(2\kappa+2)} r^{2\kappa+2-\nu} + \frac{(2r_0)^{2\kappa+1}}{(2\kappa+1)!}\epsilon i(2r_0/r)\right], \quad \kappa = 0,1,2,\cdots$$

ACKNOWLEDGMENTS

The author would like to thank Professor Abdus Salam, the International Atomic Energy Agency, and UNESCO for hospitality at the International Centre for Theoretical Physics, Trieste. He would also like to thank Bengt Nagel for many stimulating discussions.

[1] A.O. Barut and J. Kraus, J. Math. Phys. **17**, 506 (1976).
[2] A.O. Barut, in *Proceedings of the Texas Conference on Group Theory and Mathematical Physics*, September 1978, A. Böhm and E. Takasugi (Springer Lecture Notes in Physics, 1979); and in *Surveys in High Energy Physics*, Vol. 1, No. 2, January 1980.
[3] E. Predazzi and T. Regge, Nuovo Cimento **24**, 518 (1962).
[4] A. Skorupski and A. Senatorski (to be published).
[5] A.S. Wightman, Rev. Mod. Phys. **34**, 845 (1962).

Solution of boundary value problems with Laplace's equation for ellipsoids and elliptical cylinders

Leonard Eyges
Rome Air Development Center, Deputy for Electronic Technology, Hanscom AFB, Massachusetts 01731

(Received 24 April 1979; accepted for publication 15 October 1979)

If a dielectric ellipsoid or elliptical cylinder is placed in a uniform applied field it is well known that the field inside remains uniform, and is changed only by a depolarization factor that multiplies each applied field component. This paper generalizes this result. Namely, if for the three-dimensional case the potential Φ_{app} of the applied field can be expanded in the neighborhood of the ellipsoid as $\Phi_{app} = \Sigma_{l=0}^{L} \Sigma_{m=-l}^{l} D_{lm} r^l Y_{lm}(\theta,\varphi)$ where l goes from zero to a maximum value L, then it is shown that the resultant potential inside the ellipsoid, Φ_{int}, is $\Phi_{int} = \Sigma_{l=0}^{L} \Sigma_{m=-l}^{l} C_{lm} r^l Y_{lm}(\theta,\varphi)$ where the coefficients C_{lm} are found explicitly and there is no C_{lm} with $l > L$. For a dielectric constant ϵ, the limits of the above solution as $\epsilon \to \infty$ and $\epsilon \to 0$ are considered and are shown to yield respectively the solutions to the Dirichlet problem with potential zero on the boundary (grounded perfect conductor) and the Neumann problem with normal derivative of the potential zero on the boundary (ideal fluid flow). The homogeneous problem of free charge on an ellipsoidal perfect conductor is considered and it is shown to require a modification of the methods that yielded the results above. The modified method is applied to the problem of the oblate ellipsoid, and its limiting case of a disc, and it enables the easy derivation of various classical results due to Copson and others on the "problem of the electrified disc." Finally, multiple dielectric ellipsoids or elliptic cylinders are considered and it is shown that problems involving such bodies can be solved in powers of a_i/d_{ij} where a_i is a typical length of the ith body and d_{ij} is the distance between the ith and the jth. This opens the way, in the proper limit of ϵ, to the solution of a variety of problems, such as flow around multiple strips and through the slots they may form, penetration of the electric field through perforated screens, and so on.

1. INTRODUCTION

In a previous paper[1] we discussed the problem of finding the electrostatic (magnetostatic) potential when a dielectric (permeable) body is placed in an applied electrostatic (magnetostatic) potential, denoted by Φ_{app}. An integral equation for the potential was formulated and was used to derive the well-known result that the field inside a homogeneous dielectric spheroid in a uniform external field is again uniform, but reduced by the so called *depolarization factor*. A similar result holds for the analogous problem involving a two-dimensional elliptic cylinder. The basic result of this paper can be considered to be the extension to a more general *applied potential* of these results for a uniform applied potential. The derivation is very similar for the two and three dimensional cases so we sometimes present the discussion in terms of three dimensional examples and sometimes in terms of two dimensional ones, as convenience dictates. Thus, when we write about elliptic cylinders there is almost always an analogous discussion possible for ellipsoids, and conversely.

In two dimensional cylindrical coordinates ρ, φ, an arbitrary applied potential, whose sources are distant from the origin, can be written in the neighborhood of the origin as

$$\Phi_{app} = \sum_{n=0}^{\infty} (D_n \Phi_{app}^{(n)} + \tilde{D}_n \tilde{\Phi}_{app}^{(n)}), \tag{1}$$

where

$$\Phi_{app}^{(n)} = \rho^n \cos n\varphi, \quad \tilde{\Phi}_{app}^{(n)} = \rho^n \sin n\varphi. \tag{2}$$

Let an elliptic cylinder be in the potential $\Phi^{(n)}$. Then it is shown, and this is the fundamental result, that the *resultant potential in the interior*, Φ_{int}, has an expansion of the form

$$\Phi_{int} = \sum_{l=0}^{n} C_l \rho^l \cos l\varphi, \tag{3}$$

in which the C_l can be found explicitly, and, as indicated, there is no C_l with $l > n$. A similar result holds, of course, for $\tilde{\Phi}_{app}^{(n)}$. It follows by superposition that the solution can be constructed in finite terms for any applied potential that either exactly or approximately consists of a finite series of terms in an expansion like Eq. (1), and this comprises a large variety of potentials.

A second point of this paper is to show that in various limits the solution to the problem of the elliptic dielectric cylinder can be reinterpreted as the solution to other problems of physical interest. The first such limit is one that has been remarked on before: By letting $\epsilon \to \infty$ the dielectric problem becomes that of a grounded perfect conductor, i.e., a problem with the homogeneous Dirichlet condition $\Phi = 0$ on the boundary (Dirichlet-zero problem). Second is the limit $\epsilon \to 0$ which generates the solution to the homogeneous Neumann problem, with $\partial\Phi/\partial n = 0$ on the boundary (Neumann-zero problem); this is of course the problem of irrotational ideal fluid flow. Finally, and independently of these two limits, there is the well-known limit of vanishing thick-

 0022-2488/80/030571-11$1.00

ness in which the elliptic cylinder becomes a strip and the three dimensional oblate spheroid becomes a disc.

There is a further consequence of the basic result of this paper: It enables the solution of problems with multiple elliptic cylinders. To make this remark plausible consider two such cylinders each with major axis of length a and suppose that the centers of the cylinders are separated by a distance d. Then the first cylinder is acted on by the applied field $\Phi_{\rm app}$ *and* by the field of the second cylinder. This latter field is an *effective applied field* for the first cylinder and if a/d is less than unity, it can be shown that the expansion of this second applied field around the first cylinder is of the form of Eq. (1). It will follow from this fact, as we show in Sec. 5, that a solution in powers of (a/d) can be found for the problem of two cylinders in an applied field, and this solution is easily generalized to more than two cylinders.

The plan of the paper is as follows: In Sec. 2 is proven the basic result of this paper, as stated above. In Sec. 3 the limiting cases $\epsilon\to\infty$ and $\epsilon\to 0$ are discussed. Section 4 discusses a variant of the method which is necessary for homogeneous problems with conductors, but which also proves useful for inhomogeneous problems. It enables simple derivations of classical results due to Copson[2] and others[3] for a conducting disc in an applied field. Finally, Sec. 5 extends the above results to multiple bodies.

2. ELLIPTIC CYLINDER IN AN ARBITRARY APPLIED POTENTIAL

The basic tool in Ref. 1 was an integral representaton for the electrostatic potential Φ that arises when an applied electric field with potential $\Phi_{\rm app}$ acts on a dielectric body. For the present two-dimensional case it reads

$$\Phi(\boldsymbol{\rho}) = \Phi_{\rm app}(\boldsymbol{\rho}) + (\epsilon - 1)\int_L \frac{\partial\Phi(\boldsymbol{\rho}')}{\partial n'}\bigg|_- g(\boldsymbol{\rho},\boldsymbol{\rho}')\,dL', \tag{4}$$

where ϵ is the dielectric constant of the cylinder, L is the curve defining the cross section (Fig. 1), $\mathbf{n}$ is the outward pointing normal, $\boldsymbol{\rho}$ is a two dimensional vector, and $\partial\Phi/\partial n|_-$ is the normal derivative of Φ at a point just inside L. The Green function $g(\boldsymbol{\rho},\boldsymbol{\rho}')$ is

$$g(\boldsymbol{\rho},\boldsymbol{\rho}') = (1/2\pi)\ln|\boldsymbol{\rho} - \boldsymbol{\rho}'| \tag{5}$$

and the Φ defined by Eq. (4) satisfies the boundary conditions

$$\Phi_+ = \Phi_-, \qquad \frac{\partial\Phi}{\partial n}\bigg|_+ = \epsilon\frac{\partial\Phi}{\partial n}\bigg|_-, \tag{6}$$

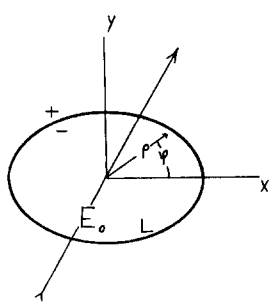

FIG. 1. Elliptic cylinder of dielectric constant ϵ in a uniform external field of magnitude E_0. A point on the inside surface is labeled by $-$, a point on the outside by $+$, and the contour itself is denoted by L.

where the subscripts $+$ and $-$, respectively, denote points just outside and inside L. The analogous equation for the three dimensional case uses the Green function $-1/(4\pi|\mathbf{r}-\mathbf{r}'|)$.[4]

We now apply Eq. (4) to the specific case of the elliptic cylinder of Fig. 1 in which the contour is defined by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{7}$$

and the applied potential is $\Phi^{(n)}_{\rm app}$ or $\tilde{\Phi}^{(n)}_{\rm app}$ of Eq. (2). The procedure will of course be very similar for $\Phi^{(n)}$ and $\tilde{\Phi}^{(n)}$ so we shall simply work out details for the former.

The integral equation (4) for $\Phi^{(n)}_{\rm app}$ is

$$\Phi(\boldsymbol{\rho}) = \rho^n\cos n\varphi + (\epsilon - 1)\int_L \frac{\partial\Phi}{\partial n'}\bigg|_- g(\boldsymbol{\rho},\boldsymbol{\rho}')\,dL'. \tag{8}$$

We need $g(\boldsymbol{\rho},\boldsymbol{\rho}')$ in its real form for $\rho<\rho'$:

$$2\pi g(\boldsymbol{\rho},\boldsymbol{\rho}') = \ln\rho' - \sum_{l=1}^{\infty}(\rho^l/l\rho'^l)\cos l(\varphi - \varphi'). \tag{9}$$

For ρ in the interior we assume for the potential there, $\Phi_{\rm int}$, the expansion

$$\Phi_{\rm int}(\boldsymbol{\rho}) = \sum_{s=0}^{\infty} C_s\rho^s\cos s\varphi. \tag{10}$$

The possible terms in $\rho^s\sin s\varphi$ have been omitted in Eq. (10) since it is clear that the applied potential, which is even about $\varphi = 0$, will also generate an interior potential that is even. The equation of the cylinder can be written in the form

$$\rho = a/[\lambda^2 + (1-\lambda^2)\cos^2\varphi\,]^{1/2} \tag{11}$$

with $a = \lambda b$, and we also have

$$\begin{aligned}\mathbf{n} &= (\hat{\boldsymbol{\rho}}\rho - \hat{\boldsymbol{\varphi}}\,d\rho/d\varphi)\|dL/d\varphi \\ &= (\hat{\boldsymbol{\rho}}\rho - \hat{\boldsymbol{\varphi}}(\rho^3/a^2)(1-\lambda^2)\sin\varphi\cos\varphi)\|dL/d\varphi,\end{aligned} \tag{12}$$

where $dL/d\varphi = [\rho^2 + (d\rho/d\varphi)^2]^{1/2}$. With Eqs. (10) and (12) we find

$$\partial\Phi/\partial n = \mathbf{n}\cdot\nabla\Phi = \frac{1}{2a^2\,dL/d\varphi}\sum_{s=1}^{\infty} sC_s\rho^{s+2}F_s(\varphi), \tag{13}$$

$$F_s(\varphi) = (1+\lambda^2)\cos s\varphi + (1-\lambda^2)\cos(s-2)\varphi. \tag{14}$$

If now Eqs. (9) and (13) are put into Eq. (8), we get, on equating coefficients of $\cos l\varphi$, and for $l\neq 0$,

$$C_l = \delta_{ln} - \sum_{s=1}^{\infty} C_sF_{ls}, \quad l = 1,2,..., \tag{15}$$

where

$$F_{ls} = \frac{(\epsilon-1)s}{4\pi a^2 l}\int_0^{2\pi}\rho^{s-l+2}F_s(\varphi)\cos l\varphi\,d\varphi. \tag{16}$$

For $l = 0$ this procedure gives an equation in which C_0 on the left hand side is calculated from $C_1, C_2, \cdots$ on the right hand side. However, since C_0 is a meaningless constant, we shall forget about this equation and consider that the basic set of equations is Eq. (15) for $l = 1,2,3,\cdots$.

It is important to note that

$$F_{ls} = 0, \quad s + l = \text{odd integer}. \tag{17}$$

This is easily shown by noting that the integral in Eq. (16) from 0 to 2π, which is twice the same integral from 0 to π, is

odd about $\varphi = \pi/2$ for $l + s$ odd. A second and crucial property of F_{ls} is that it is upper triangular:

$$F_{ls} = 0, \quad l \geqslant s + 1. \tag{18}$$

A related property holds for the Q matrices that arise in Waterman's treatment by the *extended boundary-condition method*[5] of scattering from ellipsolids and our proof is based on his. Equation (18) follows obviously from Eq. (17) for l equal to $s+1, s+3, \cdots$ so we consider the nontrivial case for which l equals $s+2, s+4, \cdots$ and define an integer m by

$$l - s = 2 + 2m, \quad m = 0,1,2,\cdots .$$

In terms of m we have, on using Eqs. (11) and (14),

$$\begin{aligned} F_{ls} \propto \int_0^{2\pi} & [1 + (1-\lambda^2)\cos^2\varphi\,]^m \\ & \times \{(1+\lambda^2)\,[\cos(2s+2m+2)\varphi + \cos(2m+2)\varphi\,] \\ & + (1-\lambda^2)[\cos(2s+2m)\varphi + \cos(2m+4)\varphi\,]\}\, d\varphi. \end{aligned} \tag{19}$$

Now the factor $[1 + (1-\lambda^2)\cos^2\varphi]^m$ can be expanded in a series of the functions $\cos 2p\varphi$ where p varies from zero to a maximum value of m so that the highest term in this series is $\cos 2m\varphi$. For m zero or positive every term in this series will be orthogonal to each term in the curly brackets of the integrand of Eq. (19). Thus, F_{ls} is shown to vanish for $l = s+2, s+4, \cdots$, so that Eq. (18) can be considered as proven.

As a consequence of the property of F_{ls} embodied in Eq. (18) we now prove the central result: For any finite n, Eqs. (15) have an explicit solution for which all C_l with $l > n$ vanish. Before proving this in general it is illuminating to first write out explicitly a special case. We choose the case $n = 4$ and then Eqs. (15) have the following form:

$$\begin{aligned} C_1 &= -C_1F_{11} + 0 - C_3F_{13} + 0 - C_5F_{15} + 0\cdots, \\ C_2 &= 0 - C_2F_{22} + 0 - C_4F_{24} + 0 - C_6F_{26}\cdots, \\ C_3 &= 0 + 0 - C_3F_{33} + 0 - C_5F_{35} + 0\cdots, \\ C_4 - 1 &= 0 + 0 + 0 - C_4F_{44} + 0 - C_6F_{46}\cdots, \\ C_5 &= 0 + 0 + 0 + 0 - C_5F_{55} + 0\cdots, \\ C_6 &= 0 + 0 + 0 + 0 - C_6F_{66}\cdots, \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \end{aligned} \tag{20}$$

Note first that the equations below the fourth equaton constitute a homogeneous set that relates the coefficients $C_5, C_6, C_7, \cdots$ to themselves. Since the determinant of this set does not vanish in general, the solution of this set is $C_5 = C_6 = C_7 = \cdots = 0$. This leaves in Eq. (20) the first four equations, but of these the first and third relates C_1 and C_3 to themselves by a homogenenous set whose determinant is nonzero so that $C_1 = C_3 = 0$. We are left with the equations for C_2 and C_4 from which we find

$$C_4 = 1/(1 + F_{44}), \quad C_2 = -C_4F_{24}/(F_{22} + 1).$$

These results can now readily be generalized to arbitrary n. Thus, in Eqs. (15) those equations beginning with the $(n+1)$st constitute a homogeneous set for $C_{n+1}, C_{n+2}, \cdots$ and we conclude that $C_{n+1} = C_{n+2} = C_{n+3} = \cdots 0$. Further, if for example n is even, the odd coefficients $C_1, C_3, \ldots, C_{n-1}$ are related to themselves by a homogeneous set and they too vanish. Of the even coefficients, $C_n = 1/(1 + F_{nn})$ from the nth equation, and given this one can work down in the set of equations to successively find $C_{n-2}, C_{n-4}, \cdots$. Thus we have the result claimed above: The interior potential for $\Phi^{(n)}_{\text{app}}$ has an expansion like Eq. (3) with no terms higher than $\rho^n \cos n\varphi$. Obviously, a similar result holds for n odd and for an applied potential $\tilde{\Phi}^{(n)}_{\text{app}}$ for which the interior expansion is in terms of the functions $\rho^n \sin n\varphi$.

As a simple example of these results consider a dielectric cylinder in a uniform field of magnitude $\mathscr{E}_0$ that makes an angle γ with the x axis. Then the applied potential is

$$\begin{aligned} \Phi_{\text{app}} &= -\mathscr{E}_0(x\cos\gamma + y\sin\gamma) \\ &= -\mathscr{E}_0(\Phi^{(1)}_{\text{app}}\cos\gamma + \tilde{\Phi}^{(1)}_{\text{app}}\sin\gamma). \end{aligned} \tag{21}$$

By applying the results above to the linear combination of $\Phi^{(1)}_{\text{app}}$ and $\tilde{\Phi}^{(1)}_{\text{app}}$ embodied in Eq. (21), it is easy to show that the potential Φ_{int} inside the cylinder is

$$\Phi_{\text{int}} = -\mathscr{E}_0(a+b)\left(\frac{x\cos\gamma}{a+\epsilon b} + \frac{y\sin\gamma}{\epsilon a + b}\right). \tag{22}$$

This result can also be derived by conformal mapping.[6]

As the simplest example after the above one for $n = 1$, take $n = 2$ in Eq. (15). Then only C_2 is nonvanishing in Eq. (10) and it is given by $C_2 = 1/(1 + F_{22})$, where

$$F_{22} = \frac{2(\epsilon - 1)\lambda}{(1+\lambda)^2}. \tag{23}$$

The three dimensional problem of an ellipsoid in an arbitrary applied potential can be treated in much the same way as above; our previous discussion[1] of a prolate spheroid in a uniform field is an example of a special case of such treatment.

3. DIRICHLET-ZERO AND NEUMANN-ZERO PROBLEMS AS LIMITING CASES

In this section we consider the solutions above in the limits $\epsilon \to \infty$ and $\epsilon \to 0$. We show that in the limit $\epsilon \to \infty$ these become the solutions of the problem of an elliptic cylinder or ellipsoid subject to the homogeneous Dirichlet conditions $\Phi_+ = 0$ on the surface; physically, this is the problem of the grounded perfect conductor. For $\epsilon = 0$ the solutions become those appropriate to the homogeneous Neumann condition $\partial\Phi/\partial n|_+ = 0$, the boundary condition of ideal fluid flow.

To show these limiting connections we take as basic Green's theorem that expresses Φ in the region outside L in terms of Φ_+ and $\partial\Phi/\partial n|_+$, the potential on the outside surface and its normal derivative there, respectively. In its complete form,[7,8] in which $\boldsymbol{\rho}$ can be either inside or outside L, it reads

$$\begin{aligned} \Phi^{(\boldsymbol{\rho})}_{0}\Big\} &= \Phi_{\text{app}}(\boldsymbol{\rho}) - \int \left((\Phi_+(\boldsymbol{\rho}') \frac{\partial g(\boldsymbol{\rho},\boldsymbol{\rho}')}{\partial n'} \right. \\ &\quad \left. - g(\boldsymbol{\rho},\boldsymbol{\rho}') \frac{\partial \Phi(\boldsymbol{\rho}')}{\partial n'}\bigg|_+ \right) dL', \quad \begin{cases} \boldsymbol{\rho} \text{ outside } L, \\ \boldsymbol{\rho} \text{ inside } L \end{cases}. \end{aligned} \tag{24}$$

There is a counterpart to this, which is less frequently quoted, but which is derived in much the same way as Eq. (22), in which in the surface integration the potential and its derivative are evaluated just inside L:

$$\left.\begin{matrix}\Phi(\rho)\\ 0\end{matrix}\right\} = \int \left(\Phi_-(\rho')\frac{\partial g(\rho,\rho')}{\partial n'} - g(\rho,\rho')\frac{\partial \Phi(\rho')}{\partial n'}\Big|_-\right) dL', \quad \begin{cases}\rho \text{ inside } L,\\ \rho \text{ outside } L.\end{cases} \tag{25}$$

We apply the upper integral identity in Eq. (24) to the Dirichlet-zero problem, i.e., for $\Phi_+ = 0$, whence it becomes

$$\Phi(\rho) = \Phi_{app}(\rho) + \int_L \frac{\partial \Phi(\rho')}{\partial n'}\Big|_+ g(\rho,\rho')\, dL', \quad \rho \text{ outside } L. \tag{26}$$

However, the basic integral equation (4) reduces to just this last equation in the limit $\epsilon \to \infty$, since in that limit we have, on applying the boundary condition (6),

$$(\epsilon - 1)\frac{\partial \Phi}{\partial n'}\Big|_- \approx \epsilon \frac{\partial \Phi}{\partial n'}\Big|_- = \frac{\partial \Phi}{\partial n'}\Big|_+ .$$

Thus, as we have remarked previously, the solution to a problem with $\Phi_+ = 0$ on the boundary of a given shape is derived from the solution for a body of that shape[9] of dielectric constant ϵ by letting $\epsilon \to \infty$.

The Neumann-zero problem is that of finding a potential that satisfies $\nabla^2\Phi = 0$ outside the body, becomes Φ_{app} at infinity, and satisfies the boundary condition $\partial\Phi/\partial n|_+ = 0$. From its derivation the Φ of Eq. (4) satisfies the first two of these conditions and from Eq. (6) satisfies the third as well if $\epsilon = 0$. It then appears that any solution of Eq. (4) satisfies the Neumann-zero conditons $\partial\Phi/\partial n|_+ = 0$ in the limit $\epsilon \to 0$.

To confirm this result from the point of view of the Green theorem of Eqs. (24) and (25) we first write the integral equation (4) in the limit $\epsilon = 0$, in which case it becomes

$$\Phi(\rho) = \Phi_{app}(\rho) - \int \frac{\partial \Phi(\rho')}{\partial n'}\Big|_- g(\rho,\rho')\, dL'. \tag{27}$$

On the other hand, a standard integral equation for the Neumann-zero case is derived from the upper line of the Green theorem of Eq. (24) by using the boundary condition $\partial\Phi/\partial n|_+ = 0$ in Eq. (24) in which case it becomes

$$\Phi(\rho) = \Phi_{app}(\rho) - \int \Phi_+(\rho')\frac{\partial g(\rho,\rho')}{\partial n'} dL', \quad \rho \text{ outside } L. \tag{28}$$

The lower of the relations in Eq. (25) and the further fact that $\Phi_+ = \Phi_-$ can now be used to show that for ρ outside L, Eq. (27) is in fact identical with the standard Eq. (28).

As an example of these limiting cases we calculate the charge density per unit length $\sigma = -(1/4\pi)\partial\Phi/\partial n|_+$ for the perfectly conducting elliptic cylinder in the uniform applied field of Eq. (21). The solution to the analagous problem with dielectric constant ϵ is given by Eq. (22). With $b = \mu a$ in this equation we have, in the limit $\epsilon \to \infty$,

$$-4\pi\sigma = \frac{\partial \Phi}{\partial n}\Big|_+ = \lim_{\epsilon\to\infty}\left(\epsilon\frac{\partial \Phi}{\partial n}\right)\Big|_-$$

$$= \lim_{\epsilon\to\infty}(\mathbf{n}\cdot\nabla(\epsilon\Phi_{int}))$$

$$= -\mathscr{E}_0(1+\mu)(x\mu\cos\gamma + y\sin\gamma)/(\mu^4x^2 + y^2)^{1/2}.$$

The case of the strip is obtained by taking the limit $\mu \to 0$. For the applied field parallel to the strip ($\gamma = 0$) we find, having used $y^2 = \mu^2(a^2 - x^2)$ from Eq. (7),

$$\sigma_{strip} = \frac{\mathscr{E}_0 x}{4\pi\sqrt{a^2 - x^2}},$$

which shows the well-known singularity near the edges.

As a second example we find, using Eq. (23) and $\mu = 1/\lambda$, the charge density on the perfectly conducting cylinder in the applied field $\Phi^{(2)}_{app}$,

$$\sigma = \frac{(1+\mu)^2(a^2 - 2x^2)}{4\pi(\mu^2x^2 + (a^2 - x^2))^{1/2}}.$$

In the limit $\mu \to 0$ this becomes the charge density on a strip

$$\sigma_{strip} = \frac{a^2 - 2x^2}{4\pi(a^2 - x^2)^{1/2}},$$

again with the characteristic singularity at the edges.

For the solution to the Neumann-zero problem we first set $\epsilon = 0$ in Eq. (22) to get

$$\Phi_{int} = -\mathscr{E}_0(a+b)\left(\frac{x\cos\gamma}{a} + \frac{y\sin\gamma}{b}\right). \tag{29}$$

Of course, Φ_{int} is not of physical interest for the Neumann problem but one can use Eq. (29) to calculate $\partial\Phi/\partial n'|_-$ and put it into Eq. (27) whereby $\Phi(\rho)$ can be calculated in the physically interesting exterior region. A check can be made on the correctness of Eq. (29), namely, the Neumann-zero problem can be considered to be one of ideal fluid flow if E_0 is taken to be the speed of the uniform flow at infinity. A quantity of particular interest in fluid flow is v_t, the tangential velocity at the surface, $v_t = -\hat{\mathbf{t}}\cdot\nabla\Phi_+$, where $\hat{\mathbf{t}}$ is the unit tangent vector. Now since Φ is continuous across L so is its tangential derivative, or $\mathbf{t}\cdot\nabla\Phi_- = \mathbf{t}\cdot\nabla\Phi_+$. With

$$\hat{\mathbf{t}} = \left(\frac{\hat{\mathbf{x}}y}{b^2} - \frac{\hat{\mathbf{y}}x}{a^2}\right)\Big/\left(\frac{x^2}{a^4} + \frac{y^2}{b^4}\right)^{1/2},$$

we have then from Eq. (29)

$$v_t = -\hat{\mathbf{t}}\cdot\nabla\Phi_{int} = \mathscr{E}_0\frac{(a+b)}{ab}\left(\frac{x\sin\gamma}{a} - \frac{y\cos\gamma}{b}\right)\Big/\left(\frac{x^2}{a^4} + \frac{y^2}{b^4}\right)^{1/2}. \tag{30}$$

This is just the result obtained by standard methods for the problem,[10,11] e.g., conformal mapping, or the separation of Laplace's equation in elliptical coordinates.

4. VARIANT METHOD FOR PERFECT CONDUCTORS

Although the method above for a dielectric body in an applied potential (an inhomogeneous problem) yields the solution for a perfect conductor in the limit $\epsilon \to \infty$, there is another class of problems with perfect conductors for which it must be modified. These are *homogeneous* problems, in which there is *no* applied field. In a typical homogeneous problem one might specify the total charge on a perfect conductor, or the (constant) potential on its surface, and ask for the potential throughout space.

The need to modify the method emerges if we try to apply the basic equation (4) to an homogeneous problem by simply letting Φ_{app} be zero. For arbitrary ϵ the resulting equation then has no solution. It becomes an eigenvalue equation,[12,13] that has solutions only for certain special val-

ues of ϵ. With the homogeneous version of Eq. (4) we cannot then adopt the procedure of solving for arbitrary ϵ, and then letting $\epsilon \to \infty$. This mathematical result has its physical counterpart. In an inhomogeneous problem, whether ϵ is large or small the same physical picture applies: Polarization charge appears on the surface of the body as the result of the action of the applied field; this polarization charge becomes the free surface charge of conductors as $\epsilon \to \infty$. However, for a homogeneous problem there is no counterpart to this physical picture for *finite* ϵ. There is no polarization mechanism without an applied field, and if one imagines charge sprayed on the surface of a dielectric, this charge is not free to move and redistribute itself *except* in the limit $\epsilon \to \infty$. Thus, for the homogeneous problem one must take the limit $\epsilon \to \infty$ from the beginning.

Although it is the homogeneous problem that forces the modification of the method of Secs. 2 and 3, this modification has advantages for the inhomogeneous one as well and we shall discuss both homogeneous and inhomogeneous problems on the same footing in this section. The basic tool will be an integral representation that can be derived from Eqs. (24) and (25) but which is simpler to write down from physical arguments. It is the three dimensional analog ($\rho \to \mathbf{r}, L \to S$) of Eq. (26) when the field point is in the interior. Equation (26) has an obvious physical interpretation if we remember that the surface charge distribution on a perfect conductor is proportional to $\partial\Phi/\partial n|_+$. The equation simply states that the potential outside a perfect conductor is the sum of the applied potential and the potential due to the surface charge. Now the potential *inside* a perfect conductor is a similar sum but it is well known that this potential is a constant, say Φ_0. The mathematical transcription of this last sentence is then the desired integral representation.[14]

$$\Phi_0 = \Phi_{\text{app}}(\mathbf{r}) + \int_S \left.\frac{\partial\Phi(\mathbf{r}')}{\partial n'}\right|_+ g(\mathbf{r},\mathbf{r}')\,dS', \quad \mathbf{r} \text{ inside } S. \tag{31}$$

We first consider Eq. (31) as applied to the problem of free charge ($\Phi_{\text{app}} = 0$) on a perfectly conducting spheroid. It is then

$$\Phi_0 = \int_S \left.\frac{\partial\Phi(\mathbf{r}')}{\partial n'}\right|_+ g(\mathbf{r},\mathbf{r}')\,dS', \quad \mathbf{r} \text{ inside } S. \tag{32}$$

In standard r, θ, φ spherical coordinates, with the z axis the axis of symmetry, the equation of the spheroid is

$$r = \xi a/[\xi^2 + (1-\xi^2)\cos^2\theta\,]^{1/2}, \tag{33}$$

where the semiminor axis is c, the semimajor is a, and $\xi = c/a$.

We now introduce a new function $\Gamma(\Omega)$ defined by

$$\Gamma(\Omega) = \frac{1}{r^3}\left.\frac{\partial\Phi}{\partial n}\right|_+ \frac{dS}{d\Omega}, \tag{34}$$

where Ω stands for θ, φ and $d\Omega$ is an element of solid angle. In terms of $\Gamma(\Omega)$, Eq. (32) becomes

$$\Phi_0 = \int \Gamma(\Omega')r'^3 g(\mathbf{r},\mathbf{r}')\,d\Omega', \quad \mathbf{r} \text{ inside } S. \tag{35}$$

We shall use the real even and odd spherical harmonics Y^e_{lm} and Y^0_{lm} defined by

$$\left.\begin{matrix} Y^e_{lm} \\ Y^0_{lm} \end{matrix}\right\} = \left(\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right)^{1/2} P_{lm}(\cos\theta) \begin{cases} \cos m\varphi, \\ \sin m\varphi \end{cases}.$$

The Green function expansion for $r < r'$ is then

$$\begin{aligned} g(\mathbf{r},\mathbf{r}') &= -1/4\pi\,|\mathbf{r}-\mathbf{r}'| \\ &= -\sum_{l=0}^{\infty}\sum_{m=0}^{l}\frac{\epsilon_m}{(2l+1)}\frac{r^l}{r'^{l+1}} \\ &\quad\times[Y^e_{lm}(\Omega)\,Y^e_{lm}(\Omega') + Y^0_{lm}(\Omega)\,Y^0_{lm}(\Omega')], \end{aligned} \tag{36}$$

where ϵ_m is the Neumann factor. Now $\Gamma(\Omega)$ is clearly independent of φ and we assume it can be expanded in the following series involving Legendre polynomials:

$$\Gamma(\Omega') = \sum_{l'=0}^{\infty} A_{l'} r'^{l'} P_{l'}(\cos\theta'). \tag{37}$$

If now Eqs. (36) and (37) are put into Eq. (35) and the coefficients of $r^l P_l(\cos\theta)$ are equated on both sides, we get the following set of equations:

$$\Phi_0\delta_{l0} = \sum_{l'=0}^{\infty} A_{l'} M_{ll'}, \tag{38}$$

where

$$M_{ll'} = -\frac{(\xi a)^{l'-l+2}}{2}\int_{-1}^{1}\frac{P_l(x)P_{l'}(x)\,dx}{[\xi^2+(1-\xi^2)x^2]^{(l'-l+2)/2}}.$$

It follows from the fact that $P_l(x)$ is even in x for l even and odd for l odd that $M_{ll'}$ is zero unless l and l' are both even or both odd. An integer s is now introduced by

$$l = l' + 2 + 2s,$$

whence

$$\begin{aligned} M_{ll'} &= -\tfrac{1}{2}(\xi a)^{-2s} \\ &\quad\times\int_{-1}^{1}[\xi^2+(1-\xi^2)x^2]^s \\ &\quad\times P_{l'}(x)P_{l'+2+2s}(x)\,dx. \end{aligned}$$

For $s \geqslant 0$ the product $P_{l'}P_{l'+2+2s}$ can be expanded in a series of Legendre polynomials whose order ranges from P_{2s+2} to $P_{2l'+2s+2}$. Also, $[\xi^2+(1-\xi^2)x^2]^s$ can be expanded in a similar series in which the polynomials range from P_0 to P_{2s}. Every polynomial in this latter series will be orthogonal to every one in the first mentioned series and $M_{ll'}$ therefore vanishes for $s \geqslant 0$. We have thus proved that $M_{ll'}$ is upper triangular

$$M_{ll'} = 0, \quad l \geqslant l'+1, \tag{39}$$

and it follows by arguments very similar to those in Sec. 2 that all A_l except A_0 are zero. From Eq. (38) we have

$$A_0 = \Phi_0/M_{00}.$$

The matrix element M_{00} has different forms for $\xi < 1$ (oblate spheroid) and $\xi > 1$ (prolate spheroid). For the former case

$$M_{00} = -\frac{\xi a^2 \tan^{-1}\sqrt{1/\xi^2-1}}{\sqrt{1-\xi^2}}.$$

From this and Eq. (34) we find for the charge density σ on the oblate spheroid

$$\sigma = -\frac{1}{4\pi}\left.\frac{\partial\Phi}{\partial n}\right|_+$$

$$= \frac{\Phi_0(1-\xi^2)^{1/2}[\xi^2+(1-\xi^2)\cos^2\theta\,]^{1/2}}{4\pi a\,\tan^{-1}(1/\xi^2-1)^{1/2}\,[\xi^4+(1-\xi^4)\cos^2\theta\,]^{1/2}}, \quad (40)$$

which is just the standard result.[8,15]

We consider next the inhomogeneous problem of an oblate spheroid in an applied potential. As we have remarked, this problem is of special interest since in the limit of zero thickness the ellipsoid becomes a disc, and there are classical results on this "problem of the electrified disc" with which to compare. In particular, Copson[2] has solved the problem of such a disc in a fairly general applied field, and we shall be able to compare our results with his.

For purposes of this comparison we here summarize Copson's findings. Consider a disc of unit radius that is centered in the plane $z=0$ of a standard spherical coordinate system. Copson solves that canonical problem of potential theory in which the potential Φ is specified on the surface of the disc, and one seeks $\partial\Phi/\partial n$ there. His solution is expressed as follows: Let the surface potential distribution be given by

$$\Phi_{\rm disc} = \sum_{n=0}^{\infty} f_n(\rho)\cos n(\varphi-\varphi_n). \quad (41)$$

Then the corresponding quantity $\partial\Phi/\partial n$, expressed in terms of the charge distribution $\sigma=(-1/4\pi)(\partial\Phi/\partial n)$, is

$$\sigma(\rho,\varphi) = \sum_{n=0}^{\infty} \sigma_n(\rho)\cos n(\varphi-\varphi_n), \quad (42)$$

where

$$\sigma_n(\rho) = -\frac{2\rho^{n-1}}{\pi}\frac{d}{d\rho}\int_\rho^1 \frac{tS_n(t)\,dt}{\sqrt{t^2-\rho^2}} \quad (43)$$

and

$$S_n(t) = \frac{1}{2\pi t^{2n}}\frac{d}{dt}\int_0^t \frac{u^{n+1}f_n(u)\,du}{\sqrt{t^2-u^2}}. \quad (44)$$

These results can be applied, for example, to the homogeneous problem of the charge distribution on an isolated perfectly conducting charged disc. For this case $\Phi_{\rm disc}$ is a constant, conveniently taken to be unity so that $f_n(\rho)=0$, $n\neq 0$. Equations (43) and (44) then yield the standard result for the charge density, *twice* the limit of Eq. (40) as $\xi\to 0$ (the disc has two sides):

$$\sigma = \frac{1}{\pi^2\sqrt{1-\rho^2}}.$$

Equations (42) through (44) also solve the problem of the grounded disc in certain applied fields. Suppose the applied potential $\Phi_{\rm app}(r,\theta,\varphi)$ is such that *on the disc* it has the form, with reference to Eq. (41),

$$\Phi_{\rm app}(r,\pi/2,\varphi) = f_n(\rho)\cos n(\varphi-\varphi_n). \quad (45)$$

Suppose also that the charge distribution on the disc is $-\sigma_n(\rho,\varphi)$. Then on the disc the potential $f_n\cos n(\varphi-\varphi_n)$ of the applied field plus the potential $-f_n\cos n(\varphi-\varphi_n)$ produced by the charge $-\sigma_n$ on the disc yields a zero total potential, as required for the grounded charged disc. For the self-consistency of this argument note that the applied potential yields no net contribution to the charge density on the disc, since $\partial\Phi_{\rm app}/\partial n$ has opposite signs on the upper and lower faces.

For reference we now calculate σ for the case

$$f_n(\rho)=\rho^n. \quad (46)$$

On substituting this value for f_n in Eq. (44), and making the change of variable $u=t\cos\varphi$, we find

$$S_n = \frac{2n+1}{2\pi}\int_0^{\pi/2}\cos^{2n+1}\varphi\,d\varphi = \frac{(2n)!!}{2\pi(2n-1)!!}.$$

Then from Eq. (43)

$$\sigma_n(\rho) = \frac{(2n)!!}{\pi^2(2n-1)!!}\frac{\rho^n}{\sqrt{1-\rho^2}}. \quad (47)$$

To solve the above grounded disc problem with the present model we consider the oblate spheroid in applied potentials which, in the limit that the spheroid becomes a disc, become those applied potentials that are considered by Copson. Now the *general* applied potential is a linear superposition over l and m of terms $\phi^{\,l}_{\tilde l\tilde m,\rm app}$ and $\phi^{\,0}_{\tilde l\tilde m,\rm app}$ defined by, for example,

$$\phi^{\,e}_{\tilde l\tilde m,\rm app} = r^{\tilde l}Y^e_{\tilde l\tilde m}. \quad (48)$$

Consider the special case $\tilde m=\tilde l$. On the surface of the disc we have, with $P_{ll}(0)=(2l-1)!!$

$$\phi^{\,e}_{\tilde l\tilde l,\rm app}(r,\pi/2,\varphi) = \left(\frac{2\tilde l+1}{4\pi(2\tilde l)!}\right)^{1/2}(2\tilde l-1)!!\;\rho^{\tilde l}\cos\tilde l\varphi$$
$$= K\rho^{\tilde l}\cos\tilde l\varphi. \quad (49)$$

Except for a phase factor this is proportional to the applied potentials of Eqs. (45) and (46) that are considered by Copson and it is these applied potentials that we shall use for the disc problem. Before we consider this special problem, however, it is illuminating to look briefly at the problem of an arbitrary grounded ellipsoid, with x,y,z semiaxes of magnitudes a,b,c in a general potential of the form of Eq. (48). Equations for the solution of this problem can be derived by a technique very similar to that in Sec. 2 for the cylinder, and to that discussed above for the homogenous spheroid problem, namely, in the integral equation (31) with $\Phi_0=0$ (grounded conductor) and with $\Phi_{\rm app}$ given by Eq. (48), the function $\Gamma(\Omega)$ of Eq. (34) is expanded as

$$\Gamma(\Omega') = \sum_{l'=0}^{\infty}\sum_{m'=0}^{l'} C_{l'm'}r'^{l'}\,Y^e_{l'm'}(\Omega'), \quad (50)$$

and the Green function expansion (36) is used to find the set of equations

$$\delta_{l\tilde l}\delta_{m\tilde m} = \sum_{l'=0}^{\infty}\sum_{m'=0}^{l'} C_{l'm'}T_{lm,l'm'}, \quad (51)$$

where

$$T_{lm,l'm'} = \frac{\epsilon_m}{2l+1}\int \frac{Y^e_{lm}\,Y^e_{l'm'}}{r'^{\,l-l'-2}}\,d\Omega'. \quad (52)$$

Certain special properties of the matrix element $T_{lm,l'm'}$ enable the solution of Eqs. (51) and we note these now. These properties derive from the fact that $r'^{\,l-l'-2}$ in the integrand of Eq. (52) is invariant under reflection in the planes $x=0$, $y=0$, and $z=0$ whereas the product $Y^e_{lm}\,Y^e_{l'm'}$ $\propto P_{lm}(\cos\theta')P_{l'm'}(\cos\theta')\cos m\varphi'\cos m'\varphi'$ is either even or odd with respect to such reflections, depending on the values of the indices l,m,l',m'. If it is odd in any such reflection, then

$T_{lm,l'm'}$ vanishes. With $\eta = \cos\theta$ consider the product $P_{l,m}(\eta)P_{l',m'}(\eta)$. Since $P_{lm}(\eta)$ is even in η for $l+m$ even and odd for $l+m$ odd, the matrix element $T_{lm,l'm'}$ vanishes unless the sum $l+m$ and the sum $l'+m'$ are both even or both odd, i.e., unless $l+m+l'+m' =$ even. The requirement of evenness about the plane $y=0$ is trivial since by definition Y^e_{lm} satisfies it. However, the requirement of evenness about the plane $x=0$ is not trivial and in fact it is readily shown to imply $l+l' =$ even. Thus, the matrix element $T_{lm,l'm'}$ vanishes when $l = l'+1, l'+3$ just as for the two dimensional case of Sec. 2. Beyond this one can show that $T_{lm,l'm'}$ also vanishes for $l = l'+2, l'+4, \cdots$. The proof of this is similar to that for Eq. (39) except that in the present case the addition theorem for the product of two *associated* Lengendre polynomials must be used. We have thus shown that $T_{lm,l'm'}$ is upper triangular

$$T_{lm,l'm'} = 0, \quad l > l'. \tag{53}$$

Much as in Sec. 2 it then follows from Eq. (51) that the only $C_{l'm'}$ that can occur in the expansion of Eq. (50) are those for which $l' \leqslant l$. It further follows from these equations that of this possible set of $C_{l'm'}$ the only ones which are in fact nonvanishing are those for which $l'+m'$ has the same parity as $\tilde{l}+\tilde{m}$. Alternately, one can draw the latter conclusion simply by remarking that it is clear physically that the response of the ellipsoid to the applied potential must have the same evenness or oddness properties around $z=0$ as the applied potential itself. As an example of the above remarks, for $\tilde{l}=4, \tilde{m}=2$ the nonvanishing terms in Eq. (50) are $C_{44}, C_{42}, C_{40}, C_{33}, C_{31}, C_{22}, C_{20}, C_{00}$.

Consider next the specialization that occurs when the ellipsoid becomes a spheroid in that the axes a and b become identical. At any given θ the cross section of the spheroid is a circle and this implies that the charge distribution, and hence Γ, must have the same dependence on φ as the applied potential. In turn, this condition selects out from the $C_{l'm'}$ allowed above for the general ellipsoid those for which $m' = \tilde{m}$. Thus, for the case $\Phi^e_{42,\text{app}}$ just considered, only the coefficients C_{42} and C_{22} are nonvanishing for the spheroid.

Finally, in the special case of the applied potential for which $\tilde{m} = \tilde{l}$, which applies to Eq. (49), it is clear that the only allowed $C_{l'm'}$ is $C_{\tilde{l}\tilde{l}}$ since we must have $m' = \tilde{m}$, and $C_{\tilde{l}\tilde{l}}$ is the only coefficient that can satisfy this condition, in addition to the condition $l' \geqslant m'$. This remark yields the solution to Eq. (51) for the oblate spheroid in $\Phi^e_{\tilde{l}\tilde{l},\text{app}}$:

$$C_{\tilde{l}\tilde{l}} = 1/T_{\tilde{l}\tilde{l},\tilde{l}\tilde{l}}, \tag{54}$$

where, from Eqs. (33) and (52),

$$T_{\tilde{l}\tilde{l},\tilde{l}\tilde{l}} = \frac{\xi^2 a^2((2\tilde{l}-1)!!)^2}{(2\tilde{l})!!} I(\xi),$$

with

$$I(\xi) = \int_0^1 \frac{(1-x^2)^{\tilde{l}}\,dx}{\xi^2 + (1-\xi^2)x^2}.$$

From Eq. (34) we find

$$\left.\frac{\partial\Phi}{\partial n}\right|_+ = \frac{r^{\tilde{l}} Y^e_{\tilde{l}\tilde{l}}}{T_{\tilde{l}\tilde{l},\tilde{l}\tilde{l}}(1/r^3)(dS/d\Omega)}.$$

We use

$$\frac{1}{r^3}\frac{dS}{d\Omega} = \frac{r}{\xi^2 a^2}[\xi^4 + (1-\xi^4)\cos^2\theta]^{1/2}$$
$$= [\xi^2\rho^2 + (a^2-\rho^2)]^{1/2}/\xi a^2$$

to find

$$\left.\frac{\partial\Phi}{\partial n}\right|_+ = \frac{(2\tilde{l})!!\, r^{\tilde{l}} Y^e_{\tilde{l}\tilde{l}}}{(2\tilde{l}-1)!!\,\xi I(\xi)\,[\xi^2\rho^2 + (a^2-\rho^2)]^{1/2}}. \tag{55}$$

Now $I(\xi)$ can be evaluated in closed but lengthy form, so Eq. (55) solves the problem of the charge distribution on the oblate spheroid in the applied field of Eq. (49). Of particular interest is the limit $\xi \to 0$, where the spheroid become a disc. In this limit we have

$$\lim_{\xi\to 0}[\xi I(\xi)] = \pi/2$$

so that the charge density σ is

$$\sigma = -\frac{1}{4\pi}\left.\frac{\partial\Phi}{\partial n}\right|_+ = -\frac{K(2\tilde{l})!!\,\rho^{\tilde{l}}}{2\pi^2(2\tilde{l}-1)!!\sqrt{a^2-\rho^2}},$$

where K is the proportionality constant of Eq. (49). This is just Copson's result (47) if we remember that this equation above gives the charge on only one side of the disc, and that for comparing with Copson we should set K equal to unity and take $-\sigma_n$ as the charge on the disc.

5. MULTIPLE BODIES

In this section we extend the above technique to two or more elliptic cylinders or ellipsoids, and hence to the various limits implied by letting ϵ become zero or infinity and by letting the cylinders and ellipsoids become strips or discs. This makes possible the solution to a large variety of problems: ideal fluid flow around two or more strips and through the slots they may form; flow through a screen formed by an array of strips or discs; calculation of field intensities in an ensemble of dielectric or conducting discs or needles, and so on.

As we remarked in Sec. 1, the basic idea is that in the problem involving several bodies one can think of any one of them, say the ith, as being in the sum of the applied potential and the *effective applied potential* that is generated by the other bodies. This effective potential could be derived, in principle, by making multipole expansions for the potentials generated by all except the ith body, in appropriate coordinates, and then expressing these potentials in terms of the coordinates of the ith body. Such multipole expansions are possible if the individual bodies can be inscribed in circles or spheres, in two or three dimensions, that do not intersect. This fact puts a geometrical constraint on the configurations to which the method can be applied. These multipole expansions are infinite expansions in powers of a_i/d_{ij}, where a_i is a typical dimension of the jth body and d_{ij} is a typical distance from the ith to the jth, and when they are expressed in terms of the coordinates of the ith system they are, in two dimensions, of the form of the applied potentials of Eq. (1).

The above remarks, which apply to bodies of arbitrary shape, are not helpful in general in solving problems of *multiple* arbitrary bodies. One cannot solve the problem of a *single* such body in a uniform applied potential, let alone in the more complicated effective potentials that other bodies

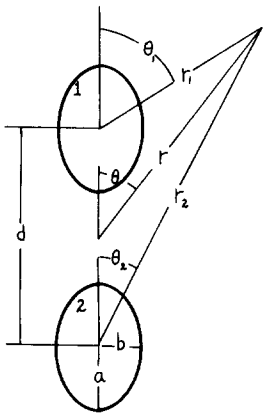

FIG. 2. Two dielectric prolate ellipsoids a distance d apart on the z axis.

generate. The situation is however different for elliptical and ellipsoidal bodies. As this paper has shown one can exactly solve problems involving elliptic cyllinders and ellipsoids in applied potentials that are *finite* series of the form of Eq. (1). The multipole effective potentials are *infinite* series of just this form, but to any given power of a_i/d_{ij} they become finite ones. In any specific problem, then, we shall see that one can decide in advance on the order of the powers of a_i/d_{ij} that one wishes to retain as an approximation, and having done that solve the truncated problem exactly thereafter.

The above remarks are meant merely to supply physical understanding and motivation. They gloss over the essential point that the effective applied potential that one body exerts on the other is not known until the problem is solved. However, they will serve as a guide for setting up the equations that determine this effective potential.

Although the method to be presented here can deal with several or many bodies, for simplicity we shall consider only two; the generalization to three or more bodies is relatively straightforward. Similarly, the two bodies can be arbitrarily oriented with respect to each other and can be in an arbitrary applied field, but the discussion of this very general case is complicated in a way that is merely fussy. We shall therefore limit the discussion to the case of two prolate ellipsoids, one above the other on the z axis. The geometry is shown in Fig. 2. Both the $\mathbf{r}_1$ and $\mathbf{r}_2$ coordinate systems shown in that figure will be used and we shall need the following transformation formula valid for $r_1 < d$:

$$\frac{P_s(\cos\theta_2)}{r_2^{s+1}} = \frac{1}{d^{s+1}} \sum_{l=0}^{\infty} (-)^l \frac{(l+s)!}{l!s!} \left(\frac{r_1}{d}\right)^l P_l(\cos\theta_1). \tag{56}$$

The more general transformation formulas of this kind that are needed for *arbitrary* relative translations of the two coordinate systems can be found in Ref. 16.

Before embarking on the formalism it is instructive to show how the solution, to lowest order of powers of a/d, can be derived from simple physical arguments. With reference then to Fig. 2 we assume that a uniform field $\mathscr{E}_0$ is applied in the z direction and we wish to find the potential inside each of the two ellipsoids, at least approximately. We shall need the formula for the internal field that arises in a *solitary* ellipsoid when it is in a uniform applied field Φ_{app} given by

$$\Phi_{\text{app}} = -\mathscr{E}_0 z = -\mathscr{E}_0 r\cos\theta. \tag{57}$$

Let the ellipsoid be oriented as in Fig. 2, but centered at the origin. Then the potential Φ_{int} inside the ellipsoid is

$$\Phi_{\text{int}} = -k\mathscr{E}_0 r\cos\theta, \tag{58}$$

where

$$k = 1/\left(1 + (\epsilon - 1)\left(1 - \frac{1}{e^2}\right)\left(1 - \frac{1}{2e}\ln\frac{1+e}{1-e}\right)\right),$$

and the eccentricity $e = (1 - b^2/a^2)^{1/2}$. Also, the external potential Φ_{ext}, i.e., the potential that is produced by the polarized ellipsoid outside its circumscribing sphere, is in the dipole approximation

$$\Phi_{\text{ext}} = k\mathscr{E}_0(\epsilon - 1)ab^2\cos\theta/3r^2.$$

To return to the two ellipsoids of Fig. 2, the applied potential of the uniform field $\mathscr{E}_0$ is still given by Eq. (58), but in r_1 coordinates it has an extra trivial constant since $z = z_1 + (d/2)$:

$$\Phi_{\text{app}} = -\mathscr{E}_0 r_1\cos\theta_1 - \mathscr{E}_0 d/2. \tag{59}$$

If it acted by itself, this potential would generate in the interior of ellipsoid #1 the potential

$$-k\mathscr{E}_0 r_1\cos\theta_1 - \mathscr{E}_0 d/2. \tag{60}$$

In addition, however, there is a contribution to the potential inside #1 due to the effect of the polarized ellipsoid #2 acting on it. To the lowest order approximation considered here we can take the polarization of ellipsoid #2 to be due only to the action of the uniform applied potential. Thereby, it produces an external potential $k\mathscr{E}_0(\epsilon - 1)ab^2\cos\theta_2/3r_2^2$. We use Eq. (56) to express this in r_1 coordinates:

$$k\mathscr{E}_0 ab^2(\epsilon - 1)\cos\theta_2/3r_2^2 \approx \frac{k\mathscr{E}_0 ab^2(\epsilon - 1)}{3}\left[\frac{1}{d^2} - \frac{2r_1}{d^3}\cos\theta_1 + \cdots\right]. \tag{61}$$

The first term in the square brackets of Eq. (61) is a trivial constant but the second term is the potential of an effective uniform applied field of magnitude $2k\mathscr{E}_0 ab^2(\epsilon - 1)/3d^3$. The first dielectric ellipsoid responds to this effective field by producing a uniform interior potential of strength k times this, i.e., a potential $-2k^2\mathscr{E}_0 ab^2(\epsilon - 1)r_1\cos\theta_1/3d^3$. Thus, the total internal potential in ellipsoid #1, the sum of that produced by the applied field and that produced by the effective applied field, is

$$\Phi_{\text{int}} = -k\mathscr{E}_0 r_1\cos\theta_1\left(1 + \frac{2kab^2(\epsilon - 1)}{3d^3}\right) + \mathscr{E}_0\left[\frac{kab^2(\epsilon - 1)}{3d^2} - \frac{d}{2}\right]. \tag{62}$$

The last term in this formula, in square brackets, is merely a constant; the essential effect of the coupling between the ellipsoids is embodied in the difference from unity of the factor $1 + 2kab^2(\epsilon - 1)/3d^3$ in the first term. Note that this difference is usually quite small. For example, for the extreme case of two spheres in contact, $a = b$ and $d = 2a$, the factor becomes $[1 + (\epsilon - 1)/4(\epsilon + 2)]$ which varies from unity for $\epsilon = 1$ to 1.25 for $\epsilon \to \infty$.

The physical argument above can be extended to higher order in powers of a/d but it is clear that its bookkeeping soon becomes complicated, so we now undertake a more systematic treatment of the problem based on an extension

of the integral equation technique of the earlier parts of this paper. We refer to Fig. 2. When we do not want to make a specific choice of either of the two vectors $\mathbf{r}_1$ or $\mathbf{r}_2$ we shall use $\mathbf{P}$ to refer to the point in space that they jointly define. From the derivation of Eq. (4), it is clear that the integral in it corresponds to the potential produced by the polarized ellipsoid, and hence that the integral representation for two bodies is obtained by adding one such integral for each body, i.e., it is

$$\Phi(\mathbf{P}) = \Phi_{\text{app}}(\mathbf{P}) + (\epsilon - 1)\int_{\text{I}} \frac{\partial \Phi}{\partial n_1'}\bigg|_{-} g(\mathbf{r}_1,\mathbf{r}_1')\, dS_1' + (\epsilon - 1)\int_{\text{II}} \frac{\partial \Phi}{\partial n_2'}\bigg|_{-} g(\mathbf{r}_2,\mathbf{r}_2')\, dS_2' . \tag{63}$$

This equation is valid everywhere, i.e, for $\mathbf{P}$ inside or outside either of the two ellipsoids. When $\mathbf{P}$ is inside ellipsoid #1 it will be convenient to rename $\Phi(\mathbf{P})$ and call it $\Phi_1(r_1)$, i.e.,

$$\Phi(\mathbf{P}) \equiv \Phi_1(\mathbf{r}_1), \quad \mathbf{P} \text{ inside ellipsoid } \#1,$$

and we similarly define $\Phi_{\text{II}}(\mathbf{r}_2)$ for $\mathbf{P}$ inside ellipsoid #2. In this notation, then, Eq. (63) is, inside ellipsoid #1,

$$\Phi_1(\mathbf{r}_1) = \Phi_{\text{app}}(\mathbf{r}_1) + (\epsilon - 1)\int_{S_1} \frac{\partial \Phi_1}{\partial n_1'}\bigg|_{-} g(\mathbf{r}_1,\mathbf{r}_1')\, dS_1' + (\epsilon - 1)\int_{S_2} \frac{\partial \Phi_{\text{II}}}{\partial n_2'}\bigg|_{-} g(\mathbf{r}_2,\mathbf{r}_2')\, dS_2' \tag{64}$$

and there is of course a similar equation for $\Phi_{\text{II}}(\mathbf{r}_2)$. As a first step in using Eq. (64) we wish to express the integral over dS_2', which is a function of $\mathbf{r}_2$, in terms of $\mathbf{r}_1$. Figure 2 shows that if $\mathbf{r}_1$ is inside ellipsoid #1, then we necessarily have $r_2 > r_2'$. We can then use the standard Green function expansion in, for the present case, the usual complex spherical harmonics

$$\begin{aligned} g(\mathbf{r}_2,\mathbf{r}_2') &= -\frac{1}{4\pi}\sum_{s=0}^{\infty}\sum_{t=-s}^{s}\frac{1}{2s+1}\frac{r_2'^s}{r_2^{s+1}} Y_{st}(\Omega_2)Y_{st}^*(\Omega_2') \\ &= -\frac{1}{4\pi}\sum_{s=0}^{\infty}\frac{r_2'^s}{r_2^{s+1}} P_s(\cos\theta_2)P_s(\cos\theta_2') \\ &\quad -\frac{1}{4\pi}\sum_{s=0}^{\infty}\sum_{t\neq 0}\frac{1}{2s+1}\frac{r_2'^s}{r_2^{s+1}} Y_{st}(\Omega_2)Y_{st}^*(\Omega_2'). \end{aligned} \tag{65}$$

The expansion above has been broken into a first sum corresponding to $t = 0$ and a second for $t \neq 0$ since it is clear by virtue of the fact that $\partial\Phi_{\text{II}}/\partial n_2'$ in Eq. (64) is independent of azimuthal angle that only the first sum yields nonzero integrals. To express the second integral in Eq. (64) in terms of r_1, θ_1 we now use Eq. (56) for $P_s(\cos\theta_2)/r_2^{s+1}$. It is clear that for $\mathbf{r}_1$ inside ellipsoid #1 we have $r_1 < d$ so that Eq. (56) is indeed applicable. Then Eq. (64) becomes

$$\Phi_1(\mathbf{r}_1) = \Phi_{\text{app}}(\mathbf{r}_1) + (\epsilon - 1)\int_{S_1} \frac{\partial \Phi_1}{\partial n_1'}\bigg|_{-} \times g(\mathbf{r}_1,\mathbf{r}_1')\, dS_1' - \frac{\epsilon - 1}{4\pi}\sum_{l=0}^{\infty}\sum_{s=0}^{\infty}\frac{(-)^l}{d^{l+s+1}}\frac{(l+s)!}{l!s!} r_1^l P_l(\cos\theta_1)\int_{S_2} \frac{\partial \Phi_{\text{II}}}{\partial n_2'}\bigg|_{-} r_2'^s P_s(\cos\theta_2')\, dS_2' , \tag{66}$$

again with a similar equation for Φ_{II}. Now Φ_1 and Φ_{II} are expanded in powers of $1/d$:

$$\Phi_1 = \sum_{q=0}^{\infty} \Phi_1^{(q)}\left(\frac{1}{d}\right)^q, \quad \Phi_{\text{II}} = \sum_{q=0}^{\infty} \Phi_{\text{II}}^{(q)}\left(\frac{1}{d}\right)^q, \tag{67}$$

and Eq. (66) becomes

$$\sum_{q=0}^{\infty} \Phi_1^{(q)}(\mathbf{r}_1)\left(\frac{1}{d}\right)^q = \Phi_{\text{app}}(\mathbf{r}_1) + (\epsilon - 1)\int_{S_1}\left(\left(\frac{\partial}{\partial n_1'}\sum_{q=0}^{\infty}\Phi_1^{(q)}(\mathbf{r}_1')\right)\bigg|_{-}\left(\frac{1}{d}\right)^q\right) g(\mathbf{r}_1,\mathbf{r}_1')\, dS_1' - \frac{\epsilon - 1}{4\pi}\sum_{q=0}^{\infty}\sum_{l=0}^{\infty}\sum_{s=0}^{\infty}\frac{(-)^l}{d^{l+s+q+1}}\frac{(l+s)!}{l!s!} r_1^l P_l(\cos\theta_1)\int_{S_2} \frac{\partial \Phi_{\text{II}}^{(q)}(\mathbf{r}_2')}{\partial n_2'}\bigg|_{-} r_2'^s P_s(\cos\theta_2')\, dS_2' . \tag{68}$$

The coefficients of various powers of $1/d$ on either side of Eq. (68) are now equated. The equating of the zeroth power yields

$$\Phi_1^{(0)}(\mathbf{r}_1) = \Phi_{\text{app}}(\mathbf{r}_1) + (\epsilon - 1)\int \frac{\partial \Phi_1^{(0)}}{\partial n_1'}\bigg|_{-} g(\mathbf{r}_1,\mathbf{r}_1')\, dS_1' , \tag{69}$$

which has the obvious physical interpretation that it represents ellipsoid #1 in the applied potential Φ_{app} only, as if ellipsoid #2 were not present, i.e., it corresponds to d becoming infinite. The solution to Eq. (69) is then the solution for a single ellipsoid in the Φ_{app} of Eq. (59) or

$$\Phi_1^{(0)}(r_1) = -k\mathscr{E}_0 r_1 P_1(\cos\theta_1) - \frac{\mathscr{E}_0 d}{2} . \tag{70}$$

The equating of coefficents of the first power of $1/d$ in Eq. (68) yields

$$\Phi_1^{(1)}(\mathbf{r}_1) = -\frac{\epsilon - 1}{4\pi}\int_{S_2} \frac{\partial \Phi_{\text{II}}^{(0)}(\mathbf{r}_2')}{\partial n_2'}\bigg|_{-} dS_2' + (\epsilon - 1)\int_{S_1} \frac{\partial \Phi_1^{(1)}(\mathbf{r}_1')}{\partial n_1'}\bigg|_{-} g(\mathbf{r}_1,\mathbf{r}_1')dS_1' . \tag{71}$$

By comparing this equation with Eq. (4) for a single body we see that the first term on the right hand side is an *effective* applied potential that represents to order $1/d$ the effect of the polarized ellipsoid #2 on ellipsoid #1. On physical grounds, however, there can be no such interaction to this order since ellipsoid #2 is uncharged and hence produces a dipole potential which at distance d falls off as $1/d^2$, not as $1/d$. This physical expectation is confirmed in that the first integral of Eq. (71) vanishes. Thus, by the divergence theorem

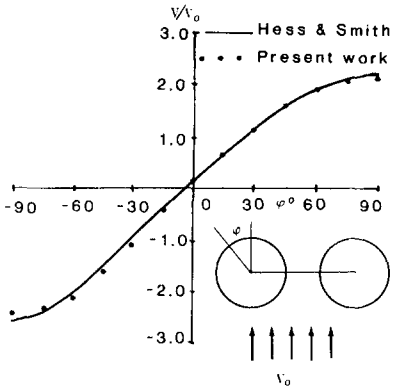

FIG. 3. Fluid speed to order $(a/d)^3$ on the surface of one of two circular cylinders of radius a, center to center distance d, in a stream with uniform velocity v_0 at infinity. $a/d = 1/3$.

$$\int_{S_2} \frac{\partial \Phi_{II}^{(0)}}{\partial n_2'}\bigg|_{-} dS_2' = \int_{V_2} \nabla\cdot\nabla\Phi_{II}^{(0)}\, dV_2' = \int_{V_2} \nabla^2\Phi_{II}^{(0)}\, dV_2' = 0\,. \tag{72}$$

We conclude that $\Phi_I^{(1)}$ is zero, since it represents the response of a polarizable body to a zero applied field, and similarly for $\Phi_{II}^{(1)}$:

$$\Phi_I^{(1)} = \Phi_{II}^{(1)} = 0. \tag{73}$$

The equating to zero of powers of $(1/d)^2$ in Eq. (68) yields three terms from the triple summation; they correspond, respectively, to values 100, 010, and 001 for q,l,s. The resulting equation for $\Phi_I^{(2)}$ can be written as

$$\Phi_I^{(2)}(\mathbf{r}_1) = \Phi_{II,\text{eff}}^{(2)}(\mathbf{r}_1) + (\epsilon - 1) \times \int \frac{\partial \Phi_I^{(2)}(\mathbf{r}_1')}{\partial n_1'}\bigg|_{-} g(\mathbf{r}_1,\mathbf{r}_1')\, dS_1'\,, \tag{74}$$

where $\Phi_{II,\text{eff}}^{(2)}$, the effective applied potential to order $1/d^2$ due to ellipsoid #2, turns out to be merely a constant. Thus, on remembering that $\Phi_{II}^{(1)} = 0$, and having used Eq. (72), we find

$$\Phi_{II,\text{eff}}^{(2)}(\mathbf{r}_2) = -\frac{(\epsilon-1)}{4\pi} \int \frac{\partial \Phi_{II}^{(0)}(\mathbf{r}_2')}{\partial n_2'}\bigg|_{-} r_2' \cos\theta_2'\, dS_2' = k\mathscr{E}_0(\epsilon-1)ab^2/3.$$

This constant effective potential of course corresponds to zero polarizing field; hence, we have the trivial solution for $\Phi_I^{(2)}$:

$$\Phi_I^{(2)} = k\mathscr{E}_0 ab^2(\epsilon-1)/3. \tag{75}$$

The equation for $\Phi_I^{(3)}$ is obtained by equating powers of $1/d^3$ in Eq. (68). This yields integral terms from the triple summation over q,l,s, corresponding to the q,l,s triads 001, 101, 110, 002, 020, 200, but of these six only one leads to nontrivial results. First, the two integrals corresponding to the triads in which $q = 1$ vanish since $\Phi_{II}^{(1)} = 0$. The integral term corresponding to the triad 020 is a constant which is in fact zero by virtue of the oddness of its integrand. The integral corresponding to 020 turns out to have the value zero, and the integral corresponding to 200 is zero by virtue of the fact that its integrand involve the normal derivative of $\Phi_{II}^{(2)}$ and $\Phi_{II}^{(2)}$ is a constant. This leaves only the 011 term and the equation for $\Phi_I^{(3)}$ becomes

$$\Phi_I^{(3)}(\mathbf{r}_1) = -\Gamma r_1 P_1(\cos\theta_1) + (\epsilon-1) \int \frac{\partial \Phi_I^{(3)}(\mathbf{r}_1')}{\partial n_1'}\bigg|_{-} \times g(\mathbf{r}_1,\mathbf{r}_1')\, dS_1'\,, \tag{76}$$

where

$$\Gamma = \frac{(\epsilon-1)}{2\pi} \int \frac{\partial \Phi_{II}^{(0)}(\mathbf{r}_2')}{\partial n_2'}\bigg|_{-} r_2' P_1(\cos\theta_2')\, dS_2' = 2kab^2(\epsilon-1)\mathscr{E}_0/3. \tag{77}$$

Equation (76) represents the problem of an ellipsoid in the uniform applied field of magnitude Γ, hence, the solution is

$$\Phi_I^{(3)} = -k\Gamma r_1\cos\theta_1 = -2k^2ab^2(\epsilon-1)\mathscr{E}_0 r_1\cos\theta_1/3. \tag{78}$$

If now the various pieces (70), (73), (75), and (78) are put together, we find the solution (61) that was derived in physical grounds.

The virtue of the present treatment is that one can now continue systematically to find $\Phi_I^{(4)}$ and $\Phi_{II}^{(4)}$ and higher order terms. Thus, to find $\Phi_I^{(4)}$ one gets ten terms from the triple summation in Eq. (68), corresponding to the various assignments of q,l,s to the number triples 003, 012, 111. Of these ten terms, many vanish for reasons similar to those discussed above and those that remain correspond to the problem of ellipsoid #1 in a potential that is a linear sum of $r_1P_1(\cos\theta_1)$, and $r_1^2P_2(\cos\theta_1)$, which can be solved by the method of the previous sections. The same procedure can of course be carried out to higher order in a/d. In this connection it should be noted that the calculation we have presented above is relatively simple in that we have not had to consider in any detail the essential coupled nature of the equations relating $\Phi_I^{(n)}$ and $\Phi_{II}^{(n)}$, i.e., to order $(a/d)^3$ above in which $\Phi_I^{(3)}$ was calculated, we needed to know only $\Phi_{II}^{(0)}$, $\Phi_{II}^{(1)}$, and $\Phi_{II}^{(2)}$ and we were able to calculate these trivially. In general, i.e, to higher orders, or for dissimilar bodies, one must work with the equations for $\Phi_I^{(n)}$ and $\Phi_{II}^{(n)}$ simulaneously.

An important question is, of course, that of the convergence of the expansion in powers of a/d. To throw some small light on this we have worked out in the manner above a problem of fluid flow around two cylinders. The geometry is shown in Fig. 3. A quantity of special interest in this kind of problem is the fluid speed on the surface and we present the results in terms of this. With only one cylinder this speed v is $2v_0\sin\varphi$, where v_0 is the speed of the uniform stream at infinity. For the two cylinders, to order $(a/d)^3$, the speed on the surface of one of the cylinders is

$$v = \frac{2v_0}{1-(a/d)^3}\left[\sin\varphi + 2(a/d)^3\cos 2\varphi\right]. \tag{79}$$

For $a/d = 1/3$, one can compare these results with those of Hess and Smith[17] who have solved the problem both numerically and analytically with excellent agreement between their two methods. From Fig. 3 we see that the simple formula (79) agrees with Hess and Smith's to within a couple of percent over most of the range of φ, and differs by at most about 5% at $\varphi = 90°$.

[1]L. Eyges, Ann. Phys. (New York) **90**, 266 (1975).
[2]E.T. Copson, Proc. Edinburgh Math. Soc. **8**, 14 (1947).
[3]I. Sneddon, *Mixed Boundary Value Problems in Potential Theory* (North Holland, Amsterdam, 1966).
[4]Note that the Green function here is defined with an extra factor $-1/4\pi$ as compared with that in Ref. 1.
[5]P.C. Waterman, J. Acoust. Soc. Am. **6**, 1417 (1969).
[6]J. Van Bladel, *Electromagnetic Fields* (McGraw Hill, New York, 1964), p. 140.
[7]P. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw Hill, New York, 1953), p. 805.
[8]J. Stratton, *Electromagnetic Theory* (McGraw Hill, New York, 1941).
[9]This fact is also pointed out in Ref. 8, p. 213.
[10]L.M. Milne Thompson, *Theoretical Hydrodynamics* (Macmillan, New York, 1968) Fifth Edition. The relevant equation is in Sec. 6.33: Application of elliptic coordinates to the streaming past an ellipse.
[11]G.K. Batchelor, *An Introduction to Fluid Dynamics* (Cambridge University, Cambridge, England, 1967). The relevant equation is Eq. (6.6.21).
[12]The eigenvalue spectrum of this equation is important for problems of infrared absorption and has been studied in that context in Ref. 13.
[13]D. Langbein, J. Phys. A **9**, 627 (1976).
[14]L. Eyges, *The Classical Electromagnetic Field* (Addison-Wesley, Reading, Mass., 1972) [reprinted by Dover, New York, 1980], p. 51.
[15]A similar-technique has been used by Fikioris to solve this problem without the use of ellipsoidal coordinates: J. Fikioris, Q. Appl. Math. **32**, 143 (1973).
[16]E.O. Steinborn and K. Ruedenberg, Adv. Quantum Chem. **7**, 1 (1973).
[17]J.L. Hess and A.M.O. Smith, *Advances in Aeronautical Sciences* (Pergamon, New York, 1967), Vol. 8, pp. 1–138.

Resonance of the axial-symmetric modes in microstrip disk resonators[a)]

W. C. Chew and J. A. Kong

Department of Electrical Engineering and Computer Science and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 17 August 1979; accepted for publication 2 November 1979)

Three methods have been used to calculate the resonant frequencies for the axial-symmetric modes in the microstrip disk resonators which also have the application as radiating antennas. The variational Galerkin's method provides the most accurate result for the estimation of both the real and imaginary parts of the resonant frequency shifts with judicious choices of basis functions. The iterative approach gives good results for the imaginary part of the frequency shift and very crude approximations for the real part. The perturbation formula is the most efficient in computational time and is useful for sufficiently large permittivity for the substrate material and small ratio of substrate thickness to disk radius. To illustrate the application of the circular microstrip disk as a radiating antenna, the radiation pattern for the axial-symmetric mode is also plotted.

I. INTRODUCTION

The advent of microwave integrated circuits has resulted in the increased interest in using microstrip disks as resonators or antennas. Watkins[1] obtained the approximate resonant frequencies for a circular microstrip disk by using a magnetic-wall-resonator model and ignoring the fringing field effect. Several authors have attempted to obtain corrections to Watkins' result through a quasistatic and qualitative argument.[2-5] However, at resonance, the disk radius is of the same order of magnitude as the wavelength and it is doubtful that the quasistatic argument is legitimate. Itoh[6] performed a full-wave analysis to find the resonant frequencies of a rectangular microstrip disk but the disk was enclosed in a waveguide resulting in no radiation loss. As such, the resonant frequencies computed are pure real.

In this paper, we shall formulate the mixed-boundary value problem for a circular disk placed on top of a general stratified half-space using Hankel transform methods. For the axial-symmetric modes, the problem reduced to the solution of dual integral equations which have been of historic interest in solving potential problems with axial symmetry.[7-10] We shall solve for the resonant frequencies when the stratified half-space corresponds to a substrate with a ground plane for microwave integrated circuits. We shall also derive a perturbation formula for the resonant frequency shift for both real and imaginary parts.

II. FORMULATION

For a circular disk in an unbounded medium carrying an arbitrary distribution of current, the $\hat{z}$ components of the electric and magnetic fields are given by

$$E_z = \pm \sum_n \int_0^\infty k_\rho \left[e_n^s(k_\rho) \sin n\phi + e_n^c(k_\rho) \cos n\phi \right] \times e^{\pm ik_z z} J_n(k_\rho \rho)\, dk_\rho, \tag{1a}$$

$$H_z = \sum_n \int_0^\infty k_\rho \left[h_n^s(k_\rho) \sin n\phi + h_n^c(k_\rho) \cos n\phi \right] \times e^{\pm ik_z z} J_n(k_\rho \rho)\, dk_\rho, \tag{1b}$$

where $k_z = (k^2 - k_\rho^2)^{1/2}$, $k^2 = \omega^2 \mu\epsilon$, the upper sign in the above is chosen for $z > 0$, the lower sign for $z < 0$, and $e_n^s(k_\rho)$, $e_n^c(k_\rho)$, $h_n^s(k_\rho)$, and $h_n^c(k_\rho)$ are the sine and cosine components of the TM and TE field components in the k_ρ-space. The discontinuity in E_z is necessary to account for the charge accumulation on the disk, whereas H_z is continuous across the plane $z = 0$. In the presence of a stratified half-space (Fig. 2) characterized by reflection coefficients R^{TM} and R^{TE} for TM and TE waves respectively, the integral summations of cylindrical waves as denoted by (1a) and (1b) are reflected off the half-space. Hence the $\hat{z}$ components of the fields in the upper half-space are

$$E_z = \sum_n \int_0^\infty k_\rho \left[e_n^s(k_\rho) \sin n\phi \right.$$

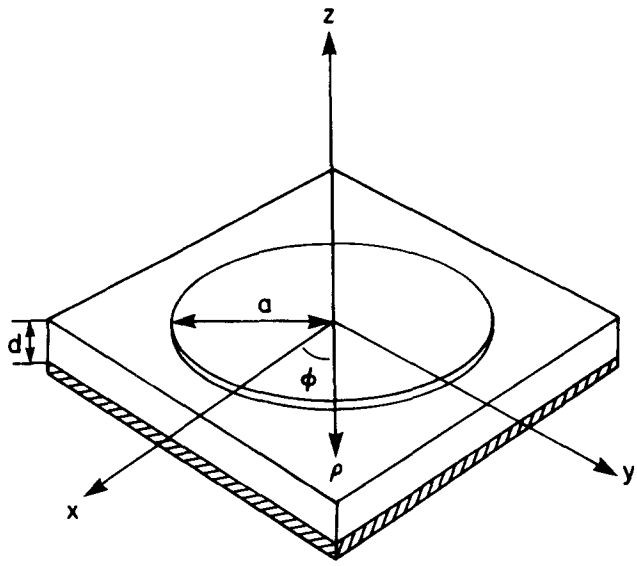

FIG. 1. Geometrical configuration of the problem.

[a)]This work was supported by the Schlumberger-Doll Research Center and by the Joint Services Electronics Program under Contract DAAG-29-78-C-0020.

 0022-2488/80/030582-10$1.00

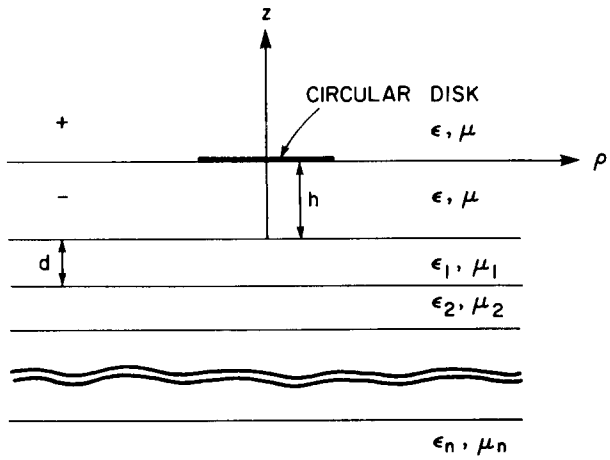

FIG. 2. A circular disk on a stratified medium.

$$+ e_n^c(k_\rho)\cos n\phi\,][\pm e^{\pm ik_z z} - R^{\mathrm{TM}}\exp(ik_z z + 2ik_z h)]\,J_n(k_\rho\rho)\,dk_\rho, \tag{2a}$$

$$H_z = \sum_n \int_0^\infty k_\rho\,[h_n^s(k_\rho)\sin n\phi + h_n^c(k_\rho)\cos n\phi\,] \times [e^{\pm ik_z z} + R^{\mathrm{TE}}\exp(ik_z z + 2ik_z h)]J_n(k_\rho\rho)\,dk_\rho. \tag{2b}$$

The fields transverse to the z axis are given by Ref. 11, p. 215.

$$\begin{aligned}\mathbf{E}_s = {}&\hat\rho\sum_n\int_0^\infty\Big(\,[e_n^s(k_\rho)\sin n\phi + e_n^c(k_\rho)\cos n\phi\,)] \\ &\times ik_z\,[e^{\pm ik_z z} - R^{\mathrm{TM}}\exp(ik_z z + i2k_z h)]J_n'(k_\rho\rho) \\ &+ \frac{i\omega\mu n}{k_\rho\rho}\,[h_n^s(k_\rho)\cos n\phi - h_n^c(k_\rho)\sin n\phi\,] \\ &\times [e^{\pm ik_z z} + R^{\mathrm{TE}}\exp(ik_z z + i2k_z h)]\,J_n(k_\rho\rho)\Big)\,dk_\rho \\ &+ \hat\phi\sum_n\int_0^\infty\Big(\,[e_n^s(k_\rho)\cos n\phi - e_n^c(k_\rho)\sin n\phi\,] \\ &\times \frac{ik_z n}{k_\rho\rho}\,[e^{\pm ik_z z} - R^{\mathrm{TM}}\exp(ik_z z + i2k_z h)]\,J_n(k_\rho\rho) \\ &\times i\omega\mu\,[h_n^s(k_\rho)\sin n\phi + h_n^c(k_\rho)\cos n\phi\,] \\ &\times [e^{\pm ik_z z} + R^{\mathrm{TE}}\exp(ik_z z + i2k_z h)]\,J_n'(k_\rho\rho)\Big)\,dk_\rho,\end{aligned} \tag{3}$$

$$\begin{aligned}\mathbf{H}_s = {}&\hat\rho\sum_n\int_0^\infty\Big(\,[h_n^s(k_\rho)\sin n\phi + h_n^c(k_\rho)\cos n\phi\,]\,ik_z \\ &\times [\pm e^{\pm ik_z z} + R^{\mathrm{TE}}\exp(ik_z z + i2k_z h)]J_n'(k_\rho\rho) \\ &- \frac{i\omega\epsilon n}{k_\rho\rho}\,[e_n^s(k_\rho)\cos n\phi - e_n^c(k_\rho)\sin n\phi\,] \\ &\times [\pm e^{\pm ik_z z} - R^{\mathrm{TM}}\exp(ik_z z + i2k_z h)]\,J_n(k_\rho\rho)\Big) \\ &\times dk_\rho + \hat\phi\sum_n\int_0^\infty\Big(\,[h_n^s(k_\rho)\cos n\phi - h_n^c(k_\rho)\sin n\phi\,] \\ &\times \frac{ik_z n}{k_\rho\rho}[\pm e^{\pm ik_z z} + R^{\mathrm{TE}}\exp(ik_z z + i2k_z h)] \\ &\times J_n(k_\rho\rho) + i\omega\epsilon\,[e_n^s(k_\rho)\sin n\phi + e_n^c(k_\rho)\cos n\phi\,] \\ &\times [\pm e^{\pm ik_z z} - R^{\mathrm{TM}}\exp(ik_z z + 2ik_z h)]\,J_n'(k_\rho\rho)\Big)\,dk_\rho.\end{aligned} \tag{4}$$

For the natural modes of the disk, the fields also satisfy the boundary condition on the disk, that is

$$\mathbf{E}_s(z=0) = 0 \quad \text{for } \rho < a. \tag{5}$$

For different n's, the above boundary condition is satisfied independently for $\sin n\phi$ or $\cos n\phi$ components. Without loss of generality, we can just consider the boundary condition (5) as

$$\int_0^\infty e_n(k_\rho)ik_z\,[1 - R^{\mathrm{TM}}e^{i2k_z h}]\,J_n'(k_\rho\rho)\,dk_\rho + \frac{i\omega\mu n}{\rho}\int_0^\infty\frac{h_n(k_\rho)}{k_\rho}\,[1 + R^{\mathrm{TE}}e^{i2k_z h}]\,J_n(k_\rho\rho)\,dk_\rho = 0, \quad \rho < a, \tag{6a}$$

$$\frac{n}{\rho}\int_0^\infty e_n(k_\rho)\frac{ik_z}{k_\rho}\,[1 - R^{\mathrm{TM}}e^{i2k_z h}]\,J_n(k_\rho\rho)\,dk_\rho + i\omega\mu\int_0^\infty h_n(k_\rho)[1 + R^{\mathrm{TE}}e^{i2k_z h}]\,J_n'(k_\rho\rho)\,dk_\rho = 0, \quad \rho < a, \tag{6b}$$

by leaving out the superscripts s and c. The other set of boundary conditions requires that $\mathbf{H}_s$ be continuous across the plane $z = 0$ for $\rho > a$, which results in

$$\frac{K_\phi}{\sin n\phi} = 2\int_0^\infty h_n(k_\rho)ik_z\,J_n'(k_\rho\rho)\,dk_\rho + \frac{2i\omega\epsilon n}{\rho}\int_0^\infty\frac{e_n(k_\rho)}{k_\rho}J_n(k_\rho\rho)\,dk_\rho = 0, \quad \rho > a, \tag{7a}$$

$$\frac{K_\rho}{\cos n\phi} = -\frac{2n}{\rho}\int_0^\infty h_n(k_\rho)\frac{ik_z}{k_\rho}J_n(k_\rho\rho)\,dk_\rho - 2i\omega\epsilon\int_0^\infty e_n(k_\rho)J_n'(k_\rho\rho)\,dk_\rho = 0, \quad \rho \geqslant a, \tag{7b}$$

where $\mathbf{K}$ denotes the current sheet distribution on the disk. We observe that in general, the TE and TM modes of a microstrip disk resonator are coupled and thus hybrid modes exist. In Eq. (7b), we require that it also be satisfied at $\rho = a$ requiring $H_\phi(\rho = a) = 0$ or the $\hat\rho$ component of the surface current $K_\rho(\rho = a) = 0$.

For the axial-symmetric modes, $n = 0$, from which we note that the TE and TM waves are decoupled. When $d \to 0$, the TE modes are evanescent and we shall look at the TM modes. Using the fact that $J_0'(x) = -J_1(x)$, we have the following equations,

$$\int_0^\infty k_\rho\,\hat K_\rho(k_\rho)G^{\mathrm{TM}}(k_\rho)\,J_1(k_\rho\rho)\,dk_\rho = 0, \quad \rho < a, \tag{8a}$$

$$\int_0^\infty k_\rho\,\hat K_\rho(k_\rho)\,J_1(k_\rho\rho)\,dk_\rho = 0, \quad \rho \geqslant a, \tag{8b}$$

where

$$\hat K_\rho(k_\rho) = +2i\omega\epsilon e(k_\rho)/k_\rho, \qquad G^{\mathrm{TM}}(k_\rho) = k_z\,[1 - R^{\mathrm{TM}}e^{i2k_z h}]\,. \tag{8c}$$

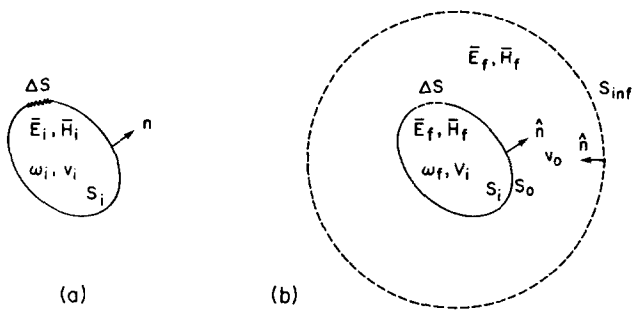

FIG. 3. Derivation of a perturbation formula. (a) Geometry before perturbation; (b) Geometry after perturbation.

Note that $\hat{K}_\rho(k_\rho)$ is the Hankel transform of $K_\rho(\rho)$, the current distribution on the disk.

For microstrip application, the disk is placed at the half-space interface and thus $h = 0$. Also, medium 2 is a perfect conductor. Using the reflection coefficient for a stratified medium,[11] it can be shown that

$$G^{\mathrm{TM}}(k_\rho) = \frac{2k_z \sin k_{1z} d}{\sin k_{1z} d + i\epsilon_r (k_z/k_{1z}) \cos k_{1z} d}, \tag{9}$$

where $\epsilon_r = \epsilon_1/\epsilon$ is the relative permittivity of the substrate.

III. ITERATIVE APPROACH

We can gain qualitative understanding of the solutions of (8) if we let $d \to 0$. In that limit,

$$G^{\mathrm{TM}}(k_\rho) \sim \frac{2d}{i\epsilon_r}(k_1^2 - k_\rho^2) + O(d^2),$$

and Eq. (8a) is approximately

$$\int_0^\infty k_\rho \hat{K}_\rho(k_\rho)(k_1^2 - k_\rho^2) J_1(k_\rho \rho)\, dk_\rho = 0, \quad \rho < a, \tag{10}$$

which can be rewritten as

$$\left(\frac{1}{\rho}\frac{\partial}{\partial \rho}\rho\frac{\partial}{\partial \rho} + k_1 - \frac{1}{\rho^2}\right) K_\rho(\rho) = 0, \quad \rho < a, \tag{10a}$$

where

$$K_\rho(\rho) = \int_0^\infty k_\rho \hat{K}_\rho(k_\rho) J_1(k_\rho \rho)\, dk_\rho. \tag{10b}$$

Equation (10a) gives rise to the solution $K_\rho(\rho) = AJ_1(k_1\rho)$. In order for it to satisfy (8b), we arrived at the eigenequation $J_1(k_1 a) = 0$ from which we can determine the resonant frequencies. The result is the same as the magnetic-wall model,[1] which is seen to be valid only when $d = 0$.

Inspired by (10) and (10a), we can rewrite (8a) as

$$\left(\frac{1}{\rho}\frac{\partial}{\partial \rho}\rho\frac{\partial}{\partial \rho} + k_1^2 - \frac{1}{\rho^2}\right)\int_0^\infty k_\rho \hat{K}_\rho(k_\rho) \times \frac{i\epsilon_r G^{\mathrm{TM}}(k_\rho)}{2d(k_1^2 - k_\rho^2)} J_1(k_\rho \rho)\, dk_\rho = 0, \quad \rho < a, \tag{8a'}$$

which is equivalent to

$$\int_0^\infty k_\rho \hat{K}_\rho(k_\rho) \frac{i\epsilon_r G^{\mathrm{TM}}(k_\rho)}{2d(k_1^2 - k_\rho^2)} J_1(k_\rho \rho)\, dk_\rho = AJ_1(k_1 \rho), \quad \rho < a. \tag{11}$$

The above equation can be solved iteratively since we notice that

$$\frac{i\epsilon_r G^{\mathrm{TM}}(k_\rho)}{2d(k_1^2 - k_\rho^2)} \sim 1 + O(d), \quad d \to 0.$$

We rewrite (11) in the form

$$K_\rho(\rho) = \int_0^\infty k_\rho \hat{K}_\rho(k_\rho) J_1(k_\rho \rho)\, dk_\rho = AJ_1(k_1 \rho) + \int_0^\infty k_\rho \hat{K}_\rho(k_\rho)\left(1 - \frac{i\epsilon_r G^{\mathrm{TM}}(k_\rho)}{2d(k_1^2 - k_\rho^2)}\right) \times J_1(k_\rho \rho)\, dk_\rho, \quad \rho < a. \tag{11a}$$

When $d \to 0$, the integral term on the right-hand side of Eq. (11a) is small. The first order solution can be obtained by substituting in the zeroth order $\hat{K}_\rho^{(0)}(k_\rho)$ into the integral and evaluating numerically. $\hat{K}_\rho^{(0)}(k_\rho)$ can be found explicitly (Ref. 12, 11.3.29),

$$\hat{K}^{(0)}(k_\rho) = \frac{Aa}{k_1^2 - k_\rho^2}[k_1 J_2(k_1 a) J_1(k_\rho a) - k_\rho J_1(k_1 a) J_2(k_\rho a)]. \tag{12}$$

When we require that $K_\rho^{(1)}(\rho = a) = 0$, we obtain an eigenequation from which we can find the resonant frequencies of the microstrip disk when $d > 0$.

IV. GALERKIN'S METHOD

Since we know that when $d \to 0$, the dual integral equations give a solution similar to the magnetic-wall model, we shall use the current-modes of the magnetic-wall model as basis functions. Consequently, we have

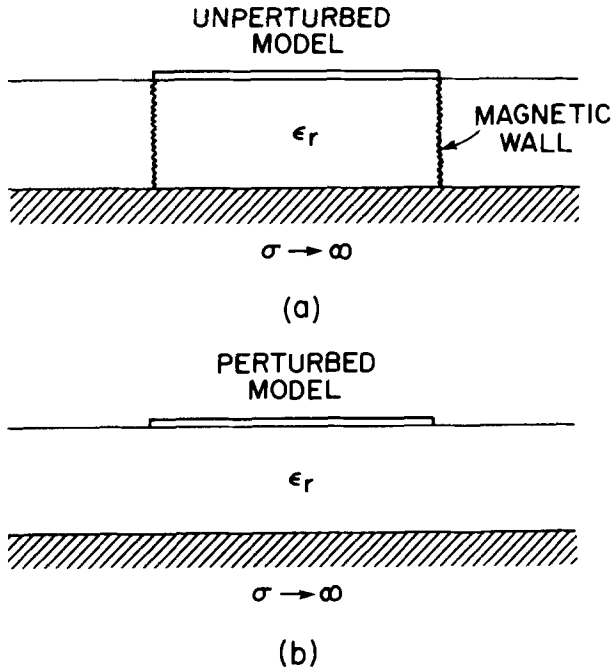

FIG. 4. Microstrip disk resonator. (a) Before perturbation; (b) After perturbation.

$$K_\rho(\rho) = \begin{cases} \sum_{m=1}^{K} b_m J_1\left(\frac{\alpha_{1m}}{a}\rho\right), & \rho < a, \\ 0, & \rho > a, \end{cases} \tag{13}$$

where $J_1(\alpha_{1m}) = 0$. As such, Eq. (8b) is automatically satisfied. It follows from (13) that

$$\hat{K}_\rho(k_\rho) = -\sum_{m=1}^{K} b_m \frac{a^2 \alpha_{1m} J_0(\alpha_{1m}) J_1(k_\rho a)}{\alpha_{1m}^2 - k_\rho^2 a^2}. \tag{14}$$

Substituting in (8a), we have

$$\sum_{m=1}^{K} -b_m a^2 \alpha_{1m} J_0(\alpha_{1m}) \times \int_0^\infty \frac{k_\rho J_1(k_\rho a)}{(\alpha_{1m}^2 - k_\rho^2 a^2)} G^{\mathrm{TM}}(k_\rho) J_1(k_\rho \rho)\, dk_\rho = 0. \tag{15}$$

We can eliminate the ρ dependence by multiplying the above equation by $\rho J_1[(\alpha_{1p}/a)\rho]$, $p = 1,2,...,K$ and integrate with respect to ρ from 0 to a. Using Parserval's theorem for Hankel transforms, we have

$$\sum_{m=1}^{K} b_m G_{mp} = 0, \quad p = 1,2,...K, \tag{16}$$

where

$$G_{mp} = a^4 \alpha_{1m} \alpha_{1p} J_0(\alpha_{1m}) J_0(\alpha_{1p}) \times \int_0^\infty k_\rho \frac{J_1^2(k_\rho a) G^{\mathrm{TM}}(k_\rho)}{(\alpha_{1m}^2 - k_\rho^2 a^2)(\alpha_{1p}^2 - k_p^2 a^2)} dk_\rho. \tag{16a}$$

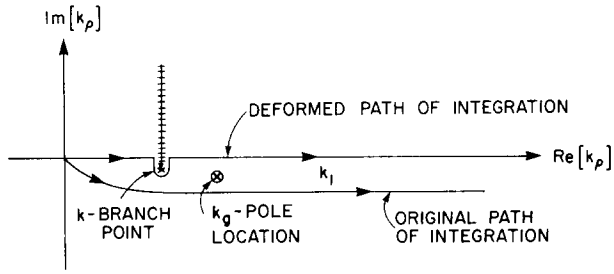

FIG. 5. Paths of integration on the k_ρ plane.

Equation (16) corresponds to a matrix equation $\mathbf{G}\cdot\mathbf{b} = 0$ where $\mathbf{G}$ is a $K \times K$ matrix and $\mathbf{b}$ is a K-element column vector. In order for (13) to be nontrivial, $\bar{b}$ has to be nontrivial. As a consequent, we arrive at our eigenequation

$$\det|\mathbf{G}| = 0, \tag{17}$$

from which we can solve for the resonant frequencies. Since the set of basis functions we have chosen in (13) is complete, our solution is exact if we let $K \to \infty$. For practical considerations, K is finite. With a judicious choice of the basis functions, usually, one or two basis functions, usually, one or two basis functions give sufficiently good results.

V. A PERTURBATION FORMULA

It is advantageous to view the microstrip disk resonator as a perturbation of the magnetic-wall model. Thus we shall derive a perturbation formula for the frequency shift from which we can see the mechanisms that cause the resonant frequencies of a disk resonator to differ from that of the magnetic-wall model.

First, we consider a resonator that consists partly of electric wall and partly of magnetic wall as shown in Fig. 3a. If the magnetic wall constitutes a small portion of the cavity wall, the current distribution on the electric wall and the fields inside the cavity do not change drastically if the magnetic wall is removed. We denote the initial $\mathbf{E}$ and $\mathbf{H}$ fields and resonant frequency by $\mathbf{E}_i$, $\mathbf{H}_i$, and ω_i, respectively, and the final $\mathbf{E}$ and $\mathbf{H}$ fields and resonant frequency by $\mathbf{E}_f$, $\mathbf{H}_f$, and ω_f respectively. It can be shown easily that

$$\nabla\cdot(\mathbf{E}_i^* \times \mathbf{H}_f + \mathbf{E}_f \times H_i^*) = i(\omega_f - \omega_i)(\mu \mathbf{H}_i^* \cdot \mathbf{H}_f + \epsilon \mathbf{E}_i^* \cdot \mathbf{E}_f). \tag{18}$$

Integrating the above over volume V_i, we obtain

$$\iint_{S_i + \Delta S} (\mathbf{E}_i^* \times \mathbf{H}_f + \mathbf{E}_f \times \mathbf{H}_i^*)\cdot\hat{n}\, ds = i(\omega_f - \omega_i) \iiint_{V_i} (\mu \mathbf{H}_i^* \cdot \mathbf{H}_f + \epsilon \mathbf{E}_i^* \cdot \mathbf{E}_f)\, dV. \tag{19}$$

The first term of the integrand on the left-hand side vanishes except over ΔS. If we consider the cavity wall on the left side to have a small wall loss, the second term of the integrand on the left-hand side is nonzero except over S_i. For this reason, we have

$$\omega_f - \omega_i = -i \frac{\iint_{\Delta S} (\mathbf{E}_i^* \times \mathbf{H}_f)\cdot\hat{n}\, dS + \iint_{S_i} (\mathbf{E}_f \times \mathbf{H}_i^*)\cdot\hat{n}\, dS}{\iiint_{V_i} (\mu \mathbf{H}_i^* \cdot \mathbf{H}_f + \epsilon \mathbf{E}_i^* \cdot \mathbf{E}_f)\, dV}. \tag{20}$$

Since $\mathbf{E}_i$ does not differ appreciably from $\mathbf{E}_f$ on ΔS, and $\mathbf{E}_f \simeq \sqrt{(\omega\mu/i\sigma)}\hat{n} \times \mathbf{H}_i$ on S_i, and $\mathbf{H}_f$ and $\mathbf{E}_f$ are approximately equal to $\mathbf{H}_i$ and $\mathbf{E}_i$ in v_i, we obtain, through appropriate approximation, that

$$\omega_f - \omega_i \simeq -i \frac{\iint_{\Delta S} (\mathbf{E}_i^* \times \mathbf{H}_f)\cdot\hat{n}\, dS + \sqrt{(\omega\mu/i\sigma)} \iint_{S_i} |\mathbf{H}_i|^2\, dS}{\iiint_{V_i} (\mu |\mathbf{H}_i|^2 + \epsilon |\mathbf{E}_i|^2)\, dV}. \tag{21}$$

Applying Poynting's theorem to V_0 outside the cavity, we obtain

$$\iint_{\Delta S} (\mathbf{E}_f \times \mathbf{H}_f^*)\cdot\hat{n}\, dS + \iint_{S_0} (\mathbf{E}_f \times \mathbf{H}_f^*)\cdot\hat{n}\, dS + \iint_{S_{\mathrm{inf}}} (\mathbf{E}_f \times \mathbf{H}_f^*)\cdot\hat{n}\, dS = -i\omega_f \iiint_{V_0} [\mu |\mathbf{H}_f|^2 - \epsilon |\mathbf{E}_f|^2]\, dV. \tag{22}$$

The second term on the left-hand side corresponds to wall loss outside the cavity, which is usually small. The third term on the

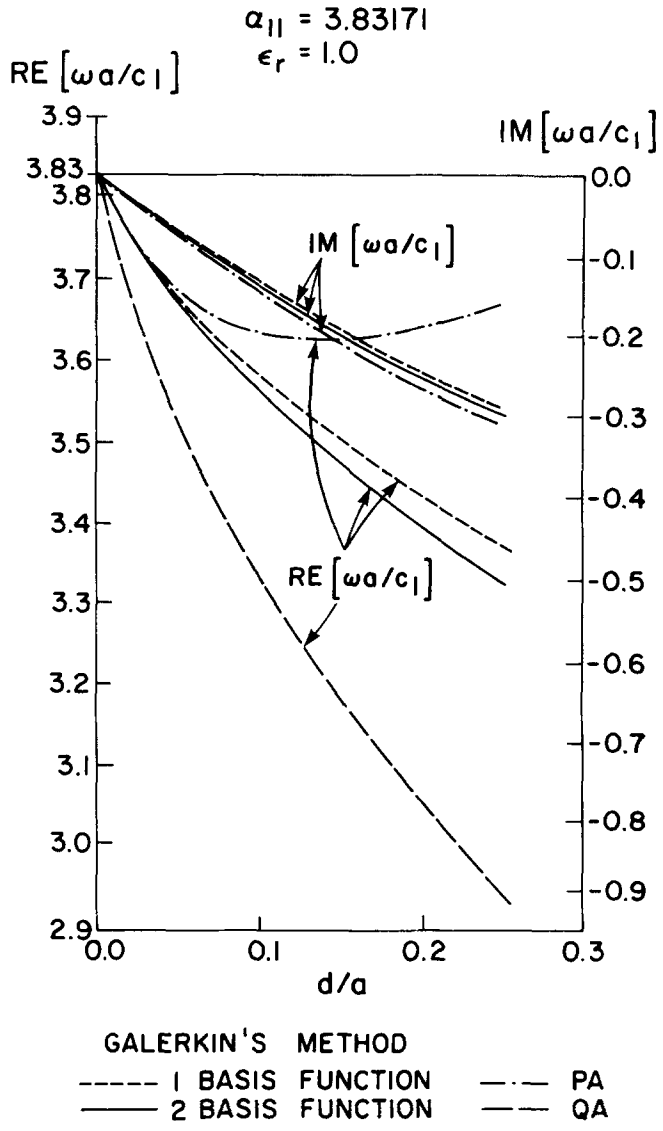

FIG. 6. Resonant frequency shift for the lowest axial symmetric TM mode when $\epsilon_r = 1.0$ (PA-perturbation approach, QA-quasistatic approach).

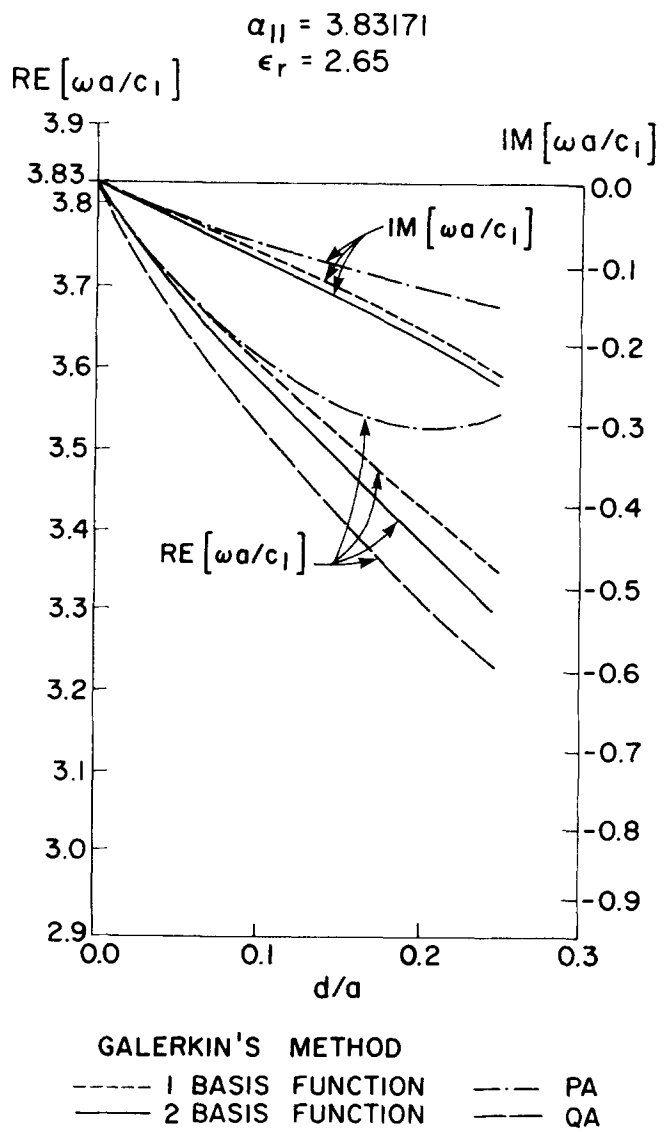

FIG. 7. Resonant frequency shift for the lowest axial symmetric TM mode when $\epsilon_r = 2.65$ (PA-perturbation approach, QA-quasistatic approach).

left-hand side does not vanish when S_{inf} recedes to infinity; it is the negative of twice the time average power radiated by the cavity. Thus (21) can be written as

$$\omega_f - \omega_i \simeq \frac{-2i\langle P_r\rangle + \omega_f \iiint_{V_0}(\mu|\mathbf{H}_f|^2 - \epsilon|\mathbf{E}_f|^2)\,dV - i\sqrt{\omega\mu/2\sigma}\iint_{S_i}|\mathbf{H}_i|^2\,dS - \sqrt{\omega\mu/2\sigma}\iint_{S_i}|\mathbf{H}_i|^2\,dS}{\iiint_{V_i}(\mu|\mathbf{H}_i|^2 + \epsilon|\mathbf{E}_i|^2)\,dV}. \tag{23}$$

From the above, we notice that the imaginary part of the frequency shift is proportional to the power radiated by the open cavity and the dissipative loss on the wall. The real part of the frequency shift is proportional to the reactive power leaking out of the cavity and the negative of the reactive power absorbed by the cavity wall.

We can obtain a zeroth order appraisal of the wall loss in the microstrip disk resonator by assuming the field inside to be that of the magnetic-wall model. For the axial-symmetric mode, the field inside the cavity is given by[1,13]

$$E_z = E_0 J_0(k_1\rho), \tag{24a}$$

$$H_\phi = -i(\epsilon_1/\mu_1)^{1/2}E_0 J_1(k_1\rho), \tag{24b}$$

where

$$k_1 a = \alpha_{1m}. \tag{24c}$$

The complex Poynting power absorbed by the top and bottom wall of the resonator is

$$P_W = 2(\omega\mu_1/i\sigma)^{1/2}\int_0^a\int_0^{2\pi}|\mathbf{H}_\phi|^2\rho\,d\rho\,d\phi. \tag{25}$$

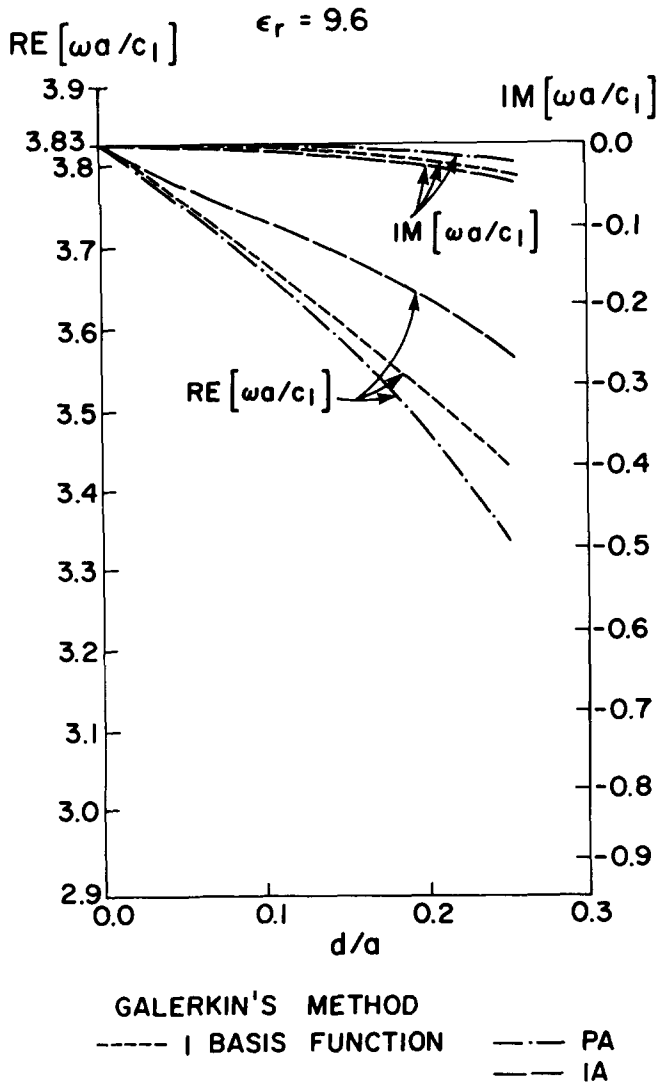

FIG. 8. Resonant frequency shift for the lowest axial symmetric TM mode when $\epsilon_r = 9.6$ (PA-perturbation approach, IA-iterative approach).

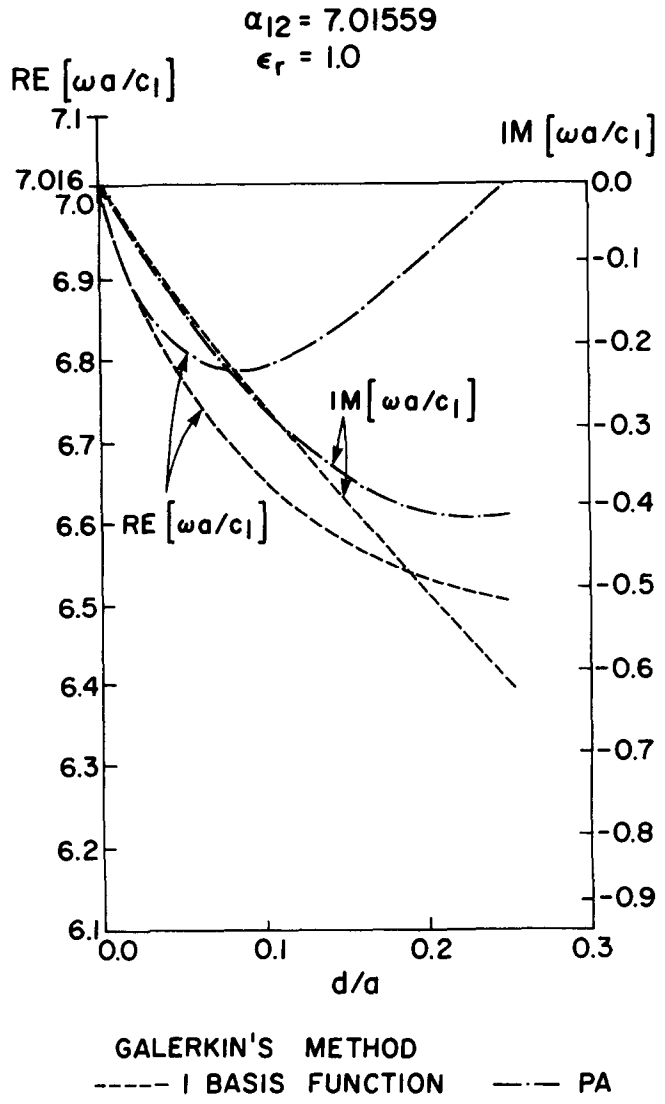

FIG. 9. Resonant frequency shift for the lowest axial symmetric TM mode when $\epsilon_r = 1.0$ (PA-perturbation approach).

The frequency shift due to wall loss from (23) is

$$\Delta\omega_W = (1/d)(\omega/i\sigma\mu_1)^{1/2} = (1/d)(\omega/2\sigma\mu_1)^{1/2}(1-i). \tag{26}$$

We see that the wall loss can cause a significant shift in frequency when $d \to 0$, but for d in the range of interest, $\Delta\omega_W$ is usually small.

If the dielectric is lossy, we obtain from (24c) that

$$\omega = \frac{\alpha_{1m}}{a[\mu_1(\epsilon_1' + i\epsilon_1'')]^{1/2}} \simeq \frac{\alpha_{1m}}{a(\mu_1\epsilon_1)^{1/2}}\left(1 - i\frac{\epsilon_1''}{2\epsilon_1'}\right). \tag{27a}$$

Therefore,

$$\Delta\omega_D \simeq -i(\omega/2)\tan\delta, \tag{27b}$$

where $\tan\delta$ is the loss tangent of the dielectric substrate. This loss tangent is usually very small ($\sim 10^{-3}$), implying that this shift is usually small.

We observed from (23) the mechanism for frequency shift of a microstrip resonator. The frequency shift due to the first two terms of the numerator, which can be interpreted as due to complex power leaking out of the resonator, can be estimated accurately by the dual integral equation approach. We have determined the frequency shift due to wall loss to zeroth order. It is of interest to note that the radiated power can be determined to zeroth order if we assume a zeroth-order current distribution on the disk. From Sec. III, we note that the zeroth-order current distribution is the same as that of the magnetic-wall model current distribution. Therefore,

$$K_\rho^{(0)}(\rho) = AJ_1(\alpha_{1m}\rho/a), \tag{28a}$$

where

$$\hat{K}_\rho^{(0)}(k_\rho) = -A\frac{\alpha_{1m}\,J_0(\alpha_{1m})\,J_1(k_\rho a)}{\alpha_{1m}^2/a^2 - k_\rho^2}. \tag{28b}$$

Making use of (8c) and (2a), we obtain the field in the upper half-space of the microstrip disk as

$$E_z^{(0)} = \frac{A}{2i\omega\epsilon}\alpha_{1m}\,J_0(\alpha_{1m})\int_0^\infty \frac{k_\rho^2\,J_1(k_\rho a)}{[(\alpha_{1m}^2/a)^2 - k_\rho^2]}(1 - R^{\text{TM}})\,e^{ik_z z}\,J_0(k_\rho\,\rho)\,dK_\rho\,. \tag{29}$$

When z and ρ are large, the above integrand is rapidly oscillating, and can be evaluated with the method of stationary phase. Using the large argument expansion of $J_0(k_\rho\,\rho)$, we find a stationary point at $k_\rho = k\sin\theta$ where $\theta = \tan^{-1}(\rho/z)$ (Ref. 11, p. 218). Thus, the leading order contribution to the above integral can be obtained by approximating the slowly varying part of the integrand with its value at the stationary point and integrate out the rapidly varying part. As such (29) becomes

$$E_z^{(0)} \sim \frac{A}{2i\omega\epsilon}\alpha_{1m}\,J_0(\alpha_{1m})\frac{k^2\sin\theta\cos\theta J_1(ka\sin\theta)}{(\alpha_{1m}/a)^2 - k^2\sin^2\theta}[1 - R^{\text{TM}}(\theta)]\int_0^\infty \frac{k_\rho}{k_z}e^{ik_z z}\,J_0(k_\rho\,\rho)\,dk_\rho, \quad r\to\infty, \tag{30}$$

where $r = (\rho^2 + z^2)^{1/2}$. We have deliberately left a slowly varying part in the integral so that it can be integrated exactly (Ref. 11, p. 213). We find

$$E_z^{(0)} \sim \frac{-A}{2\omega\epsilon}\alpha_{1m}\,J_0(\alpha_{1m})\frac{k^2\sin\theta\cos\theta\,J_1(ka\sin\theta)}{(\alpha_{1m}/a)^2 - k^2\sin^2\theta}[1 - R^{\text{TM}}(\theta)]\frac{e^{ikr}}{r}, \quad \text{where } r\to\infty. \tag{30a}$$

It can be shown that contributions from singularities on the complex plane give rise to higher order terms,[11] and thus are not important in calculating the radiation field. Due to axial symmetry, the radiation field of the disk resonator has only E_θ component, where $E_\theta = -E_z/\sin\theta$ is given by

$$E_\theta^{(0)} \sim \frac{A}{2\omega\epsilon}\alpha_{1m}\,J_0(\alpha_{1m})\frac{k^2\cos\theta\,J_1(ka\sin\theta)}{(\alpha_{1m}/a)^2 - k^2\sin^2\theta}[1 - R^{\text{TM}}(\theta)]\frac{e^{ikr}}{r}. \tag{30b}$$

From (30b), we can find the radiation pattern and the time average power radiated by the cavity.

In order to make use of (23), we need to relate the radiated power to the energy stored in the cavity. We make use of the energy stored in the unperturbed model which is the magnetic wall model. It can be shown that with the current given by (28a), the field inside the unperturbed model is

$$E_{1z} = \frac{A}{i\omega\epsilon_r}\left(\frac{\alpha_{1m}}{a}\right)J_0\left(\alpha_{1m}\frac{\rho}{a}\right), \quad \rho < a. \tag{31}$$

The total time average energy stored by the cavity can be found and is

$$\langle W_T\rangle = \frac{1}{2}\epsilon_1\int_{-d}^0\int_0^{2\pi}\int_0^a |E_{1z}|^2\,\rho\,d\rho\,d\phi\,dz = \frac{|A|^2}{2\omega\epsilon_1}\pi\alpha_{1m}^2\,dJ_0^2\alpha_{1m}). \tag{32}$$

The frequency shift due to radiation loss is

$$\Delta\omega_r = -i\frac{\langle P_r\rangle}{2\langle W_T\rangle} = \frac{-i}{4\eta\epsilon_1 d}\int_0^{\pi/2} d\theta\sin\theta\left|\frac{k_1^2\cos\theta\,J_1(ka\sin\theta)}{(\alpha_{1m}/a)^2 - k^2\sin^2\theta}[1 - R^{\text{TM}}(\theta)]\right|^2, \text{ where } \eta = (\mu/\epsilon)^{1/2}. \tag{33}$$

Similarly, the real frequency shift due to reactive power leakage can be estimated from (20) by

$$\Delta\omega_x \simeq \text{Im}\left(\hat{\rho}\cdot\int_0^{2\pi}\int_{-d}^0 (\mathbf{E}_i^*\times\mathbf{H}_f)\Big|_{\rho=a}\,adz\,d\phi\right)4\langle W_T\rangle\,. \tag{34}$$

We choose to have $\mathbf{E}_i$ in the above formula because it is known and is given by (31). We can obtain a zero-order estimate of $\mathbf{H}_f$ as follows. If the zeroth order field in the upper-region is given by (29), by matching boundary conditions, the zeroth order $\mathbf{E}$ field in region 1 is given by

$$E_{1z}^{(0)} = -\frac{A}{2i\omega\epsilon_1}\alpha_{1m}\,J_0(\alpha_{1m})\int_0^\infty \frac{k_\rho^2\,J_1(k_\rho a)}{(\alpha_{1m}/a)^2 - k_\rho^2}\frac{\cos k_{1z}(z+d)}{\cos k_{1z}\,d}(1 + R^{\text{TM}})\,J_0(k_\rho\,\rho)\,dk_\rho\,. \tag{35}$$

It can be shown similar to (3) and (4) that the $\mathbf{H}$ field is

$$\mathbf{H}_1^{(0)} = \hat{\phi}\frac{A}{2}\alpha_{1m}\,J_0(\alpha_{1m})\int_0^\infty \frac{k_\rho\,J_1(k_\rho a)}{(\alpha_{1m}/a)^2 - k_\rho^2}\frac{\cos k_{1z}(z+d)}{\cos k_{1z}\,d}(1 + R^{\text{TM}})\,J_1(k_\rho\,\rho)\,dk_\rho\,. \tag{36}$$

Therefore,

$$\int_{-d}^0\int_0^{2\pi}\mathbf{E}_i^*\times\mathbf{H}_f\Big|_{\rho=a}\cdot\hat{\rho}a\,d\phi\,dz = \frac{\pi A^2}{i\omega\epsilon_1}\alpha_{1m}^2\,J_0^2(\alpha_{1m})\int_0^\infty k_\rho\frac{J_1^2(k_\rho a)}{[(\alpha_{1m}/a)^2 - k_\rho^2]}\frac{\tan k_{1z}\,d}{k_{1z}}(1 + R^{\text{TM}})\,dk_\rho\,. \tag{37}$$

As such $\Delta\omega_x$ is given by

$$\Delta\omega_x = -\frac{\omega}{2d}\,\mathrm{Re}\left(\int_0^\infty k_\rho \frac{J_1^2(k_\rho a)}{[(\alpha_{1m}/a)^2 - k_\rho^2]}\frac{\tan k_{1z}\, d}{k_{1z}}(1 + R^{\mathrm{TM}})\, dk_\rho\right). \tag{38}$$

The above can be integrated numerically to give $\Delta\omega_x$.

VI. A NOTE ON NUMERICAL COMPUTATION

In computing (11a) and (16a) numerically, we have to define the path of integration for the integral. In searching for the zero's of the eigenequations, the resonant frequencies can be complex and have a small negative imaginary part. The function $G^{\mathrm{TM}}(k_\rho)$ as given by (9) has branch points at $k_\rho = k$ and poles between $k < k_\rho < k_1$. When the time harmonic frequency is complex, these singularities are slightly below the real axis. Thus the path of integration has to be deformed slightly below the real axis so that the migration paths of these singularities do not cross the integration path (when they do so, the integrals are undefined and diverge). The path of integration on the k_ρ plane is shown in Fig. 5.

For small d, there is only one pole in the aforementioned region, corresponding to the lowest guided TM mode in the dielectric substrate and can be located easily using the Newton–Raphson method. The Newton–Raphson method converges quickly if a good initial guess for the location of the pole is given. For small d and $\epsilon_r \neq 1$, the pole location k_g is approximately given by

$$k_g \simeq k + (k^3 d^2/2)[(\epsilon_r - 1)/\epsilon_r]^2,$$

which serves as a good initial guess to the pole location.

To reduce the number of complex arithmatics, it is advisable to deform the path of integration to the real axis deforming around the branch point and pole as shown in Fig. 5. Also, the proximity of the pole location to the integration path is detrimental to the efficient numerical evaluations of the integrals. As such, it is advantageous to subtract out the pole singularity and integrate it analytically. The remaining integrand is a smooth function and can be integrated out efficiently using Gaussian quadrature.

For example, in (11a), noting that $G^{\mathrm{TM}}(k_\rho)$ is an even function, we can rewrite

$$\frac{i\epsilon_r G^{\mathrm{TM}}(k_\rho)}{2d(k_1^2 - k_\rho^2)} = \frac{i\epsilon_r}{2d}\left[\frac{1}{(k_\rho^2 - k_g^2)}\left(\frac{H(k_\rho)}{k_1^2 - k_\rho^2} - \frac{H(k_g)}{k_1^2 - k_g^2}\right) + \frac{H(k_g)}{(k_\rho^2 - k_g^2)(k_1^2 - k_g^2)}\right]. \tag{39}$$

When the above is substituted in (11a), the last term in (39) gives rise to

$$I_{g1} = -\frac{Ai\epsilon_r H(k_g)}{2d(k_1^2 - k_g^2)}\int_0^\infty \frac{k_\rho a}{k_\rho^2 - k_g^2}\left(\frac{k_\rho J_1(k_1 a) J_2(k_\rho a) - k_1 J_2(k_1 a) J_1(k_\rho a)}{k_\rho^2 - k_1^2}\right) J_1(k_\rho \rho)\, dk_\rho\,. \tag{40}$$

When $\rho < a$, the above can be integrated exactly using contour integration (Ref. 14, p. 15) giving

$$I_{g1} = \frac{A\epsilon_r H(k_g)}{2d(k_1^2 - k_g^2)}\left[\frac{\pi}{(k_g^2 - k_1^2)}\left(\frac{a}{2}\left[k_g J_1(k_1 a) H_2^{(1)}(k_g a) - k_1 J_2(k_1 a) H_1^{(1)}(k_g a)\right] J_1(k_g \rho) - J_1(k_1 \rho)\right)\right]. \tag{41}$$

Similarly, the pole contribution in (16a) can be integrated out giving

$$I_{g2} = a^4 \alpha_{1m}\alpha_{1p} J_0(\alpha_{1m}) J_0(\alpha_{1p}) H(k_g)\left(\frac{\delta_{mp}}{2\alpha_{1p}^2(\alpha_{1p}^2 - k_g^2)} + \frac{\pi i}{2}\frac{J_1(k_g a) H_1^{(1)}(k_g a)}{(k_g^2 a^2 - \alpha_{1p}^2)(k_g^2 a^2 - \alpha_{1m}^2)}\right). \tag{42}$$

In (38), the path of integration is well defined if we assume an infinitesimal loss for the media. However, it is still advantageous to subtract out the pole location. We can write

$$\frac{\tan k_{1z}\, d}{k_{1z}}[1 + R^{\mathrm{TM}}(k_\rho)] = \left(\frac{T(k_\rho) - T(k_g)}{k_\rho^2 - k_g^2}\right) + \frac{T(k_g)}{k_\rho^2 - k_g^2}. \tag{43}$$

The pole contribution is thus

$$I_{g3} = \frac{\omega}{2d}\,\mathrm{Re}\left(\frac{-\pi i}{2} T(k_g)\frac{J_1(k_g a) H_1^{(1)}(k_g a)}{k_g^2 - (\alpha_{1m}/a)^2}\right). \tag{44}$$

For the program that searches the roots of the eigenequation Muller's method was used since it does not involve differentiation of a complicated expression.

VII. RESULTS AND CONCLUSIONS

In Figs. 6–10, we show some plots of the resonant wave number times the radius ($\omega a/c_1$) as a function of the ratio of the substrate thickness to disk radius. Figures 6–8 show the variations of $\omega a/c_1$ as a function of d/a for the lowest axial-symmetric TM resonant mode while Figs. 9 and 10 show the plots for the next lowest axial-symmetric TM resonant mode.

In Figs. 6 and 7, results for Galerkin's method using one and two basis functions are shown. It is seen that $\mathrm{Im}(\omega a/c_1)$

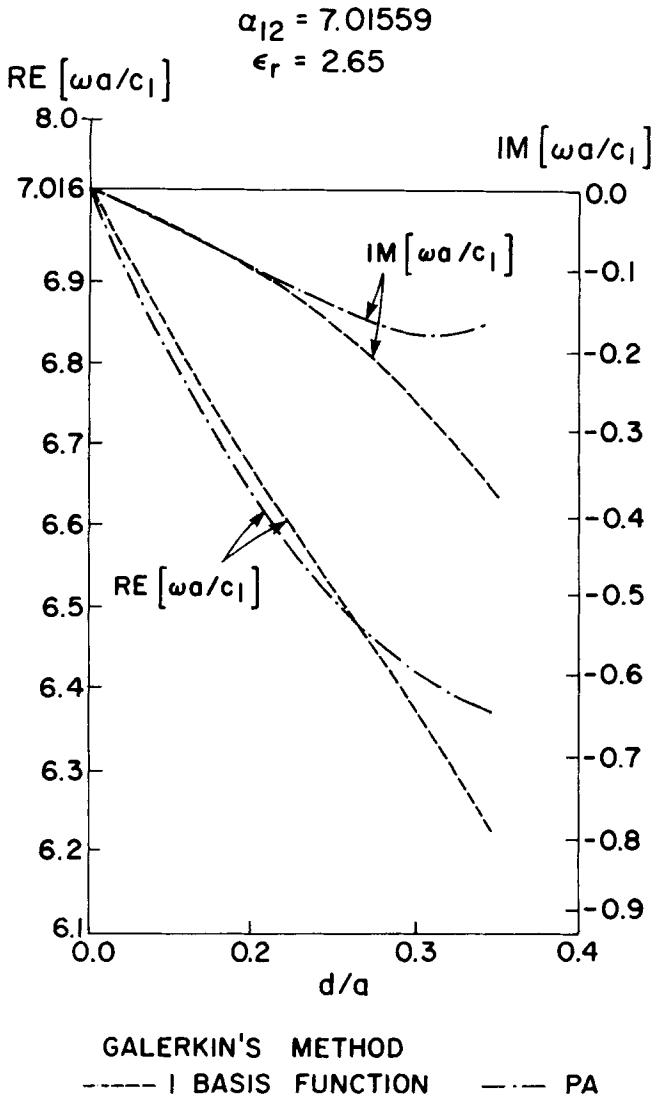

FIG. 10. Resonant frequency shift for the second lowest axial symmetric TM mode when $\epsilon_r = 2.65$ (PA-perturbation approach).

using one or two basis functions are essentially the same while $\mathrm{Re}(\omega a/c_1)$ shows some discrepancies which increase for increasing d/a. As we have seen from the perturbation formula, $\mathrm{Im}(\omega a/c_1)$ shift is due to radiation loss in the far field while $\mathrm{Re}(\omega a/c_1)$ variation is due to reactive power loss in the near field. Therefore, we can conclude that the addition of more basis functions in our computation improves the near field (fringing field) but affects the far field very little. As we increase our number of basis functions, the resulting current distribution in (13) approaches the actual current distribution which we know has singular derivatives at the edge ($\nabla\cdot\mathbf{K} = i\omega\sigma_{s'}$, where $\sigma_{s'}$ the charge distribution is singular at the edge). This charge singlarity at the edge determines the singular behavior of the near field.

In Fig. 8, we show the computation using the iterative approach as in (11a). We see large discrepancy in $\mathrm{Re}(\omega a/c_1)$ as d/a increases, and $\mathrm{Im}(\omega a/c_1)$ agrees well with results obtained by Galerkin's method. The ill convergence of the iterative process in solving dual integration equations is discussed by Leppington and Levine.[15] This is because the second term on the right-hand side of (11a) is not small when d is fixed and $k_\rho \to \infty$. As the integration is from $k_\rho = 0$ to $+\infty$, the validity of neglecting the second term in calculating the zeroth order solution when d is fixed is clearly violated for the large k_ρ contribution of the integral. The large k_ρ contribution of the integral corresponds to the singular part of the current or near field. Thus the iterative approach gives crude approximation to $\mathrm{Re}(\omega a/c_1)$ whose variation is intimately related to the near field while it approximate $\mathrm{Im}(\omega a/c_1)$ fairly well which is associated with the far field.

In Figs. 6–10, the use of the perturbation approach is shown alongside with other plots. We see that the validity of the perturbation approach is valid over a larger range of d/a for larger ϵ_r. The perturbation approach is valid when the field inside the original magnetic-wall model is not drastically changed when the magnetic wall is removed. For large ϵ_r most of the energy is trapped inside the disk resonator and thus the field resembles that of the magnetic-wall model when d/a is small, say, compared to ϵ_r near one. This explains our aforementioned observation.

In general, we find that the frequency shift due to radiation loss increases monotonically for increasing d/a and this shift decreases for increasing ϵ_r. The real frequency shift, which is proportional to

$$\iiint_{V_0} [\mu |\bar{H}|^2 - \epsilon(\bar{r})|\bar{E}|^2]\, dV,$$

might increase or decrease for increasing ϵ_r depending on the near field configuration of the resonant mode on the outside.

We have presented three methods of computing resonant frequency shift. The methods agree within their domains of validity. Both the Galerkin's method and iterative methods involve the search for complex zeros on the complex plane and numerical integration. Thus they are time consuming, especially for the iterative method, which involves the numerical integration of a function which decays as $1/k_\rho^2$ over an infinite range. Therefore, the iterative method is unsuitable for numerical computation both for its time consuming factor and poor approximation to $\mathrm{Re}(\omega a/c_1)$. Galerkin's method is a variational method and thus the ap-

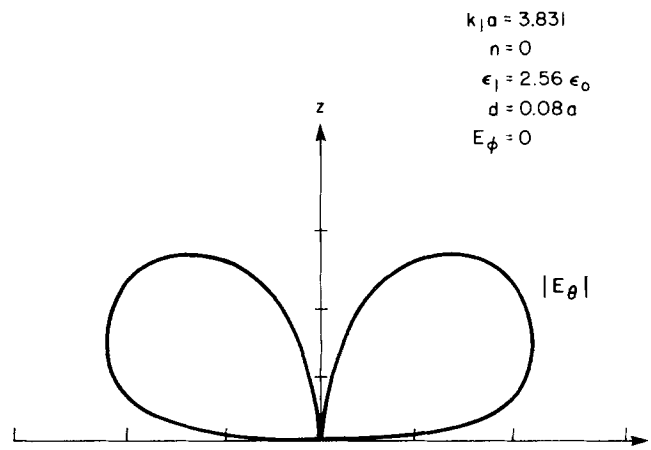

FIG. 11. Radiation pattern due to the lowest axial symmetric TM mode.

proximate results do not depart drastically from the actual result if judicious choice of basis function is made. It is also a method which can provide a result to high precision depending on the amount of computation. On the other hand, the perturbation approach gives a good approximaton to the correct result for ϵ_r sufficiently large and d/a small. For example, for $\epsilon_r \geqslant 2.65$, it provides good results for $d/a \leqslant 0.1$. Since the computation time associated with the perturbation approach is considerably less compared with the other methods, it is a superior approach within its range of validity, which in general is valid for microstrip disk resonators. It is also a much better approximation compared with the quasi-static approximation shown in Figs. 6 and 7. In Fig. 11, we also show the radiation pattern due to the lowest axial-symmetric mode.

[1]J. Watkins, "Circular resonant structures in microstrip," Electron Lett. **5**, 524–25 (1978).

[2]S. Mao, S. Jones, and G.D. Vandelin, "Millimeter-wave integrated circuits," IEEE Trans. Microwave Theory Tech. (Special Issue on Microwave Integrated Circuits), MTT-**16**, 455–61 (1968).

[3]I. Wolff and N. Knoppik, "Rectangular and circular microstrip disk capacitors and resonators," IEEE Trans. Microwave Theory Tech. MTT-**22**, 857 (1974).

[4]S.R. Borkar and R.F.M. Yang, "Capacitance of a circular disk for applications in microwave integrated circuits," IEEE Trans. Microwave Theory Tech. MTT-**23**, 583–91 (1975).

[5]L.C. Shen, S.A. Long, M.R. Allerding, and M.D. Walton, "Resonator frequency of a circular disc printed circuit antenna," IEEE Trans. Antenna Propagation (Communication), AP-**25**, 595–96 (1977).

[6]T. Itoh, "Analysis of microstrip resonators," IEEE Trans. Microwave Theory Tech. MTT-**22**, 946 (1974).

[7]E.C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Oxford, New York, 1937), p. 337.

[8]C.J. Tranter, "On some dual integral equations occurring in potential problems with axial symmetry," Q.J. Mech. Appl. Math. **3**, 411 (1950).

[9]J.C. Cooke, "A solution of Tranter's dual integral equation problem," Q.J. Mech. Appl. Math. **9**, 103 (1956).

[10]I.N. Sneddon, *Mixed Boundary Value Problems in Potential Theory* (Wiley, New York, 1966).

[11]J.A. Kong, *Theory of Electromagnetic Waves* (Wiley, New York, 1975).

[12]M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

[13]S.A. Long, L.C. Shen, and P.B. Morel, "Theory of the printed circuit antenna," Proc. IEE, **125**, 925–28 (1978).

[14]C.T. Tai, *Dyadic Green's Functions in Electromagnetic Theory* (International, Scranton, 1971).

[15]F. Leppington and H. Levine, "On the capacity of the circular disk condenser at small separation," Proc. Cambridge Philos. Soc. **68**, 235 (1970).

Dynamic polarizability of the double-quadratic kink[a)]

S. E. Trullinger
Department of Physics, University of Southern California, Los Angeles, California 90007

(Received 12 December 1978)

We discuss a simple model of ferroelectric domain walls in terms of solitary-wave (kink) solutions of the equation of motion for the double-quadratic (DQ) chain. The response of the DQ kink to an external field is calculated and a very simple expression is obtained for the dynamic polarizability of the DQ kink. The simplicity of this model makes it more attractive than the usual ϕ^4 model for studying the interaction of ferroelectric domain walls with external electric fields.

I. INTRODUCTION

Solitary wave or "kink" solutions of nonlinear wave equations have been rapidly gaining popularity in condensed matter physics.[1] Their remarkable stability and other particlelike properties have prompted their use as models for a wide variety of nonlinear structures.[1] One type of solitary wave in particular, the so-called[2] "ϕ^4 kink," had received a great deal of attention recently[2-8] as a model for domain walls in displacive ferroelectrics. The behavior of this kink in condensed matter contexts has been well studied. For example, statistical mechanics investigations[2-4,6,9] have shown that ϕ^4 kinks behave as elementary excitations at finite (but low) temperature, and phenomenologies which treat them as an "ideal gas" of "particles" have been developed.[2,3,9] In addition, perturbation theories have been developed[5-8,10,11] to study the influence of various types of perturbing forces.

A topic of particular importance in the context of domain walls in ferroelectrics is the manner in which kinks respond to an applied electric field (either static or oscillatory). The response of the ϕ^4 kink to an oscillating field has been studied recently by Theodorakopoulos and co-workers,[7] using an expansion technique proposed by Fogel, *et al.*,[10,11] which makes use of the completeness property of small oscillations about the kink waveform. Lee and Trullinger[8] have recently studied the generalized susceptibility of kink solutions of the nonlinear Klein–Gordon equation, and applied their general formalism to the sine-Gordon[10] and ϕ^4 kinks as examples. Even for the special case of homogeneous applied fields, however, the expressions for the long-wavelength susceptibility (termed the dynamic polarizability) of the ϕ^4 kink are extremely cumbersome to evaluate and require the performance of infinite summations. This feature detracts a great deal from the usefulness of the ϕ^4 kink as a model for studying the polarizability of domain walls.

In this paper, we investigate an alternate model for domain walls in ferroelectrics which leads to a much simpler expression for the dynamic polarizability and therefore lends itself quite easily to further study of the interaction of domain walls with applied electric fields. The model arises by replacing the double-well ϕ^4 potential by another potential which also has double-well character but a different analytic form, namely the double-quadratic (DQ) potential.[12] This potential consists of two displaced parabolas (see Fig. 1) and gives rise to solitary wave solutions of the unperturbed equation of motion. The forms of the kink solution and accompanying small oscillations are simple enough to allow the entire calculation described below to be carried out analytically. The final result for the dynamic polarizability of the DQ kink is quite simple and should allow researchers in the field to proceed quite far in further investigations.

In Sec. II below, we describe the system of double-quadratic oscillators in detail. We present the kink solution to the equation of motion which obtains in the displacive limit and characterize the small oscillations about the kink waveform. The completeness property of these small oscillations (phonons) is then used in Sec. III in applying the general formalism of Ref. 8 to obtain the dynamic polarizability of the DQ kink. In addition to the simple expression for the polarizability, plots are presented for its real and imaginary parts as a function of frequency, and we show that the Kramers–Kronig relations are satisfied. In Sec. IV we give a brief summary.

II. THE DOUBLE-QUADRATIC KINK

In this section we describe an alternate model to the ϕ^4 model proposed by Krumhansl and Schrieffer[2] for a one-dimensional ferroelectric. The kink solution (domain wall) obtained for this new model has the virtue that its dynamic polarizability can be expressed in simple closed form, as opposed to the cumbersome series form found[8] for the ϕ^4 kink. In addition to describing the DQ kink solution, we also examine the nature of small oscillations of the displacement field in the presence of a kink, since these play a fundamental role in obtaining the dynamic polarizability in Sec. III. The analysis presented in this section closely follows that used by Trullinger and DeLeonardis[13] in their examination of the statistical mechanics of the double-quadratic chain.

The system under consideration consists of a one-dimensional chain of harmonically coupled oscillators governed by the following Hamiltonian:

$$H = lA \sum_i \left(\frac{1}{2} \dot{\phi}_i^2 + \frac{1}{2} \frac{c_0^2}{l^2} (\phi_{i+1} - \phi_i)^2 + \omega_0^2 V(\phi_i) \right), \quad (2.1)$$

where ϕ_i is the dimensionless displacement coordinate of the

[a)]Research supported by the National Science Foundation under grant No. DMR-77-08445.

 0022-2488/80/030592-07$1.00

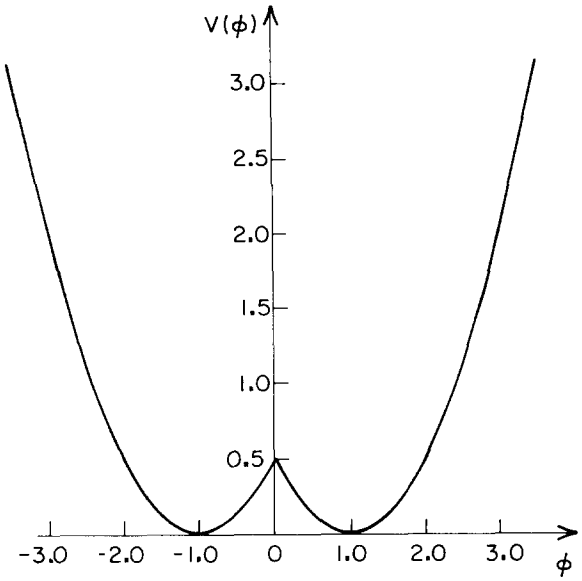

FIG. 1. The double-quadratic (DQ) potential.

ith oscillator, l is the equilibrium spacing between nearest neighbors, and $V(\phi)$ is an "on-site" potential having the form of a double-quadratic (DQ) well[12] (as shown in Fig. 1):

$$V(\phi) = \tfrac{1}{2}(|\phi| - 1)^2. \tag{2.2}$$

The first term in Eq. (2.1) represents the kinetic energy carried by the displacement field (a dot denotes a time derivative) and the second term represents harmonic coupling (strain energy) between displacements at neighboring lattice sites. The constant c_0 is the characteristic velocity in the system and represents the limiting velocity of the kink (see below). The constant ω_0 is the characteristic frequency of oscillation and represents the limiting frequency of long-wavelength phonons. The overall constant A sets the energy scale and has dimensions of (energy)$\times$(length)$^{-1}$ $\times$(time)^2.

We restrict ourselves to the displacive limit[2] when the coupling between sites is strong enough to ensure that variations of ϕ from site to site are quite small, at least at low temperatures. In this limit we may replace the site index i by a continuous position variable x so that ϕ becomes a continuous function of x and t, $\phi = \phi(x,t)$. The relevant length scale then becomes $d = c_0/\omega_0$ ($d \gg l$) and it is in this limit that nonlinear kinks become well-defined[14,15] elementary excitations with long lifetimes and as such behave[10] very much like extended particles.

In the continuum (displacive) limit the Hamiltonian (2.1) is replaced by

$$H = A\int dx\{\tfrac{1}{2}[\dot\phi(x,t)]^2 + \tfrac{1}{2}c_0^2[\phi_x(x,t)]^2 + \omega_0^2 V(\phi)\}, \tag{2.3}$$

where $\phi_x(x,t) \equiv (\partial/\partial x)\phi(x,t)$ replaces the finite difference $(\phi_{i+1} - \phi_i)/l$. The excitations we are concerned with arise as solutions of the Euler–Lagrange equation of motion following from Eq. (2.3):

$$\ddot\phi - c_0^2\phi_{xx} + \omega_0^2(|\phi| - 1)\mathrm{sgn}\phi = 0. \tag{2.4}$$

The linear "phonon" solutions of Eq. (2.4) have the form

$$\phi \pm 1 = \phi_0 \cos(kx - \omega_k t), \tag{2.5}$$

where the magnitude of ϕ_0 must be less than 1 ($|\phi_0| < 1$) but is *not* required to be infinitesimally small, since the individual potential wells are *perfectly* harmonic for $|\phi_0| < 1$. The dispersion relation for these solutions is given by

$$\omega_k^2 = \omega_0^2 + c_0^2 k^2, \tag{2.6}$$

which is the continuum limit of the discrete lattice dispersion relation,

$$\omega_k^2 = \omega_0^2 + 4(c_0/l)^2 \sin^2(lk/2). \tag{2.7}$$

The solitary-wave (kink) solutions of Eq. (2.4) evolve the displacement field from one minimum of the DQ well to the other minimum and thus represent the domain walls in the system. Because of the covariance of Eq. (2.4) we may solve first for the static kink waveform and then "boost" the solution to any frame moving with velocity v ($|v| < c_0$). The static kink solutions can be obtained quite easily by considering the regions $x > 0$ and $x < 0$ separately. For the kink solution we impose the conditions $\phi(+\infty) = 1$, $\phi(0) = 0$, $\phi(-\infty) = -1$, while for the antikink solution the sign of ϕ is reversed. We find the following static kink waveform:

$$\phi_K^{(0)} = \pm(\mathrm{sgn}x)[1 - \exp(-|x|/d)], \tag{2.8}$$

where $+(-)$ is appropriate to the kink (antikink) solution. The traveling kink solutions are obtained by boosting to velocity v:

$$\phi_K^{(v)} = \pm \mathrm{sgn}\left[\frac{x - vt}{(1 - v^2/c_0^2)^{1/2}}\right] \times\left[1 - \exp\left(-\frac{|x - vt|}{d(1 - v^2/c_0^2)^{1/2}}\right)\right]. \tag{2.9}$$

In Fig. 2 we have plotted the waveform of the kink in its rest frame.

The energy required to create a kink (or antikink) can be obtained by substituting the static kink solution (2.8) into Eq. (2.3) and performing the integration over x. The kink rest energy is thus found to be

$$E_K^{(0)} = A\omega_0 c_0 = M_K c_0^2, \tag{2.10}$$

and the energy of a kink moving with velocity v is simply

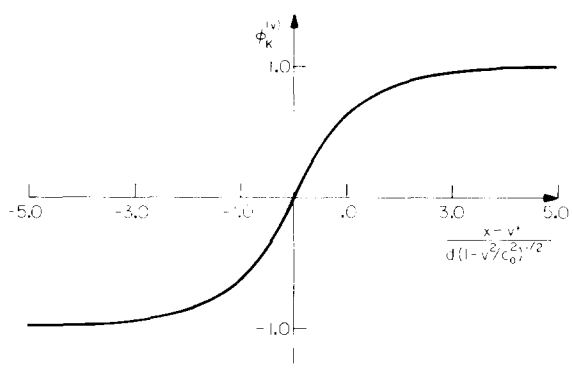

FIG. 2. The waveform of the traveling kink, $\phi_K^{(v)}$, viewed in its rest frame. The antikink waveform is obtained by reflection through the horizontal axis.

$$E_K^{(v)} = A\omega_0 c_0 (1 - v^2/c_0^2)^{-1/2}. \tag{2.11}$$

The "rest mass" of the kink is

$$M_K = A/d. \tag{2.12}$$

We now turn our attention to an examination of small oscillations in the presence of a kink. We suppose that a kink is at rest at the origin ($x = 0$). The subsequent analysis can be made applicable to a moving kink as well by simply transforming to the kink rest frame. We write

$$\phi(x,t) = \phi_K^{(0)}(x) + \psi(x,t), \tag{2.13}$$

where $\phi_K^{(0)}(x)$ is the static kink waveform (2.8) and $\psi(x,t)$ is assumed to be a small deviation. Using the fact that $\phi_K^{(0)}$ satisfies Eq. (2.4), we find that $\psi(x,t)$ is governed by

$$\ddot{\psi}(x,t) - c_0^2 \psi_{xx}(x,t) + \omega_0^2 V''(\phi_K^{(0)}(x))\psi(x,t) = 0. \tag{2.14}$$

Assuming a solution of the form

$$\psi(x,t) = f(x)e^{-i\omega t}, \tag{2.15}$$

and making use of the relation,

$$V''(\phi_K^{(0)}(x)) = 1 - 2d\delta(x), \tag{2.16}$$

where $\delta(x)$ is the Dirac delta function, Eq. (2.14) reduces to

$$-c_0^2 f_{xx}(x) - 2d\omega_0^2 \delta(x) f(x) = (\omega^2 - \omega_0^2) f(x). \tag{2.17}$$

This equation has the form of Schrödinger's equation for a "particle" in the presence of a delta function potential well. There exists exactly one "bound state" with eigenvalue,

$$\omega_b^2 = 0, \tag{2.18}$$

and corresponding normalized eigenfunction,

$$f_b(x) = d^{-1/2} \exp(-|x|/d). \tag{2.19}$$

We note that this bound state solution for small oscillations about the kink simply corresponds to the "translation mode"[8-11,16-19] of the kink and its presence is required by Goldstone's theorem, i.e., the translation mode is the Goldstone mode which restores translation invariance upon introduction of the kink to the system.

Since the delta function potential is even in x, we may classify the remaining "continuum" states by their parity under reflection through $x = 0$. Both classes of eigenfunctions must satisfy the continuity condition

$$\lim_{x \to 0^-} f(x) = \lim_{x \to 0} f(x) \equiv f(0), \tag{2.20}$$

and the relation

$$\left.\frac{df}{dx}\right|_{x=0} - \left.\frac{df}{dx}\right|_{x=0^-} = \frac{2}{d} f(0), \tag{2.21}$$

which is obtained by integrating Eq. (2.17) over an infinitesimal neighborhood of $x = 0$. The continuum states of odd parity are unaffected by the presence of the delta function potential since $f(x)$ vanishes at $x = 0$ if $f(x)$ is odd. Thus,

$$f_{k,-}(x) = S_k \sin kx, \tag{2.22}$$

where the minus sign ($-$) denotes odd parity and S_k is a normalization constant. The states of even parity are found to have the form

$$f_{K,+}(x) = \frac{C_k}{(1 + k^2 d^2)^{1/2}} [kd \cos kx - \sin k|x|], \tag{2.23}$$

where C_k is a normalization constant. The eigenvalue ω_k for the even state labeled by k is the same as the eigenvalue for the odd state labeled by k and is given by the dispersion relation (2.6), i.e., the frequencies of the small oscillations are *unaffected* by the presence of a kink. Note that for both the even and odd states, k and $-k$ do not label *distinct* states. Thus, we make the arbitrary choice that odd states be labeled by negative k values and even states by positive k values.

For convenience in what follows, we now employ the dimensionless quantities

$$z = x/d, \quad \tau = \omega_0 t, \quad \kappa = kd, \quad \bar{\omega}_\kappa = \omega_k/\omega_0. \tag{2.24}$$

In terms of these quantities, we have

$$f_b(z) = \exp(-|z|), \tag{2.25a}$$

$$f_{\kappa,-}(z) = \frac{1}{\sqrt{\pi}} \sin \kappa z \quad (\kappa < 0), \tag{2.25b}$$

$$f_{\kappa,+}(z) = \frac{1}{\sqrt{\pi}(1 + \kappa^2)^{1/2}} [\kappa \cos \kappa z - \sin \kappa |z|] \quad (\kappa > 0). \tag{2.25c}$$

The normalization constants have been chosen so that the eigenfunctions satisfy the orthonormality conditions,

$$\int_{-\infty}^{+\infty} dz f_b(z) f_b(z) = 1, \tag{2.26a}$$

$$\int_{-\infty}^{+\infty} dz f_{\kappa,-}(z) f_{\kappa',-}(z) = \tfrac{1}{2}\delta(\kappa - \kappa') \quad (\kappa, \kappa' < 0), \tag{2.26b}$$

$$\int_{-\infty}^{+\infty} dz f_{\kappa,+}(z) f_{\kappa',+}(z) = \tfrac{1}{2}\delta(\kappa - \kappa') \quad (\kappa, \kappa' > 0), \tag{2.26c}$$

$$\int_{-\infty}^{+\infty} dz f_{\kappa,-}(z) f_{\kappa',+}(z) = 0, \tag{2.26d}$$

$$\int_{-\infty}^{+\infty} dz f_b(z) f_{\kappa,\pm}(z) = 0, \tag{2.26e}$$

and the completeness condition,

$$f_b(z) f_b(z') + \int_{-\infty}^{0} d\kappa f_{\kappa,-}(z) f_{\kappa,-}(z') + \int_{0}^{\infty} d\kappa f_{\kappa,+}(z) f_{\kappa,+}(z') = \delta(z - z'). \tag{2.27}$$

Since the set of functions $\{f_b, f_{\kappa,-}, f_{\kappa,+}\}$ is complete, we may use them to expand[10] deviations of the kink waveform in the presence of external perturbations. In particular, we shall use them in this manner in the next section where we give a discussion of the kink response to an external field in the presence of linear damping.

III. DYNAMIC POLARIZABILITY OF THE DQ KINK

In this section we use the general formalism of Lee and Trullinger[8] to calculate the dynamic polarizability of the double-quadratic (DQ) kink described in the previous section. We obtain a very simple closed-form expression for the polarizability and present plots of its real and imaginary parts for representative values of the damping constant.

We consider the effect of an external electric field, $E(z,\tau)$, together with linear damping on the kink solution $\phi_K^{(0)}(z)$. Since the dimensionless electric field E couples lin-

early to the displacement field ϕ (i.e., to the displacement of charged ions), the equation of motion (2.4) now assumes the dimensionless form

$$\frac{\partial^2\phi}{\partial\tau^2} - \frac{\partial^2\phi}{\partial z^2} + V'(\phi) + \Gamma\frac{\partial\phi}{\partial\tau} = E(z,\tau), \tag{3.1}$$

where Γ is a dimensionless damping constant and $V'(\phi)$ is given by

$$V'(\phi) = (|\phi| - 1)\mathrm{sgn}\phi. \tag{3.2}$$

The form of $E(z,\tau)$ is left arbitrary for the present. We specialize later to the case where E may be taken to be spatially uniform (at least over a distance large compared with the domain wall width).

We seek a solution of Eq. (3.1) of the form

$$\phi(z,\tau) = \phi_K^{(0)}(z) + \Delta\phi(z,\tau), \tag{3.3}$$

i.e., we assume that a response $\Delta\phi$ to the disturbance is superimposed on the initially static solitary wave. We assume that $|\Delta\phi|$ is small[20] if $E(z,\tau)$ is small in magnitude. Substitution of (3.3) into (3.1) then yields the following equation for $\Delta\phi$:

$$\frac{\partial^2\Delta\phi}{\partial\tau^2} - \frac{\partial^2\Delta\phi}{\partial z^2} + V''(\phi_K^{(0)}(z))\Delta\phi + \Gamma\frac{\partial\Delta\phi}{\partial\tau} = E(z,\tau), \tag{3.4}$$

where

$$V''(\phi_K^{(0)}(z)) = 1 - 2\delta(z). \tag{3.5}$$

The solution of the linear equation (3.4) can be expressed formally in terms of a linear integral operator whose kernel we denote by $K(z,z';\tau-\tau')$:

$$\Delta\phi(z,\tau) = \int_{-\infty}^{\tau} d\tau' \int_{-\infty}^{+\infty} dz'\, K(z,z';\tau-\tau')E(z',\tau'). \tag{3.6}$$

The upper limit of the τ' integration in (3.6) is bounded by τ due to causality. It is convenient to define

$$K(z,z';\tau) \equiv 0 \quad (\tau < 0), \tag{3.7}$$

and rewrite Eq. (3.6) as

$$\Delta\phi(z,\tau) = \int_{-\infty}^{+\infty} d\tau' \int_{-\infty}^{\infty} dz'\, K(z,z';\tau')E(z',\tau-\tau'). \tag{3.8}$$

We now introduce Fourier time transforms of $\Delta\phi$, E and K:

$$\Delta\hat{\phi}(z,\Omega) = \int_{-\infty}^{+\infty} d\tau\, e^{-i\Omega\tau}\Delta\phi(z,\tau), \tag{3.9a}$$

$$\hat{E}(z,\Omega) = \int_{-\infty}^{+\infty} d\tau\, e^{-i\Omega\tau}E(z,\tau), \tag{3.9b}$$

$$\alpha(z,z';\Omega) = \int_{-\infty}^{+\infty} d\tau\, e^{-i\Omega\tau}K(z,z';\tau). \tag{3.9c}$$

It then follows from Eq. (3.8) that

$$\Delta\hat{\phi}(z,\Omega) = \int_{-\infty}^{+\infty} dz'\, \alpha(z,z';\Omega)\hat{E}(z',\Omega). \tag{3.10}$$

In order to express $\alpha(z,z';\Omega)$ in terms of the complete set of eigenfunctions, $\{f_b, f_{\kappa,+}, f_{\kappa,-}\}$, discussed in Sec. II, we first expand $E(z,\tau)$ and $\Delta\phi(z,\tau)$ in terms of this complete set:

$$E(z,\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\Omega\, e^{i\Omega\tau}\hat{E}(z,\Omega) \tag{3.11a}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\Omega\, e^{i\Omega\tau}\int_m \hat{E}_m(\Omega)f_m(z), \tag{3.11b}$$

and

$$\Delta\phi(z,\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\Omega\, e^{i\Omega\tau}\Delta\hat{\phi}(z,\Omega) \tag{3.12a}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\Omega\, e^{i\Omega\tau}\int_m \Delta\hat{\phi}_m(\Omega)f_m(z), \tag{3.12b}$$

where the generalized summation symbol $\int_m$ denotes the following when applied to any function g_m subscripted by m:

$$\int_m g_m \equiv g_b + 2\int_{-\infty}^{0} d\kappa\, g_{\kappa,-} + 2\int_0^{\infty} d\kappa\, g_{\kappa,+}, \tag{3.13}$$

i.e., it indicates the addition of a "bound state" portion (b) to ordinary integrations of continuum portions (κ, + and κ, −). The functions $\hat{E}_m(\Omega)$, for example, represent the coefficients $\hat{E}_b(\Omega)$, $\hat{E}_{\kappa,+}(\Omega)$, $\hat{E}_{\kappa,-}(\Omega)$ in the expansion of $\hat{E}(z,\Omega)$ in terms of the complete set $\{f_b(z), f_{\kappa,+}(z), f_{\kappa,-}(z)\}$.

Substituting Eqs. (3.11b) and (3.12b) into Eq. (3.4) and using Eq. (2.4) (in its dimensionless form) we obtain

$$\int_{-\infty}^{+\infty}\frac{d\Omega'}{2\pi}e^{i\Omega'\tau}\int_m \Delta\hat{\phi}_m(\Omega')[\bar{\omega}_m^2 - \Omega'^2 + i\Gamma\Omega']f_m(z) = \int_{-\infty}^{+\infty}\frac{d\Omega'}{2\pi}e^{i\Omega'\tau}\int_m \hat{E}_m(\Omega')f_m(z). \tag{3.14}$$

Multiplying both sides of Eq. (3.14) by $e^{-i\Omega\tau}f_j(z)$ and integrating over z and τ, we obtain

$$\Delta\hat{\phi}_j(\Omega) = \hat{E}_j(\Omega)/[\bar{\omega}_j^2 - \Omega^2 + i\Gamma\Omega], \tag{3.15}$$

where we have used Eqs. (2.26). Substitution of Eq. (3.15) into Eq. (3.12b) yields

$$\Delta\phi(z,\tau) = \int_{-\infty}^{+\infty}\frac{d\Omega}{2\pi}e^{i\Omega\tau}\int_j \frac{\hat{E}_j(\Omega)f_j(z)}{\bar{\omega}_j^2 - \Omega^2 + i\Gamma\Omega} = \int_{-\infty}^{+\infty}\frac{d\Omega}{2\pi}e^{i\Omega\tau}\int_j\int_{-\infty}^{+\infty} dz' \times \frac{\hat{E}(z',\Omega)f_j(z')f_j(z)}{\bar{\omega}_j^2 - \Omega^2 + i\Gamma\Omega}, \tag{3.16}$$

where this last equality is obtained using the inverse transform of Eq. (3.11b),

$$\hat{E}_j(\Omega) = \int_{-\infty}^{+\infty} dz'\, \hat{E}(z',\Omega)f_j(z'), \tag{3.17}$$

and Eqs. (2.26) and (3.13). The Fourier transform of Eq. (3.16) with respect to τ then gives

$$\Delta\hat{\phi}(z,\Omega) = \int_{-\infty}^{+\infty} dz'\left\{\int_j \frac{f_j(z)f_j(z')}{\bar{\omega}_j^2 - \Omega^2 + i\Gamma\Omega}\right\}\hat{E}(z',\Omega). \tag{3.18}$$

By comparing Eq. (3.18) with Eq. (3.10), we obtain an expression for $\alpha(z,z';\Omega)$ in terms of $\{f_j\}$:

$$\alpha(z,z';\Omega) = \int_j \frac{f_j(z)f_j(z')}{\bar{\omega}_j^2 - \Omega^2 + i\Gamma\Omega}. \tag{3.19}$$

In this paper we wish to consider only that portion of the response corresponding to the deformation of the kink waveform. For this purpose we drop the term with $j = b$

from the j sum in Eq. (3.19), since this term presents the contribution of the translation mode without deformation. We denote this modified response function by $\alpha'(z,z';\Omega)$, i.e.,

$$\alpha'(z,z';\Omega) = \sum_j{}' \frac{f_j(z)f_j(z')}{\bar{\omega}_j^2 - \Omega^2 + i\Gamma\Omega}, \tag{3.20}$$

where the prime on the generalized summation symbol indicates that the term with $j = b$ is to be excluded. From Eq. (3.13), we obtain then

$$\alpha'(z,z';\Omega) = 2\int_{-\infty}^{0} d\kappa \frac{f_{\kappa,-}(z)f_{\kappa,-}(z')}{\kappa^2 + 1 - \Omega^2 + i\Gamma\Omega} + 2\int_{0}^{-\infty} d\kappa \frac{f_{\kappa,+}(z)f_{\kappa,+}(z')}{\kappa^2 + 1 - \Omega^2 + i\Gamma\Omega}. \tag{3.21}$$

Upon making use of Eqs. (2.25b) and (2.25c) and performing the κ integrations, this becomes

$$\alpha(z,z';\Omega) = -\frac{i}{2Q}e^{iQ|z+z'|} + \frac{i}{2Q}e^{iQ|z-z'|} + \frac{i}{2Q}\frac{Q+i}{Q-i}e^{iQ(|z|+|z'|)} + \frac{i}{2Q}e^{iQ||z|-|z'||} + \frac{2}{1+Q^2}e^{-(|z|+|z'|)}, \tag{3.22}$$

where Q denotes the branch of $(\Omega^2 - 1 - i\Gamma\Omega)^{1/2}$ having a positive imaginary part.

In general, $f_{\kappa,\pm}(z)$ is nonvanishing in the limit as $|z|\to\infty$. Therefore, $\alpha'(z,z';\Omega)$ contains a portion which measures the response of the displacement field far from the center ($z = 0$) of the kink. This can be seen from Eq. (3.22) by setting $z' = z$ and letting $|z|$ tend to infinity; there remains a portion due to the second and fourth terms, namely,

$$\lim_{|z|\to\infty} \alpha'(z,z;\Omega) = i/Q. \tag{3.23}$$

In order to obtain the *intrinsic* response of the kink, we must substract the response, $\alpha^{(0)}(z,z';\Omega)$, of the system in the *absence* of a kink:

$$\alpha^{(0)}(z,z';\Omega) = 2\int_{-\infty}^{0} d\kappa \frac{f^{(0)}_{\kappa,-}(z)f^{(0)}_{\kappa,-}(z')}{\kappa^2 + 1 - \Omega^2 + i\Gamma\Omega} + 2\int_{0}^{\infty} d\kappa \frac{f^{(0)}_{\kappa,+}(z)f^{(0)}_{\kappa,+}(z')}{\kappa^2 + 1 - \Omega^2 + i\Gamma\Omega}, \tag{3.24}$$

where the kink-free functions $f^{(0)}_{\kappa,\pm}$ are simply

$$f^{(0)}_{\kappa,-}(z) = \frac{1}{\sqrt{\pi}}\sin\kappa z \quad (\kappa < 0) \tag{3.25a}$$

and

$$f^{(0)}_{\kappa,+}(z) = \frac{1}{\sqrt{\pi}}\cos\kappa z \quad (\kappa \geqslant 0). \tag{3.25b}$$

The κ integrations in Eq. (3.24) are readily performed to yield

$$\alpha^{(0)}(z,z';\Omega) = (i/Q)e^{iQ|z-z'|}. \tag{3.26}$$

The kink response is then given by

$$\bar{\alpha}(z,z';\Omega) = \alpha'(z,z';\Omega) - \alpha^{(0)}(z,z';\Omega) = -\frac{i}{2Q}e^{iQ|z+z'|} - \frac{i}{2Q}e^{iQ|z-z'|} + \frac{i}{2Q}\frac{Q+i}{Q-i}e^{iQ(|z|+|z'|)} + \frac{i}{2Q}e^{iQ||z|-|z'||} + \frac{2}{1+Q^2}e^{-(|z|+|z'|)}. \tag{3.27}$$

Setting $z' = z$ and taking the limit as $|z|\to\infty$, we see that now the second and fourth terms cancel to give

$$\lim_{|z|\to\infty} \bar{\alpha}(z,z;\Omega) = 0, \tag{3.28}$$

as required for the response of the kink alone.

It is often convenient to deal with the Fourier transform of the response function, i.e., the generalized susceptibility[8] given by

$$\hat{\alpha}(p,q;\Omega) \equiv \int_{-\infty}^{+\infty} dz \int_{-\infty}^{+\infty} dz'\, e^{-ipz} e^{-iqz'} \bar{\alpha}(z,z';\Omega). \tag{3.29}$$

From Eq. (3.18), we see that $\hat{\alpha}(p,q;\Omega)$ gives the pth Fourier component of the kink response to the qth component of the applied electric field. Substitution of Eq. (3.27) into Eq. (3.29) yields the simple result,

$$\hat{\alpha}(p,q;\Omega) = \frac{4Q(i+Q)}{(1+Q^2)(Q^2-q^2)(Q^2-p^2)} + \frac{8}{(1+Q^2)(1+q^2)(1+p^2)}. \tag{3.30}$$

Equation (3.30) may be written explicitly in terms of Ω by recalling the definition of Q

$$Q = (\Omega^2 - 1 - i\Gamma\Omega)^{1/2}, \quad \mathrm{Im}Q > 0. \tag{3.31}$$

Thus,

$$\hat{\alpha}(p,q;\Omega) = \frac{4[i(\Omega^2 - 1 - i\Gamma\Omega)^{1/2} + \Omega^2 - 1 - i\Gamma\Omega]}{\Omega(\Omega - i\Gamma)(\Omega^2 - \bar{\omega}_q^2 - i\Gamma\Omega)(\Omega^2 - \bar{\omega}_p^2 - i\Gamma\Omega)} + \frac{8}{\Omega(\Omega - i\Gamma)\bar{\omega}_q^2\bar{\omega}_p^2}, \tag{3.32}$$

where we have recalled that $1 + q^2 = \bar{\omega}_q^2$ and $1 + p^2 = \bar{\omega}_p^2$.

In the special case where $E(z,\tau)$ is uniform in space (at least over a distance large compared to the kink width) we may focus on $\hat{\alpha}(p,0;\Omega)$. In addition, we may be concerned only with the integrated deviation of the kink waveform,

$$\Delta\bar{\phi}(\Omega) = \int_{-\infty}^{+\infty} dz \left[\Delta\phi(z,\Omega) - \lim_{|z|\to\infty} \Delta\hat{\phi}(z,\Omega)\right], \tag{3.33}$$

since this corresponds to the total induced "dipole moment" of the kink (domain wall). In this situation, we have

$$\Delta\bar{\phi}(\Omega) = \hat{\alpha}(\Omega)\hat{E}(\Omega), \tag{3.34}$$

where $\hat{\alpha}(\Omega) = \hat{\alpha}(p = 0, q = 0;\Omega)$ is termed the dynamic polarizability of the kink. Equation (3.30) then reduces to the very simple form

$$\hat{\alpha}(\Omega) = 4(i + Q + 2Q^3)/Q^3(1 + Q^2), \tag{3.35}$$

or explicitly in terms of Ω to

$$\hat{\alpha}(\Omega) = \frac{4}{\Omega(\Omega - i\Gamma)}[2 + (\Omega^2 - 1 - i\Gamma\Omega)^{-1} + i(\Omega^2 - 1 - i\Gamma\Omega)^{-3/2}], \tag{3.36}$$

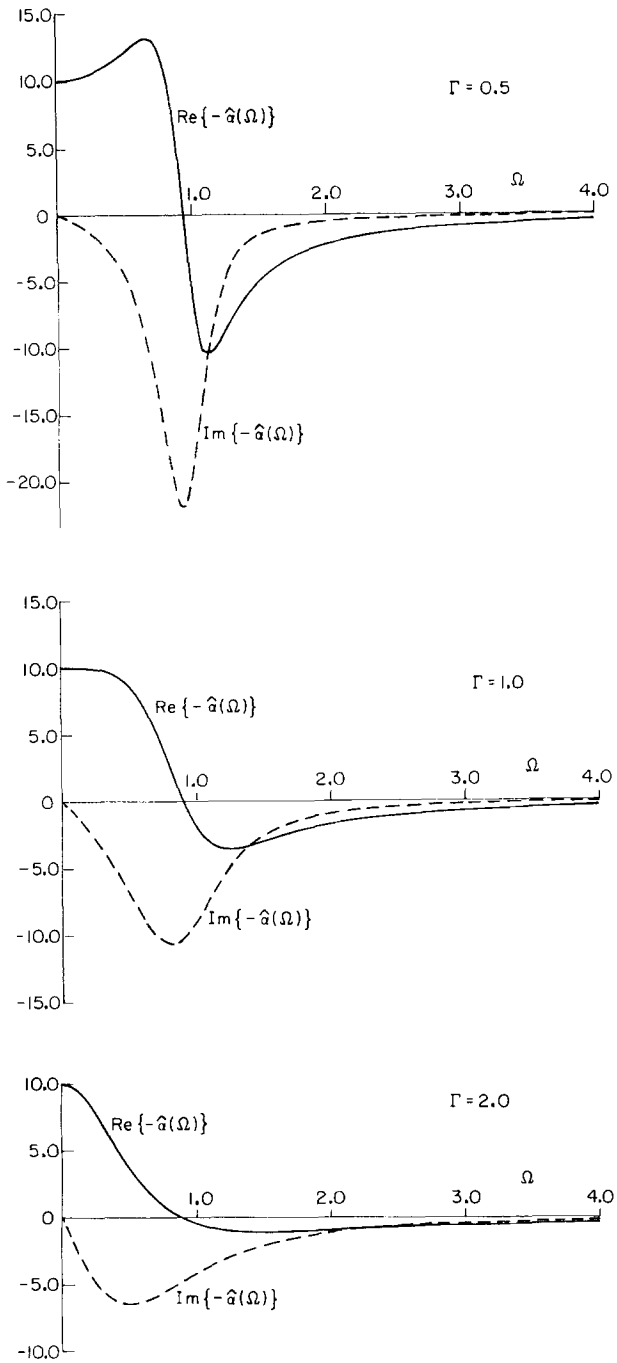

FIG. 3. Dynamic polarizability of the DQ kink vs. frequency. The real and imaginary parts of $-\hat{\alpha}(\Omega)$ are shown as solid and dashed curves, respectively, for three example values (0.5, 1.0, 2.0) of the damping constant Γ.

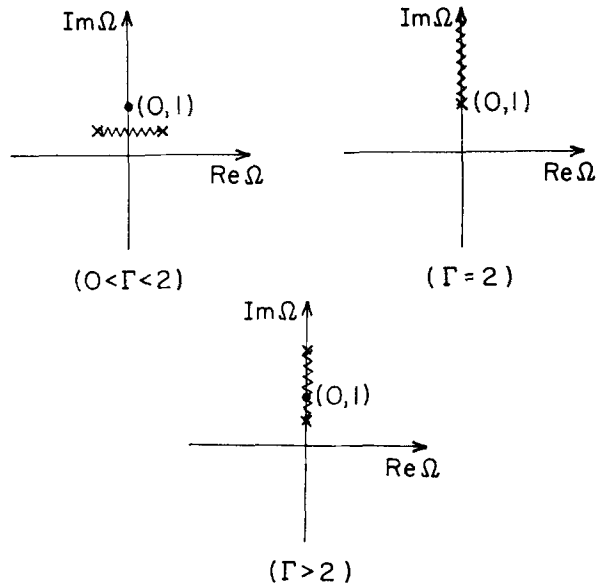

FIG. 4. Pole structure of the dynamic polarizability in the complex Ω-plane. The location of poles (which are also branch points) of $\hat{\alpha}(\Omega)$ are indicated schematically by crosses for the three regions of Γ. Also indicated are branch cuts which avoid the lower half-plane.

where it is to be remembered that the branch chosen for the square root must have a positive imaginary part. In Fig. 3 we have plotted the real and imaginary parts of $-\hat{\alpha}(\Omega)$ for three representative values of the damping constant Γ. These curves are very similar to those for a harmonic oscillator in the underdamped ($\Gamma = 0.5$), critically damped ($\Gamma = 1.0$), and overdamped ($\Gamma = 2.0$) regimes, respectively.

The analytic properties of $\hat{\alpha}(\Omega)$ in the complex Ω plane may be easily investigated with the use of Eq. (3.35). Values of Ω which give $Q = 0$ are the poles (and branch points) of $\hat{\alpha}(\Omega)$. Denoting these points by Ω_0 we have

$$\Omega_0^2 - i\Gamma\Omega_0 - 1 = 0, \tag{3.37}$$

which has the solutions

$$\Omega_0 = \begin{cases} \dfrac{i\Gamma}{2} \pm [1 - (\Gamma/2)^2]^{1/2}, & 0 < \Gamma < 2, \\ i, & \Gamma = 2, \\ i\left(\dfrac{\Gamma}{2}\right) \pm [(\Gamma/2)^2 - 1]^{1/2}, & \Gamma > 2. \end{cases} \tag{3.38}$$

Although it may appear to be so at first glance from Eq. (3.35), the points with $Q = i$ are *not* poles of $\hat{\alpha}(\Omega)$, since the numerator also vanishes at $Q = i$.

In Fig. 4 we show the poles of $\hat{\alpha}(\Omega)$ in the complex Ω plane. Possible branch cuts are also indicated. We see that for all positive values of Γ, the dynamic polarizability $\hat{\alpha}(\Omega)$ is regular in the lower half of the Ω plane, and vanishes as $\Omega \to \infty$. Thus, the Kramers–Kronig relations[21] are satisfied:

$$\mathrm{Re}\{\hat{\alpha}(\Omega)\} = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\mathrm{Im}\{\hat{\alpha}(\Omega')\}}{\Omega' - \Omega} d\Omega' \tag{3.39a}$$

$$\mathrm{Im}\{\hat{\alpha}(\Omega)\} = -\frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\mathrm{Re}\{\hat{\alpha}(\Omega')\}}{\Omega' - \Omega} d\Omega'. \tag{3.39b}$$

IV. SUMMARY

In this paper we have discussed a simple alternate (DQ) to the ϕ^4 model for domain walls in ferroelectrics. The kink solutions of the continuumized equation of motion were discussed and the small oscillations about the kink were used in obtaining a closed-form expression for the generalized susceptibility of the kink and its long-wavelength limit, the dynamic polarizability $\hat{\alpha}(\Omega)$. This quantity provides a measure of the induced dipole moment of the DQ kink in the presence of a uniform oscillating electric field. The particularly simple

form for $\hat{\alpha}(\Omega)$ makes the DQ kink a much more attractive model than the ϕ^4 kink for studying the interaction of ferroelectric domain walls with external electric fields. It is hoped that the results presented here will prove useful to specialists in the field.

ACKNOWLEDGMENT

The author wishes to thank R.M. DeLeonardis for his assistance in obtaining the curves shown in Fig. 3.

[1]See, for example, *Solitons and Condensed Matter Physics*, edited by A.R. Bishop and T. Schneider, Springer Series in Solid State Sciences, Vol. 8 (Springer-Verlag, Berlin, 1978).
[2]J.A. Krumhansl and J.R. Schrieffer, Phys. Rev. B **11**, 3535 (1975).
[3]S. Aubry, J. Chem. Phys. **64**, 3392 (1976).
[4]T. Schneider and E. Stoll, Phys. Rev. B **13**, 1216 (1976).
[5]W. Hasenfratz and R. Klein, Physica A **89**, 191 (1977).
[6]Y. Wada and J.R. Schrieffer, Phys. Rev. B **18**, 3897 (1978).
[7]N. Theodorakopoulos, S. Hanna, and R. Klein, in *Solitons and Condensed Matter Physics*, edited by A.R. Bishop and T. Schneider, Springer Series in Solid State Sciences, Vol. 8 (Springer-Verlag, Berlin, 1978), p. 158.
[8]K.C. Lee and S.E. Trullinger, "Generalized Susceptibility of a Solitary Wave", J. Math. Phys. (in press).
[9]J.F. Currie, J.A. Krumhansl, A.R. Bishop, and S.E. Trullinger, "Statistical Mechanics of One-Dimensional Solitary-Wave-Bearing Scalar Fields: Exact Results and Ideal-Gas Phenomenology" (to be published).
[10]M.B. Fogel, S.E. Trullinger, A.R. Bishop, and J.A. Krumhansl, Phys. Rev. Lett. **36**, 1411 (1976); Phys. Rev. B **15**, 1578 (1977).
[11]M.B. Fogel, S.E. Trullinger, and A.R. Bishop, Phys. Lett. A **59**, 81 (1976).
[12]The DQ well was first suggested by S. Aubry (private communication) as a simple kink-bearing potential, and has been used subsequently by B. Horovitz, J.A. Krumhansl, and Eytan Domany, Phys. Rev. Lett. **38**, 778 (1977), in a study of solitons in coupled-chain systems.
[13]S.E. Trullinger and R.M. DeLeonardis, "Statistical Mechanics of the Double-Quadratic Chain: Exact Results and Ideal-Gas Phenomenology for Non-Reflectionless Solitary Waves", Phys. Rev. A (in press).
[14]J.F. Currie, S.E. Trullinger, A.R. Bishop, and J.A. Krumhansl, Phys. Rev. B **15**, 5567 (1977).
[15]J.S. Zmuidzinas and S.E. Trullinger, "Effects of Discreteness on Sine-Gordon Solitons" (to be published).
[16]R. Rajaraman, Phys. Rep. C **21**, 229 (1975).
[17]R. Jackiw, Rev. Mod. Phys. **49**, 681 (1977).
[18]Julio Rubinstein, J. Math. Phys. **11**, 258 (1970).
[19]J. Goldstone and R. Jackiw, Phys. Rev. D **11**, 1486 (1975).
[20]If $E(z,\tau)$ contains a term which is constant in time, then $\Delta\phi$ is small only after the removal of secularities due to the translation of the kink. See Ref. 10.
[21]L.D. Landau and E.M. Lifshitz, *Statistical Physics* (Addison–Wesley, Reading, Massachusetts, 1969), 2nd ed., p. 390.

ERRATA

Erratum: Linearized analysis of inhomogeneous plasma equilibria: General theory [J. Math. Phys. 20, 413 (1979)]

H. Ralph Lewis
Los Alamos Scientific Laboratory, P. O. Box 1663, M. S. 642, Los Alamos, New Mexico 87545
Keith R. Symon
Department of Physics, University of Wisconsin, Madison, Wisconsin 53706

(Received 12 June 1979; accepted for publication 22 June 1979)

1) The sentence beginning in the fourth line after Eq. (II.21) should read "See the note added in proof, Ref. 12."

2) Equation (IV.23) should read

$$J^{*}_{srn'} = \int dQ\, dP \Gamma^{(c)} \mathcal{F}_s(Q,P,p_k)\delta(H_s - E_s) \times z^{*}_{n'}(Q)u_{sr}(Q,P)\big|_{p_k = p_k^0}. \quad \text{(IV.23)}$$

3) Equation (IV.24) should read

$$\langle n' | D | n \rangle = \lambda_n \delta_{nn'} - \delta_{\kappa\kappa'} \sum_{s(r)} \frac{K_{sr}(\omega) J^{*}_{srn'} H'_{srn}}{\mu_{sr} - \omega}. \quad \text{(IV.24)}$$

4) Equation (IV.25) should read

$$J_{srn} = H'_{srn}. \quad \text{(IV.25)}$$

5) In the sixth line following Eq. (V.6), "amplitude and phases" should read "amplitudes and phase."

6) In the third line before Eq. (A1), $\phi(r)$ should read $\phi(\mathbf{r})$.

7) The equation following Eq. (Al2) should read

$$\delta H_s = \sum_i \delta \mathbf{r} \cdot (\nabla \phi_i) \frac{\partial H_s}{\partial \phi_i} + \sum_j \left(\delta q_j \frac{\partial H_s}{\partial q_j} + \delta p_j \frac{\partial H_s}{\partial p_j} \right).$$

8) In Ref. 2, "D.C. RD. C. Robinson" should read "D.C. Robinson."

9)In Ref. 8, "**23**, 892 (1973)." should read "B **23**, 892 (1973)."

10)In Ref. 10, "K.R." should read "K.R. Symon."

11)In the seventh line of the right-hand column of p. 423, "indeces" should read "indices."

 0022-2488/80/030599-01$0.00